



RICS

the mark of
property
professionalism
worldwide

DAUPHINE
UNIVERSITÉ PARIS

**CIB W113 Papers on
Law and Dispute Resolution**

LAW AND DISPUTE RESOLUTION

Environmental Law

- Public nuisance re-visited (John Pointing, Kingston University, UK) 7
- Surveying the sustainable and environmental legal and market challenges for real estate (Colleen Theron, LexisNexis, UK) 20
- UK climate change law: securing a low carbon economy by recourse to experts and government obligations? (Mark Stallworthy, Swansea University, UK) 31
- Understanding property rights in carbon: a methodological inquiry (Jason Prior and Spike Boydell, University of Technology Sydney, Australia) 45
- Building regulations and control in the face of climate change (Henk Visscher, Frits Meijer and Laure Itard, Delft University of Technology, The Netherlands) 62
- Managing Urbanisation and Environmental Protection in Australian Cities: Approaches for Integrating Biodiversity and Urban Growth in Sydney (Peter Williams, The University of New South Wales, Australia) 72
- Sustainable Development Legal Ownership in Danish Urban Planning Law (Helle Ina Elmer, Ministry of the Environment, Denmark) 87
- Severn Tidal Barrage Project: a legal paradox? A case study looking at the possible conflict between climate change policy and environmental legislation in the UK (Tilak Ginige, Fraser Ball and Natalie Sparks, Bournemouth University, UK, and Ann Thornton, Lindsay Carrington Ecological Services Ltd, UK) 101
- Assessing the roles of regulation, litigation, management and environmental advocacy in protecting groundwater resources in the USA (Robert W. Schultz and Donald C. Daniels, AMEC Geomatrix, USA, and Thomas Vandenburg, Dongell Lawrence Finney LLP, USA) 130
- Legal challenges on controlling carbon capture and storage and land planning (Inga-Liisa Paavola, University of Helsinki, Finland) 142
- An Integrated Product Policy Approach to Products and their Impact on Energy (Rosalind Malcolm, University of Surrey, UK) 153
- Retrofits and renewals in buildings for energy efficiency and building regulations and control in local authorities in South Africa (Joachim E Wafula and Alfred Talukhaba, University of the Witwatersrand, South Africa) 173
- The quest for sustainable homes and reduced emissions from buildings: getting it right at the planning stage (Julie Adshead, University of Salford, UK) 191

Building Regulations

- Equivalent performance in Dutch Building Regulations (Nico Scholten, Expertiscentrum Regelgeving Bouw, The Netherlands) 207
- The validity of daylight calculations in rights to light cases (Peter Defoe and Ian Frame, Anglia Ruskin University, UK) 221
- Probabilism, the way out for performance based building regulations (Nico Scholten, Expertiscentrum Regelgeving Bouw, The Netherlands) 249
- Automated compliance checking using building information models (David Greenwood, Stephen Lockley, Sagar Malsane and Jane Matthews, Northumbria University, UK) 266

- Uncertainty of building regulation for refurbishment projects in Malaysia (Azlan Shah Ali, University of Malaya, Malaysia, Adi Irfan Che Ani, Universiti Kebangsaan Malaysia, Malaysia, and Wan Zahari Wan Yusoff, Universiti Tun Hussein Onn Malaysia) 275
- Study of building contractors' compliance to building regulations in Cape Town (Bimbo Windapo and Keith Cattell, University of Cape Town, South Africa) 283

Construction Law

- The duty to co-operate in construction contracts - an international review (Evaldas Klimas, Vilnius University, Lithuania) 298
- Legal Implications of New Equipment Anti-Idling Requirements (Thomas Korman and Lonny Simonian, California Polytechnic State University, USA) 312
- When clients do not pay: a critical analysis of the legal remedies available to the South African building and civil engineering contractors and consultants (Tinus Maritz, University of Pretoria, South Africa, and Dirk Robertson, Greene Consulting Engineers, South Africa) 320
- Legal Issues Associated with Building Information Modelling (Lonny Simonian and Thomas Korman, California Polytechnic State University, USA) 341
- A New Frontier for the Construction Court (Michael Reynolds, Anglia Ruskin University, UK) 350
- Building Information Models as contract documents: common practice for the US construction industry - a preliminary report (Darren Olsen and J. Mark Taylor, Auburn University, USA) 364
- Legal issues of partnering in construction: an analysis of case law (Faraziera Mohd Raslim, Nur Emma Mustaffa and Jamaluddin Yaakob, Universiti Teknologi Malaysia, Malaysia) 374
- Dispute Resolution: A Preliminary Report. Recent Changes to the AIA Documents (J. Mark Taylor, Auburn University, USA) 388
- A conceptual model for effective maintenance of productivity records: a tool for facilitating the analysis of disruption claims (Nuhu Braimah, Brunel University, UK) 401
- Investigating the relationship between construction contract documentation incompleteness and loss-and-expense claims: practical considerations (Florence Walker and Stephen Pryke, University College London, UK) 414
- Is there a need for public private partnership projects in Singapore? (Asanga Gunawansa, National University of Singapore) 440
- BIM Implementation and the Law: An International Literature Review (Brodie McAdam, University of Salford, UK) 459
- Defect Liability Period: Employer's Right and Contractor's Liabilities Examined (Ayodeji Sunday Ojo, Ministry of National Development, Seychelles) 467
- Standard form of contract for decoration, repair and maintenance works for use by laymen (Paul H.K. Ho, City University of Hong Kong) 482
- Inhibiting and enhancing impingements on second year undergraduate construction law education (Tim McLernon and Robert Eadie, University of Ulster, UK) 495

Alternative Dispute Resolution

- Dispute resolution alternatives: problems, preference and process (Carl van Zyl, Basie Verster and Stephan Ramabodu, University of the Free State, South Africa) 513
- A critical analysis of the potential application of dispute resolution boards in the Chilean construction industry (Alfredo Serpell and Maria Alejandra Medina Oyuela, Pontificia Universidad Católica de Chile, Chile) 527
- An investigation of construction lawyer attitudes to the use of mediation in Scotland (Andrew Agapiou and Bryan Clark, Strathclyde University, UK) 544
- Mediators: towards a code of good practice for construction mediators (Penny Brooker, University of Wolverhampton, UK) 560
- Opinions of legal professionals regarding the selection of appropriate resolution method in construction disputes (Deniz Ilter, Istanbul Technical University, Turkey) 587
- The development of statutory adjudication in the UK, 1998 to the present day (Peter Kennedy, Lisa Cattnach and Edward McLuskey, Glasgow Caledonian University, UK, and Janey Milligan, Construction Dispute Resolution Ltd, UK) 595
- A proposal for a 'Dual Scheme' of statutory adjudication for the building and construction industry in Australia (Michael Brand, The University of New South Wales, Australia, and Philip Davenport, Solicitor of the Supreme Court of New South Wales, Australia) 609

International Law

- Comparative analysis of aspects of damages; consequential loss and classification (Maree Chetwin, University of Canterbury, New Zealand) 626
- The capitalization rates for the valuation of the land uses in the Spanish law context (Vicente Caballer, Natividad Guadalajara and Elena De la Poza, Universidad Politécnica de Valencia, Spain) 638
- International contracts and the choice of law in New Zealand (Nicky Richardson, University of Canterbury, New Zealand) 651
- The legal consequences of breach of government undertakings to investors (Asanga Gunawansa, National University of Singapore) 664

Housing Law

- Children's citizenship: the dilemma of housing and property law (Simon Hoffman, Swansea University, UK) 679
- The housing policies for remodeling and renovation in Spain (María Luisa Gómez Jiménez, University of Málaga, Spain) 690
- Balancing housing need and prevention of homelessness in England (Patricia Ng) 702

Professional Ethics

- Black Economic Empowerment in South Africa: an update (Charles Nsibandé, Bogoje Properties, South Africa, and Elizabeth Pienaar, University of the Witwatersrand, South Africa) 720

- Who owns the copyright of architectural works and designs? (Phebe Mann, University of Reading, UK) 733
- Identity as a manifestation of organisational culture (Sai On Cheung, Hoi Yan Pang, So Yan Tam and Hoi Yan Chan, City University of Hong Kong) 748
- A single ethical code for the construction industry: what do young professionals think? (Jim Mason, University of the West of England, UK) 762
- Constructing a Professional Credentialing Scheme for the Construction Industry (Jenny Casmus and Melissa Daigneault, Texas A&M University, USA) 775

Planning Law

- Spatial planning to the test of property rights in Switzerland: an innovative land management approach to coordinate spatial planning goals with property rights interests (Philippe Weber, University of Lausanne, Switzerland) 787
- Implications of Commons Act 2006 for the registration of land as a town or village green (Phebe Mann, University of Reading, UK) 802
- The contemporary urban commons - a case study of Darling Harbour, Sydney (Spike Boydell, University of Technology Sydney, Australia, and Glen Searle, University of Queensland, Australia) 809
- Compensation Rights for Land Use Regulation: The Findings of Cross-national Comparative Research (Rachelle Alterman, Technion - Israel Institute of Technology) 827

Doctoral Session

- Key note address: New horizons in teaching law (Michael Reynolds, Anglia Ruskin University) 842
- The proposed directive on industrial emissions: what will change regarding the best available technique (BAT), the Finnish perspective (Eeva-Maija Puheloinen, University of Helsinki, Finland) 850
- Minority protection doctrines: from company law and equity to strata title (Rachel Leow, National University of Singapore) 860
- Land registration and information asymmetry problem in sub-Saharan Africa urban real estate markets (Stanislaus Adiaba, Felix Hammond, David Proverbs, Jessica Lamond, Colin Booth and Robin Bloch, University of Wolverhampton, UK) 877
- Conceptualising the Singapore Real Estate Investment Trust (Seah Xiande Aaron, National University of Singapore) 894
- Does the Planning Act 2008 Enshrine and Heighten the Statutory Immunity Defence? (Francis Moor, University of Salford, UK) 930
- A principled response to pre-contractual remuneration (Kelvin Koh) 953

Property Law

- The function and efficiency of anticipated recording in China's Real Estate Transactions (Li Bin, Harbin Institute of Technology, China) 973

• Where will European Property Law go next? Tenancies and Property after Hutten-Czapka v Poland and FEANTSA v Slovenia (Jane Ball, University of Sheffield, UK)	988
• The challenges of integrating the carbon reduction scheme into commercial leases (Susan Bright, New College, University of Oxford, UK)	1006
• A commercial right to buy? (Jill Morgan, Swansea University, UK)	1020
• Judges, child trespassers and occupiers' liability in the built environment (Luke Bennett, Sheffield Hallam University, UK)	1032
• Maps, land and land title - are they synonymous? (Carl Calvert, University College London, UK, and Pierre Clergeot and Nicolas Chauvin, l'Ecole Supérieure des Géomètres et Topographes, France)	1065
• Constructing the notion of market value of a property in the context of liability of valuer (Stanislawa Kalus, University of Silesia in Katowice, Poland)	1076
• Property loss due to coastal erosion: judicial, legislative and policy interventions (Mick Strack, University of Otago, New Zealand)	1093
• Adverse possession of boundary land: lessons from abroad (Una Woods, University of Limerick, Ireland)	1105
• Assessment on problems of issuances and transferences of strata title in Malaysia: the perception of the owner (S.N Kamaruzzaman, Naziah Muhamad Salleh and S.A.F AlZawawi, University of Malaya, Malaysia)	1122

Public nuisance re-visited?

John Pointing

Barrister and Senior Lecturer in Property Law, School of Surveying & Planning, Kingston University.

Abstract:

Recent decisions involving environmental disasters, such as the Buncefield refinery explosion and the failure of Corby Borough Council to remediate the disused Corby steelworks site, have suggested that the orthodox position on public nuisance is no longer sustainable. Instead of conceptualising it as a property tort operating on a larger scale than private nuisance, public nuisances may also be seen as a separate, personal rights based tort. The flexibility of public nuisance has been apparent over several centuries and this continues to make it a useful cause in a civil action involving threats to the life, health and safety of the public as well as an environmental offence for egregious crimes.

Keywords: common law, environmental protection, public nuisance, tort

Introduction

Since time immemorial the boundaries of nuisance have disturbed tidy-minded lawyers and jurists. Often, confusion results from the assumption that a public nuisance is merely a larger-scale form of private nuisance affecting a large (but unquantified) section of the public. On the other hand, public nuisance can be seen as an environmental tort (and offence), as the last ditch protecting the right not to be adversely affected by the unlawful act or omission whose effect was ‘to endanger the life, health, property, or comfort of the public’ (Archbold, 2010: 31-40). This conceptualisation posits a quite separate tort from an unlawful interference in a proprietary right, germane to private nuisance.

The tension between property-based and rights-based forms of public nuisance has been brought into focus by a number of important cases, culminating in the Court of Appeal’s decision in respect of the failure of Corby District Council to properly remediate the site of

the old Corby steelworks.¹ An important implication of *Corby* is that it strengthens the position of public nuisance as a separate, rights-based tort from private nuisance: one that does not depend on the victim having a proprietary interest in land. Dyson LJ quoted with approval the position in the United States, where: ‘Unlike a private nuisance, a public nuisance does not necessarily involve interference with use and enjoyment of land’.² He added:

...it does not follow that the right which is interfered with in a public nuisance case is properly to be regarded as a right to enjoy property. The essence of the right that is protected by the tort of private nuisance is the right to enjoy one's property.... The essence of the right that is protected by the crime and tort of public nuisance is the right not to be adversely affected by an unlawful act or omission whose effect is to endanger the life, safety, health etc of the public.³

For a very long time the seminal work on the boundaries of nuisance has been Professor Newark's piece published over 60 years ago in the *Law Quarterly Review* (Newark, 1949). This article is very appealing to traditional property lawyers, who prefer their hedges neat, their boundaries tightly drawn. Professor Newark's position is simple: the proper place for nuisance is that it is a tort to land, or one directed against the plaintiff's enjoyment of rights over land, including interference with an easement or profit (Newark, 1949: 482). Any deviation from this position creates uncertainty and conceptual turmoil amounting to heresy. The Court of Appeal in *Corby*, in turning away from such orthodoxy, rejected Professor Newark's monistic formulation, finding that public and private nuisance should be seen as separate torts protecting different rights.⁴

There is some evidence that straying from orthodoxy produces perplexing and confused results, so vindicating Professor Newark's anxieties. For example, a government department has managed recently to confuse itself so completely that in its guidance to local authorities on implementing the Licensing Act 2003 it proclaims that a public nuisance, such as noise emanating from licensed premises, includes: ‘low-level nuisance perhaps affecting a few

¹ *Corby Group Litigation v Corby Borough Council* [2008] EWCA Civ 463.

² American Law Institute, *Restatement of the Law, Second, Torts 2d* (1979) chapter 40 para 821B (h).

³ *Corby Group Litigation* (n 1 above) at [29].

⁴ *Corby Group Litigation* (n 1 above) per Dyson LJ at [27-30].

people living locally' (DCMS, June 2007: para 2.33). This matter will be considered further towards the end of this paper.

Judges who stray beyond the proper boundaries are brought to order

The prim and proper monistic conception of nuisance becomes adulterated once the idea is admitted that personal injuries could be compensated in an action for public nuisance. The basic proposition placed before the Court of Appeal as a preliminary issue in the *Corby* litigation by the defendant local authority was that compensation for personal injuries should not be available to the victims of a public nuisance.⁵ In the final action heard in the High Court, Corby District Council were found liable in negligence, for breach of statutory duty and in public nuisance.⁶ The local authority were held responsible for birth defects caused from the exposure of 18 mothers to toxic chemicals released because of the inadequate remediation of the old Corby steelworks which they had supervised.

The earlier, Court of Appeal decision in *Corby* had been on a preliminary issue about whether the import of two recent House of Lords decisions in *Canary Wharf*⁷ and *Transco*⁸ was such that earlier cases awarding compensation for personal injuries in public nuisance had been wrongly decided. But the overarching authority employed by the local authority in *Corby* at the Court of Appeal was none other than Professor Newark, or rather, his venerable text. So the question arises of whether *The Boundaries of Nuisance* truly deserves the sacerdotal qualities heaped upon it?

Professor Newark places the blame for setting 'the law of nuisance on the wrong track' on Mr Justice Fitzherbert, for it was he who 'sent subsequent generations wrong in their law' in deciding that an action could be maintained for damages for personal injuries caused by an obstruction to the highway (Newark, 1949: 483-4). Thus it was his Lordship who, in 1535, during the reign of Henry VIII, uttered the words unlinking nuisance as an exclusive tort to land and allowing the possibility of claiming an award of special damages for the consequences of an obstruction to the highway (Spencer, 1989: 74). No doubt Fitzherbert CJ was doing what all good common law judges do when trying to find a justification for providing a remedy to a meritorious claim: he strained the law.

⁵ Ibid at [10].

⁶ *Corby Group Litigation v Corby District Council* [2009] EWHC 1944 (TCC).

⁷ *Hunter v Canary Wharf* [1997] AC 655.

⁸ *Transco plc v Stockport MBC* [2003] UKHL 61; [2004] 2 AC 1.

As if a man make a trench across the highway, and I come riding that way by night, and I and my horse together fall in the trench so that I have great damage and inconvenience in that, I shall have an action against him who made the trench across the road because I am more damaged than any other man.⁹

For Lord Cooke, giving his minority judgement in *Canary Wharf*, 'if this was indeed an indiscretion on Fitzherbert's part, to rue it now might seem a little late'.¹⁰ Quite so.

Over the centuries, the main focus of nuisance cases has been on incompatible uses of land between neighbours. These cases have primarily been to do with private nuisance: with interferences in the use or enjoyment of land. Some have been concerned with physical damage to the land; relatively few cases have dealt with public nuisance as their central point. The House of Lords in *Canary Wharf* had little to say about public nuisance. In respect of the interference in television signals, Lord Cooke found 'no material difference... between public and private nuisance'.¹¹ This is surprising given the proportion of Her Majesty's subjects who had been materially affected by the dust nuisance and from interference in their reception of television signals caused by the construction of the Canary Wharf tower. But the position of the majority of the House of Lords in *Canary Wharf* represents the high point of monistic orthodoxy, in which public nuisance is eschewed as an independent tort and private nuisance firmly put in its place as a property-based tort (Kidner, 1998; Wightman, 1998).

Public nuisance in the 21st Century

Civil actions in private and in public nuisance have been available for a long time. Public nuisance is also a crime; an either-way offence for which an unlimited fine and/ or a maximum sentence of life imprisonment are available on conviction in the Crown Court.¹² The crime of public nuisance is defined in very broad terms in *Archbold* (2010: 31-40) as follows.

A person is guilty of a public nuisance (also known as common nuisance), who (a) does an act not warranted by law, or (b) omits to discharge a legal duty, if the effect of the act or omission is to endanger the life, health, property, or comfort of the public,

⁹ Y.B. 27 Hen. 8, Mich. pl. 10.

¹⁰ *Hunter v Canary Wharf* (n 7 above) at 718.

¹¹ *Ibid*, at 722.

¹² In one of the most serious public nuisances in modern times, a sentence of 17 years' imprisonment was upheld by the Court of Appeal in *R v Bourgass* [2007] 2 Cr App R (S) 40. The convict had been the prime mover in a conspiracy to commit acts of terrorism (but charged as a conspiracy to cause a public nuisance) involving the use of poisons and explosives intended to destabilise the community by causing disruption, fear and injury.

or to obstruct the public in the exercise or enjoyment of rights common to all Her Majesty's subjects.

This definition was accepted by Lord Bingham in *R v Rimmington* as 'clear, precise, adequately defined and based on a discernible rational principle', a definition which applies as much to the tort of public nuisance as to the crime.¹³

Civil actions in public and private nuisance are not mutually exclusive. In determining the preliminary issues arising from the Buncefield oil storage explosion, the High Court rejected a monistic view of nuisance:

It is accordingly difficult to discern any difficulty in categorizing the incident at Buncefield as a public nuisance ... A very large number of people were affected. Those who had an interest in land suffered private nuisance. The explosion endangered the health and comfort of the public at large. Subject to establishing a loss which was particular, substantial and direct ... there is a claim in public nuisance.¹⁴

Environmental forms of public nuisance will include those caused by smells, noise, waste deposits and water pollution, but there is no exhaustive list for this common law form of nuisance. Examples include: quarry-blasting;¹⁵ emission of noxious smells from a chicken-processing factory;¹⁶ storage of large amounts of inflammable material;¹⁷ allowing refuse and filth to be deposited on vacant land in a densely populated part of London;¹⁸ holding an all-night 'rave' in a field;¹⁹ holding noisy events, such as motocross.²⁰

In the civil action of *Attorney-General v PYA Quarries Ltd*, Denning LJ²¹ decided that the basic requirement for a public nuisance is for it to be so:

¹³ *R v Rimmington; R v Goldstein* [2006] 1 AC 459; [2005] UKHL 63 at [7, 36]. Lord Bingham's speech is a comprehensive analysis of the law of public nuisance, and his words on the rationality of the principle have great authority.

¹⁴ *Colour Quest Ltd and Others v Total Downstream UK Plc and Others (Rev 1)* [2009] EWHC 540 (Comm) at [434], per David Steel J; reversed in part (but not on the public nuisance parts) by *Colour Quest Ltd v Total Downstream UK Plc* [2010] EWCA Civ 180.

¹⁵ *A-G v PYA Quarries Ltd* [1957] 2 QB 169.

¹⁶ *Shoreham-by-Sea UDC v Dolphin Canadian Proteins* (1972) 71 LGR 261.

¹⁷ *R v Lister and Biggs* (1857) 26 LJMC 196.

¹⁸ *A-G v Tod Heatley* [1897] 1 Ch 560.

¹⁹ *R v Shorrock* [1993] 3 All ER 917. In this case the defendant farmer was found guilty of public nuisance for a 'rave' organised by another person but taking place on the farmer's land whilst he was elsewhere. Mr Shorrock was found guilty on the basis that he knew or ought to have known that there was a real risk of the event causing a public nuisance.

²⁰ *East Dorset DC v Eaglebeam Ltd* [2006] EWHC 2378 (QB).

²¹ Per Denning LJ in *A-G v PYA Quarries Ltd* [1957] 2 QB 169, at 190–191. The applicability of this passage to criminal proceedings in public nuisance was doubted by Lord Rodger in *R v Rimmington; R v Goldstein* [2005] UKHL 63 at [44].

widespread in its range or so indiscriminate in its effect that it would not be reasonable to expect one person to take proceedings on his own responsibility to put a stop to it, but that it should be taken on the responsibility of the community at large. So if a class of people or a neighbourhood suffer sufficiently from an unreasonable interference in their use of land or from physical damage to their property, then a public nuisance action could be brought by the local authority. A person could sue those responsible²² in a relator action brought in the name of the Attorney-General if he can prove that he suffered special damage over and above that of the community at large.²³

A prosecution for public nuisance may be appropriate for egregious environmental crimes. In the major oil pollution incident caused by the grounding of the *Sea Empress*, the Milford Haven Port Authority pleaded guilty in the Crown Court to a strict liability offence under section 85(1) of the Water Resources Act 1991. They received a fine of £4 million, reduced on appeal to £750,000 by the Court of Appeal. The conviction was the result of a plea bargain in which the Environment Agency dropped the public nuisance charge on the indictment in exchange for a guilty plea for the lesser charge. The late Michael Hill QC, who prosecuted Milford Haven Port Authority on behalf of the Environment Agency, conceded (during a conference held by the United Kingdom Environmental Law Association) several years later that dropping the more serious charge may have been a mistake.²⁴

Unlike a person, who can sue only by way of a relator action in the name of the Attorney-General, a local authority can bring an action in public nuisance in its own name under section 222 of the Local Government Act 1972. The local authority may do so where it ‘considers it expedient for the promotion or protection of the interests of the inhabitants of its area’. In *Railtrack plc v Wandsworth LBC*²⁵ the local authority sought an injunction and a declaration that the company should be held responsible for the costs of cleaning the pavements from the faecal deposits of pigeons roosting in the girders underneath the railway

²² The person responsible could be a local authority, as in the case of *Corby Group Litigation v Corby DC* [2009] EWHC 1944 (TCC), where the council was found liable for public nuisance in causing, allowing or permitting the dispersal of dangerous or noxious contaminants that resulted in birth defects.

²³ Establishing what comprises special damage is problematic. In the Delphic words of Mr Justice More-Bick, in *Jan de Nul v NV Royale Belge* [2002] 2 Lloyd's Rep 700 at 715: ‘In the end the question whether the plaintiff's injury is sufficiently “special” and “direct” must depend very much on the facts of the case.’

²⁴ The Crown Court prosecution is reported in *Environment Agency v Milford Haven Port Authority And Andrews (The “Sea Empress”)* [1999] 1 Lloyd's Rep 673.

²⁵ [2001] EWCA Civ 1236; [2002] QB 756.

bridge of a south London railway station. The council succeeded in obtaining a declaration and the court found the company liable in public nuisance for damage caused by the pigeons.

A common element is necessary for a public nuisance to be proven, as distinct from a series of nuisances affecting different individuals as separate instances. A common element may be present where there is a sufficiently large number of individual private nuisances and where the offence affects individual victims simultaneously. In *PYA Quarries Ltd*, Romer LJ²⁶ opined:

Some public nuisances (for example, the pollution of rivers) can often be established without the necessity of calling a number of individual complainants as witnesses. In general, however, a public nuisance is proved by the cumulative effect which it is shown to have had on the people living within its sphere of influence. In other words, a normal and legitimate way of proving a public nuisance is to prove a sufficiently large collection of private nuisances.

The courts have given no precise guidance as to what might constitute a sufficient number of individuals. The Court of Appeal in *PYA Quarries Ltd*, somewhat quaintly, referred to ‘a class of Her Majesty’s subjects’, but did not place a number on the persons needing to be affected by the quarrying activity to bring it within the scope of a public nuisance.²⁷ Romer LJ did say that it was not necessary to prove that every person in the locality of the nuisance needed to be affected, adding that for an injunction to be granted in public nuisance a ‘representative cross-section of the class’ would have to be affected.²⁸ The use of the word ‘representative’ in this context is curious since it implies a sample; perhaps his Lordship meant it to mean a significant proportion of the total number that could be affected by the nuisance.

A series of separate acts affecting individual members of the community would not constitute the necessary common element required to establish a public nuisance. In the criminal case of *R v Rimmington*, in which the defendant was accused of distributing hate mail to a number of individuals, Lord Rodger found that ‘a core element of the issue of public nuisance is that the defendant’s act should affect the community, a section of the public,

²⁶ *A-G v PYA Quarries Ltd* [1957] 2 QB 169, 187.

²⁷ *Ibid*, 184.

²⁸ *Ibid*.

rather than simply individuals'.²⁹ Consequently, the prosecution for a public nuisance failed as each of the letters constituted a separate act of vile racism.

There is another, crucial way of distinguishing public from private nuisances besides the 'common element' requirement. The Court of Appeal in *Corby*³⁰ rejected the formulation of public nuisance as needing to be based on an interference with rights to enjoy land in favour of the right not to be adversely affected by the unlawful act or omission whose effect was to endanger the life, health or safety of the public. The latter formulation is thus akin to a personal rather than a proprietary right. The Court of Appeal's decision in *Corby* relied heavily on that of the House of Lords in *R v Rimmington* and *R v Goldstein*.³¹ The use of a non-proprietary rights-based formulation in these decisions could be seen to give a new lease of life to public nuisance both as a crime and as a tort. *Corby* also provides further justification for allowing recovery of damages for persons whose health, life and safety have been particularly endangered by the unlawful act or omission in question.

The crucial case for distinguishing public nuisance from private nuisance and for defining it as a separate, personal rights-based tort sufficiently affecting the public with a common element, is the High Court decision in *Colour Quest*.³² Although concerned with preliminary issues in the Buncefield litigation, the consideration of public nuisance is given a very full treatment which has not been subject to challenge at the Court of Appeal.³³ In the High Court, David Steel J came to the following conclusion:

No suggestion emerges from the authorities that, where a sufficient body of the public has been subjected to the nuisance, the only claim lies in public nuisance and any claim in private nuisance is barred or vice versa:

- a) Private nuisance involves interference with someone's private right to enjoy his own land. Public nuisance involves the endangering of the health, comfort or property of the public.
- b) It follows that a collection of private nuisances can constitute a public nuisance: but it does not follow either that in consequence the claim in private nuisance is subsumed

²⁹ [2005] UKHL 63 at [47]. After *Hunter v Canary Wharf* [1997] AC 655, in which the House of Lords overruled the Court of Appeal decision in *Khorasandjian v Bush* [1993] QB 727, it would not be arguable to claim that the sending of race hate letters could constitute a private nuisance, since this act would not be a tort against a property right.

³⁰ *Corby Group Litigation v Corby Borough Council* [2008] EWCA Civ 463.

³¹ [2005] UKHL 63.

³² *Colour Quest Ltd and Others v Total Downstream UK Plc and Others (Rev 1)* [2009] EWHC 540 (Comm).

³³ *Colour Quest Ltd v Total Downstream UK Plc* [2010] EWCA Civ 180.

or that a public nuisance involving interference with health or comfort cannot be freestanding.³⁴

The basis for his Lordship's argument is that public and private nuisances should be seen as separate torts, protecting different categories of rights. But where there is a common injury, a private nuisance can become a public nuisance. Just before the passage quoted above, his Lordship had found that:

a private owner's right to the enjoyment of his own land is not a right enjoyed by him in common with other members of the public, nonetheless any illegitimate interference, being the very same interference contemporaneously suffered by other members of the public, constitutes a common injury satisfying the public nature of a public nuisance.³⁵

Whilst this may be a correct statement of the law, some will find this position defies logic. At the very least it is unfortunate to be stuck with such a fuzzy boundary between a private and public nuisance.

Noise and the Licensing Act 2003

Public nuisances can result from the behaviour of crowds. Even before the Industrial Revolution, the owner of premises attracting a motley crowd largely drawn from the 'ragged, criminal and dangerous classes' could be liable in nuisance (Himmelfarb, 1984: ch 15). In *Betterton*, Holt CJ found that:

playhouses are not in their own nature nuisances; but only as they draw together great numbers of people and coaches, and sharpers³⁶ thither, which prove generally inconvenient to the places adjacent.³⁷

In *R v Moore*, the defendant used his premises for the lawful activity of pigeon-shooting, which attracted great crowds. In finding against him Lord Tenterden opined:

The defendant asks us to allow him to make a profit to the annoyance of all his neighbours; if not, it is said we shall strain the law against him. If a person collects together a crowd of people to the annoyance of his neighbours, that is a nuisance for which he is answerable. And this is an old principle.³⁸

³⁴ *Colour Quest Ltd and Others v Total Downstream UK Plc and Others (Rev 1)* [2009] EWHC 540 (Comm) at [432].

³⁵ *Ibid.*, at [430].

³⁶ The *Oxford English Dictionary* defines a 'sharpener' as 'a cheat, swindler, rogue; one who lives by his wits and by taking advantage of others; esp. a fraudulent gamester'. Today's equivalent might be a merchant banker or a market trader.

³⁷ *Betterton's Case* (1680) Holt 538.

³⁸ *R v Moore* (1832) 3 B & Ad 184; 110 ER 68.

Such cases illustrate the existence of an established line of authority in which public nuisance succeeded as a cause of action to control noise and the associated problems of crowds where the problem existed outside the defendant's premises. It was not a requirement that the offending activity be confined to what took place within the premises as the problem was compounded on the surrounding land. This line of authority fell into abeyance by the end of the nineteenth century, but is relevant to the noise implications of crowds in and around public houses as this comes within the scope of the Licensing Act 2003.

The Licensing Act 2003 requires licensing authorities to form a judgment about what constitutes public nuisance and decide what is necessary to avoid it by attaching conditions to premises licences and club premises certificates.³⁹ Avoidance of public nuisances – such as from noise, light pollution, noxious smells and litter - is therefore a licensing objective. Licensing and responsible authorities are also required to assess the impact of the licensable activity on persons living and working in the vicinity when coming to conclusions on licensing conditions (DCMS, June 2007).⁴⁰

Local authorities and the police are the front-line authorities for making judgements about whether a state of affairs might amount to a public nuisance. Unfortunately, neither of these entities has much experience in bringing actions in public nuisance. Local authority environmental health departments utilize their statutory powers under section 80 of the Environmental Protection Act 1990, but these apply in respect of statutory nuisance which has its own procedure. The modern police force is placed in the invidious position of having very little (if any) experience or expertise with respect to the prosecution of environmental public nuisances.

The problem of controlling public nuisance for regulators is further compounded because changes to the licensing system came into effect less than two years before the ban on smoking was imposed. The application of the Health Act 2006 to ban smoking in all enclosed work places and public spaces has included pubs, restaurants, members' clubs and entertainment venues. It has resulted in increased noise disturbance from nicotine addicts and

³⁹ Licensing Act 2003, s 4.

⁴⁰ Licensable activities includes the provision of regulated entertainment as well as the provision of alcohol. Schedule 1 to the Licensing Act 2003 sets down types of regulated entertainment.

their associates congregating outside premises where licensable activities take place. Government advice is that licensing conditions to control a public nuisance cannot be imposed where the problem is caused beyond the vicinity of the premises (DCMS, June 2007: para 2.39). However, it might be argued that where a disturbance amounts to a public nuisance, although manifested beyond the vicinity of the regulated premises, it would not have occurred but for the licensable activity taking place on such premises. There may be some grounds for concluding, on the basis of the line of authority from *Betterton's Case*,⁴¹ that the scope of liability in public nuisance is wider than the guidance allows.

Advice from government has been supplied munificently. The Department of Culture, Media & Sport (DCMS) has tried to smooth the implementation of the Licensing Act 2003 and reconcile it with the object of the Anti-social Behaviour Act 2003. Buried within its copious guidance for local authorities is the following advice:

It is important to remember that the prevention of public nuisance *could ...include low-level nuisance perhaps affecting a few people living locally* as well as major disturbance affecting the whole community (my emphasis) (DCMS, June 2007: para 2.33).

The authors of the guidance fail to show how their advice might be reconciled with *A-G v PYA Quarries*,⁴² and it would appear to be advice given without consideration of any kind of case law authority. It probably reflects policy objectives based on a very loose interpretation of the scope of the Anti-social Behaviour Act 2003, or, perhaps, a manful and retrospective attempt by the government to try and make the Licensing Act 2003 more effective than Parliament had originally intended.⁴³

Concluding comments

When giving judgment in *Hunter v Canary Wharf*, Lord Cooke commented on the principle of ‘give and take’ - central to a proprietorial formulation of nuisance but vague – that:

The principle may not always conduce to tidiness, but tidiness has not had a high priority in the history of the common law. What has made the law of nuisance a

⁴¹ (1680) Holt 538.

⁴² [1957] 2 QB 169.

⁴³ The DCMS guidance is inconsistent with that provided by DEFRA, which states, correctly, in its *Guidance to Local Authorities The Noise Act 1996 as amended by the Anti-social Behaviour Act 2003 and the Clean Neighbourhoods and Environment Act 2005: Guidance to Local Authorities in England* (Defra: March 2008), at [143]: ‘Although “public nuisance” is not defined in statute for the purposes of the Licensing Act 2003 or the Anti-social Behaviour Act 2003, it is generally taken by the courts to be a nuisance which affects the public at large where it would not be reasonable to expect an individual to take proceedings to resolve the matter...’

potent instrument of justice throughout the common law world has been largely its flexibility and versatility.⁴⁴

Tension between a property-based and a rights-based conceptualisation of public nuisance could be interpreted, therefore, as a feature of the common law in which logic is subordinate to justice.

Perhaps this tension cannot be resolved without recourse to a ‘shapely code’ (Radzinowicz, 1985, para. 6). Some may conclude that the guidance provided to local authorities by the DCMS in June 2007 on public nuisance constitutes the kind of error that Professor Newark warned us about long ago in *The Boundaries of Nuisance* (Newark, 1949). The guidance appears to suggest that noise amounting to no more than a private nuisance suffered by a few residents would be enough to trigger a licensing review or result in the closure of premises, or even that a low level interference not amounting to a nuisance in law could do so. This is misleading and remains a source of confusion for enforcers, licensees and the public generally. It could be that the position taken on public nuisance is a simple mistake committed by government advisers having little expertise in nuisance law. The problem with this interpretation is that it implies that the large number of individual civil servants consulted beforehand about the precise wording to be used in the guidance would each have had to have made the same error. A policy to make the guidance appear more relevant and enforceable than it really could be in controlling the British binge-drinking culture is perhaps a more plausible explanation.

But tension in the law of public nuisance would seem to have played a part and subsequent case law has not helped to resolve the difficulties. In a hastily constituted application for leave for judicial review of a licensing case concerning the fortunes of *The Endurance*, a public house in Soho, Burton J did not find the guidance defective though invited to do so by counsel for the company, who had argued the point fully and effectively. Somewhat lamely, his Lordship decided that the guidance ‘was not unlawful’.⁴⁵ In this case the noise generated by the licensed premises had been found by the district judge at first instance to be a public nuisance, having been ‘higher on the scale than something that fell within the category of simply a private nuisance’. Burton J accepted this finding, though without questioning what

⁴⁴ *Hunter v Canary Wharf* [1997] AC 655, 711.

⁴⁵ *R (on the application of Hope and Glory Public House Ltd) v City of Westminster Magistrates’ Court* [2009] EWHC 1996 (Admin) at [64].

sort of scale the learned district judge had had in mind, and so his comments as regards the lawfulness of the guidance were *obiter*.⁴⁶ Seen in another light, his Lordship may have been doing what all good common law judges do: he strained the law to suit the justice of the case. Let us hope that Mr Justice Burton does not follow Mr Justice Fitzherbert in being reviled for straying beyond the boundaries of nuisance for the next 450 years or so.⁴⁷

References:

- Archbold (2010) *Criminal Pleading, Evidence and Practice*. London: Sweet & Maxwell
- DCMS (June 2007) *Guidance issued under section 182 of the Licensing Act 2003*. London: HM Government, Department of Culture, Media & Sport
- DEFRA (March 2008) *The Noise Act 1996 as amended by the Anti-social Behaviour Act 2003 and the Clean Neighbourhoods and Environment Act 2005: Guidance to Local Authorities in England*. London: HM Government, Department for Environment, Food & Rural Affairs
- Himmelfarb G., (1984) *The Idea of Poverty*. London: Faber & Faber
- Kidner R., (1998) 'Nuisance and rights of property', *Conveyancer & Property Lawyer*, 267
- Newark F.H., (1949) 'The Boundaries of Nuisance', 65 *Law Quarterly Review*, 480
- Parpworth N., (2008) 'Public nuisance in the environmental context', *Journal of Planning & Environment Law*, 1526
- Radzinowicz, Sir Leon., (1985). *The Law Commission: Codification of the Criminal Law, a Report to the Law Commission*, 143 Law Commission.
- Spencer J. R., (1989) 'Public Nuisance – A Critical Examination', 48 *Criminal Law Journal*, 55
- Wightman J., (1998) 'Nuisance – the Environmental Tort? *Hunter v Canary Wharf* in the House of Lords', 61 *Modern Law Review*, 870

⁴⁶ Ibid.

⁴⁷ Y.B. 27 Hen. 8, Mich. pl. 10.

Surveying the sustainable and environmental legal and market challenges for real estate

Colleen Theron

LexisNexis

Colleen.Theron@lexisnexis.co.uk

Introduction

It has been estimated that the built environment in its widest sense (including construction) is responsible overall for about 40% of carbon dioxide emissions, as well as 40% of all energy used. Commercial property is a major contributor to this.

Despite the recent negative publicity challenging the scientific evidence underpinning climate change, the consensus seems to remain that climate change must be tackled and cutting carbon emissions is critical to achieving this. As a result (and notwithstanding the global economic downturn and the depressed property market), legislation and policy developments at global and national level has continued to crystallize solutions for reducing greenhouse gas emissions (GHG), including the built environment. In the US, states and local governments have embraced green real estate initiatives, ahead of federal legislation. Since 2005 there has been an increase in approved green building polices from 13 to 31.

In the UK the drive towards a low carbon economy has led to a number of policy and legislative initiatives to tackle the reduction of carbon in buildings, such as the low carbon transition plan and the consultation on zero carbon for non domestic buildings. The agenda is also been driven by directives from the European Union which member states are bound to implement. Owners and managers of existing commercial property buildings are finding themselves bound by a raft of new environmental legislation, such as energy efficiency standards for buildings; a mandatory carbon emission reduction scheme (where the organization produces electricity above a threshold of 6000kM per half hour) and revised building regulations.¹ Legislative and market pressure is also being put on the property industry to design, construct, use and demolish building in a way that both mitigates and adapts to climate change. As the impact of Climate change and resource pressures deepen, the effect on asset value will increase. The creation of the coalition government in the UK has not deterred from the theme of a 'green economy', including incentivisation for 'green growth' and 'decarbonisation' of the economy.²

This paper examines how environmental and sustainability issues are challenging 'business as usual' models in the real estate sector, outlining key mandatory and voluntary requirements for 'green buildings'.

Is sustainability 'green' ?

¹ Available at: <http://www.lexisnexis.com/Community/environmental-climatechangelaw/landing/EmergingIssues.aspx>

² Available at: <http://www.guardian.co.uk/environment/2010/may/12/coalition-environment-policy>

There is widespread and considerable interest in the topic of sustainability and sustainable construction. However, there is also a good deal of confusion over the terms used. The term 'sustainable development' is often confused with environmental protection, and many limit the scope of sustainability to consideration only of environment issues.³

The concept of 'sustainable development' is not always interpreted in the same way by different countries. The standard definition is the one contained in Brundtland's report, namely '*Sustainable development that meets the needs of the present without compromising the ability of future generations to meet their needs*'⁴

'Sustainability' as a legal term has its origins in the 'soft law' of international conventions. This means that it is too vague for breach to give rise to legal sanction.⁵ In the UK the statutory guidance issued to the Environment Agency to supplement the Environment Act 1995 adopts the 'Brundtland' definition of sustainable development.

The concept of 'sustainability' is also often used interchangeably with the concept of Corporate Social Responsibility (CSR) at corporate levels, but clarity on the differences is important for the way business behaves. The European Union started to develop the concept of CSR in 2000/2001 in line with the strategy adopted in Lisbon in 2000. The so called Lisbon objective called on the EU to become the foremost economy in the world, focusing on sustainable economic growth and greater social cohesion by 2010.⁶ The EU's sustainability strategy is also tied up with the Lisbon objective. From a business perspective CSR came first and was primarily concerned with social matters. Sustainable development emerged from the environmental protection debate. The Brundtland Commission concluded that social, ecological and economic concerns must be given equal weight. CSR tends to be restricted to ecological and social challenges and economic contributions to sustainability are not considered in detail. Whilst the concepts overlap they are applied differently.

Sustainability, in the context of the business world, is used to refer to how environmental, social and economic considerations are integrated into corporate strategy and capital markets for the long term. Companies, historically, have separated financial issues from non-financial aspects of their business. This started changing with the onset of increased pressure by stakeholders and society for more transparency about governance and the impact that companies have on their surrounding environment. Ceres⁷ states that the 'license to operate' can no longer be taken for granted by business. Companies, they suggest, cannot consider sustainability challenges in isolation. In a recent Pricewaterhouse Coopers survey of 140 chief executives of US-based multilateral companies found that 85% of them believe that sustainable development will be even more important to their business model in five years time than it is today. This should extend to real estate investors and major developers.

³ Schleich H. And Lindholm A and Flakenback H. () *Environmental sustainability-drivers for the real estate investor'*

⁴ Brundtland report (1987): <http://www.un-documents.net/wced-ocf.htm>

⁵ Keay, A.(1990), "Insolvency and Environmental Principles: A case Study in Conflict of Public Interests" *Environmental Law Review* , Vol 3, No. 2

⁶ Future E.V and Institute for Ecological Economy Research GmbH (IOW) "*Significance of the CSR debate for sustainability and the requirements for companies*". 2004

⁷ Ceres is a national coalition of investors, environmental groups and other public interested organisations working with companies to address sustainability challenges such as water scarcity and climate change. See: www.ceres.org

In the real estate sector, environmentally sustainable buildings are also referred to as 'green buildings'. Although, 'green buildings' have been used as a synonym for sustainability, this has often led to sustainability being understood from an environmental perspective, neglecting the social and economic perspectives. Despite the lack of a universally agreed definition of 'sustainable buildings', as the market evolves and as new metrics and regulation are developed and implemented, consensus may emerge. The practice of 'green building' is seen as the practice of creating structures and using processes that are environmentally responsible and resource efficient throughout a building's life cycle from siting to design, construction, operation, maintenance, renovation and deconstruction. There is an assumption that sustainable property will perform beyond the baseline of compliance and that it will offer considerable benefits over and above conventional or merely compliant property.

What are the drivers?

A major driver for sustainability in real estate sector has been the growth in environmental legislation to regulate the impact a property has on the environment through its whole life cycle. The overall increase in legislation at both international, EU and domestic level does not show signs of any abatement, which will continue to drive change. It is believed that the environmental and social aspects of sustainability will also impact on property performance.

However, there are relatively few studies that have been undertaken about the business case to promote sustainable development in the real estate sector. A RICS report (2010) published by the RICS Foundation, in conjunction with Kingston University, London issued a report entitled 'Is sustainability reflected in commercial property process: an analysis of the evidence base'⁸. The report assessed the evidence that exists on whether sustainability is reflected in commercial property prices. It found that although the last ten years has shown an increase in reports on the business case for sustainable property, very few large scale empirical studies have been undertaken. The greatest numbers of reports are US based. The conclusions in the report emphasize the need for a clear definition on sustainable buildings and what sustainability features really matter to tenants and building occupiers. The provision of meaningful benchmarks will be able to support values in the preparation of valuations. To complement this report, the RICS believes that further research is needed in relation to the financial performance of sustainable buildings.

A joint study undertaken by the University of Regensburg and Helsinki University of Technology⁹ (Schleich report) identifying the drivers of sustainability in real estate practice, confirms the limited availability of research in this area. Whilst some mention is made of the corporate drivers, the study focused on three main property led drivers, namely increased rental level, decreased property costs and decreased risks.

Increased Rental Level

⁸ Available at: http://www.rics.org/site/download_feed.aspx?fileID=5752&fileExtension=PDF

⁹ Schleich H. And Lindholm A and Flakenback H. (2009) *Environmental sustainability-drivers for the real estate investor*, eRes 2009_133

⁹ Brundtland report (1987): <http://www.un-documents.net/wced-ocf.htm>

Those studies that have analyzed the comparative data of energy costs of green buildings to other conventional buildings show reductions of energy use and savings between 6 and 30%.¹⁰

Lower Operating Costs

The Schleich report highlights that surveys conducted by Jones Lang LaSalle (2008) amongst corporate occupants globally, and a survey undertaken by Cushman and Wakefield (2007, 2008) show signs of a willingness to pay for rental premiums for sustainable space, the range of the size of the premium varying from 2-12 percent.¹¹

It is believed that a tenant's willingness to pay premiums for sustainable real estate is as a result of increased occupant productivity, potential image benefits and lower running costs.

Decreased risks

Lower operating costs are considered to be an issue promoting sustainable buildings. Empirical evidence indicates that operating costs of certified buildings are lower compared to non-certified buildings. A study in 2008 by Miller indicated lower operating costs on a sample of 243 certified and 2,000 non-certified buildings.¹²

The decreased risk associated with sustainable buildings arises from the perception of lower vacancy rates, however this is not conclusive.

Reputation

There is a corporate driver for improving reputation, by publishing CSR reports and Carbon Disclosure reports. The drive towards mandatory reporting across the globe will no doubt see an increase in more information being available about property companies. Also, as more companies embrace 'sustainability' into the way they develop their strategies and reporting, this should lead to more transparent information. Theoretically, such positioning by corporations should translate through market mechanisms into increased demand for sustainable property. Although the principles of sustainability may be embedded in the policies of some property owners and occupiers, translating them into property decisions has been difficult.

External factors

Another external driver is the issue of how much weight asset managers of institutional investors are giving to sustainability and climate related issues. A study by Jones Lang LaSalle in 2008 sought to make the claim that by investing in accredited or sustainable stock either does or may yield higher returns and other benefits.

Are real estate investors interested in energy efficiency?

¹⁰ Ibid.

¹¹ Ibid.

¹² Ibid

In a 2010 report “Energy Efficiency in Real Estate Portfolios”¹³ published by Ceres¹⁴, and Mercer¹⁵, outlined the business case for investing in energy efficiency on the basis that it enhances value in real estate portfolios.

The report aims to provide direct and indirect real estate investors with the background information, academic and industry research, case studies, key steps and best practices for integrating energy efficiency across their portfolios.

The report also states that fiduciaries responsible for these portfolios may assume unnecessary risk and overlook substantial opportunities to enhance returns if they fail to factor energy efficiency into their real estate investment decisions. A further Ceres report on Asset managers practices¹⁶ highlights specific best practices that asset managers are purporting to use to incorporate climate risks into their due diligence, corporate governance and portfolio valuation. However, the report states that there are very few asset managers actually doing this.

The issues raised in these reports highlight that climate risks and opportunities are rising on the investment agenda as environmental, social and economic implications of climate change are beginning to crystallize.

Legal standards for sustainability

The life cycle of a commercial property is affected by both legal and voluntary standards for sustainability. Pressure is mounting for all parties involved in real estate to do more to achieve sustainability throughout the lifecycle of a building, from design to demolition.

Building design

Legal standards

The Building Regulations 2000, as amended (BR2000) set out a minimum sustainability standards relating to, for example energy and water efficiency. The local authority may refuse to issue a completion certificate if the new buildings fail to meet the BR 2000 standards.

The impact of the Flood and Water Management Act 2010 will also require developments to be designed to prevent surface water run-off.

The main voluntary tool for assessing the sustainability of new commercial buildings is the British Research Establishment Environmental Assessment Method (BREEAM). BREEAM awards new commercial buildings a ‘rating’ measuring their level of sustainability over and above the minimum requirements of the BR 2000. It does not, however address existing buildings.

¹³ www.ceres.org/Document.Doc?id=519

¹⁴ Ibid. Ceres is a national coalition of investors, environmental groups and other interest organisations working with companies to address sustainability challenges such as global climate change.

¹⁵ Mercer is a global consultancy service

¹⁶ Ceres (January 2010) ‘Investors Analyze Climate Risks and Opportunities: A survey of Asset Managers’ Practices’

The Green Rating ¹⁷ is a new alternative to BREEAM . It is an assessment of the energy efficiency of a building that looks at both the building materials and the waste generated by the building. It is an international standard intended to allow companies with an international property portfolio to accurately measure the efficiency of their sites.

Other sustainability assessment tools are LEED (Leadership in Energy and Environmental Design) developed in the US and Canada; and Green Star and NABERS (National Australian Built Environmental Rating System) developed in Australia.

The Planning system

Local Authorities

The UK planning system increasingly requires the promotion of sustainable development in planning applications.

The Planning Act 2004 requires Local Planning authorities to exercise their planning powers with the objective of contributing to the achievement of sustainable development, taking into account Government Planning Policy Statements (PPS)(such as PPS 1 and 22) promoting on -site renewable or low carbon energy generation) when drafting development plans.

The result is that developers are increasingly required to ‘front load’ their planning applications with details at the outline planning stage of sustainability considerations. Developers also have to include sustainability standards for new developments as a planning obligation under Section 106 Agreements.

Environmental Impact Assessment

Under the Town and Country Planning (EIA) (England and Wales) Regulations 1999, planning applications for major developments are required to be accompanied by an environmental statement detailing the likely impacts on the environment, and measures to be taken to mitigate those impacts.

Construction of New Buildings

JCT Contracts

The standard form of contract ¹⁸ used by the construction industry (drafted by the Joint Contracts Tribunal) has recently been amended to include sustainability clauses. The sustainability clauses and guidance do not impose a rigid set of targets upon parties, but instead create a framework in which a contract can include sustainability criteria.

Waste management

Waste is of increasing environmental and economic significance as the increasing regulatory pressure and rising cost of landfill taxes is making waste management a significant cost issue for many organizations.

¹⁷ This assessment was launched in June 2009 by AXA Real Estate Investment and ING Real Estate amongst others.

¹⁸ Available at: <http://www.jctltd.co.uk/assets/Building%20a%20sustainable%20future%20together%202009%20Web.pdf>

The Site Waste Management Regulations 2008, require developers of property worth over £300, 000 to prepare a Site Waste Management Plan (SWMP). The SWMP will record details of the construction project, estimate the types and quantities of waste that will be produced and confirm the actual waste types generated and how they have been managed.

Buildings also have to be constructed by using the correct materials. Both these requirements will place additional burdens on the principal contractors to ensure that any sub- contracts include the necessary provisions to ensure that sustainability and environmental issues are covered.

Land Remediation

The RICS highlights that many sustainable building rating systems take into account land use. Valuers should take land use into account where buildings are constructed on Brownfield land and watercourse setbacks. Where land has been constructed on previously developed land, issues of potential contamination must be considered as this may bear a risk of outlay in terms of cost and /or insurance against potential problems in the future.

It is not only valuers that should consider the potential risk of contamination. Developers intending to build property will need to carry out an environmental assessment of the site, and remove all environmental and health risks before construction can begin.

In the UK the Sustainable Remediation Forum UK (SuRF) has published a Framework for assessing the Sustainable Remediation of Soil and Groundwater.¹⁹ The document presents a framework for assessing the sustainable remediation of soil and groundwater remediation and incorporating the criteria of sustainable development into contaminated land management strategies. Its aim is to realize a number of benefits, including contributing to sustainable development at a number of levels and to positively demonstrate corporate and environmental responsibility.

Ownership and Occupation of a building

Energy Performance of Buildings

The Energy Performance of Buildings (Certificates and Inspections) (England and Wales) Regulations 2007 require that the energy performance of completed building be assessed. All buildings that are bought, sold, rented or leased need an Energy Performance Certificate (EPC). An EPC has to be provided by the developer to the local authority building control, failing which the issuing of a certificate of practical completion may not be possible.

The sale or rental of a property

Valuation

When buying property, the issue is what price is or should be paid (either as a capital value or the rental value). In commercial property terms, the question that arises is whether these should reflect sustainability credentials.

¹⁹ Available at http://www.claire.co.uk/index.php?option=com_content&task=view&id=182&Itemid=78

The Royal Institute for Chartered Surveyors (RICS) is the professional body for valuers who are the main agents involved in assessing the value of buildings. The RICS is responsible for laying down guidelines for valuing assets for purposes, such as secured lending or financial reporting. Valuers, should, therefore, have access to reliable and timely information, including information on sustainability criteria if these are going to affect the valuation process. The RICS acknowledges that there are fundamental aspects of sustainability that affect property and potentially its value.

In September 2009, the RICS issued an Information Paper on Sustainability and Commercial Property Valuation²⁰. This paper was drafted as a step towards embedding sustainability as a core consideration in the valuation process. The paper outlines ways in which sustainability can be defined; how a building's green credentials may be assessed and how such characteristics might be reflected within a valuation - whether quantitatively or qualitatively. The Red Book was published in January 2010 with more substantial guidance on these issues.

The RICS guidance note on Contamination, the environment and sustainability embodies 'best practice' and reminds members that whilst they are not required to follow the advice and recommendations, when an allegation of professional negligence is made against a surveyor, the court is likely to take account of the contents of any relevant guidance notes published by the RICS, in deciding whether or not the surveyor has acted with reasonable competence.²¹

Valuation is key to investment and the provision of information is critical to ensuring that sustainability is built into investment decisions.

Energy Performance Certificates

The sellers and landlords of UK Building are required by law to provide an EPC to potential buyers and tenants, setting out information on the building's energy efficiency. This is intended to allow potential buyers to make informed decisions on property investment based on a property's green credentials.

Green leases

The adoption of lease arrangements that either encourage or require landlords and tenants to manage the asset in accordance with sustainability principles (so called green leases) is growing. The RICS recommends in its Valuation Information Paper No 13²² that the presence of a green lease is to be evaluated as it could mark a risk reduction factor within the appraisal or conversely, it could result in a lower rental bid if it contains onerous terms.

In the UK the Better Buildings Partnership (BBP)²³ provides a green lease toolkit that aims to improve sustainability and reduce the environmental footprint of existing buildings. #

Carbon Reduction Commitment Energy Efficiency Scheme

²⁰ See http://www.rics.org/site/download_feed.aspx?fileID=5751&fileExtension=PDF

²¹ RICS Practice Standards, UK; page 01.

²² Available through the RICS membership

²³ http://www.bbpc.org.uk/resources/publications/green_lease_toolkit.html

The picture for commercial property is somewhat different from residential with the UK Government using the Carbon Reduction Commitment (CRC) Energy Efficiency Scheme to try and improve energy efficiency in this sector.

The CRC is a mandatory emissions trading scheme for the UK, designed to reduce CO2 emissions of UK properties in both the commercial and public sector. The CRC commenced in April 2010 and will run for three set time periods (known as phases). The first phase will run for three years and the subsequent phases will each last for seven years. Poor compliance can potentially result in increased costs, bad publicity and criminal prosecution.

The CRC will have a significant impact on landlords, tenants and the investment market and the RICS recognizes that its impact has to be fully appraised and reference to the CRC will be included in subsequent editions of their guidance notes.

Alteration and Retrofit

The alteration or extension of a commercial building in England and Wales is governed by the Building Regulations 2000. Any building work will need to meet minimum sustainability standards covering energy efficiency and water (amongst others). If any alterations or retrofits are being carried out, the developers will have to consider whether planning permission is required.

Building materials such as cladding, ceiling and floor panels and carpets and walling are important sustainability considerations. The life cycle value of materials should be considered as part of the rental value of a building.

Demolition

Planning permission is required prior to demolishing a building. Developers will also have to comply with the Site Waste Management Regulations 2008 and BR2000 throughout the demolition process.

Challenges for businesses

The real estate sector can no longer afford to ignore the movement to establish a low- carbon sustainable global economy. Corporate bottom lines will be impacted by emerging ESG issues, particularly where disclosure must now be included in financial filings. Impacts will flow from the physical risks arising from climate change (whether or not anthropogenic) or the physical and economic risks arising from the scarcity of energy and water resources, or from legal and regulatory burdens which may be imposed by reference to those risks. The overall governance and disclosure by companies relating to ESG issues must be improved. More details are required on actions and policies or assessing and managing environmental risks.

Companies within the real estate sector should also be aware of the growing trend of the global business and human rights agenda. Companies are also going to have to track their policies and operations on the local communities within which they operate.²⁴

²⁴ <http://www.euractiv.com/en/print/socialeurope/csr-corporate-social-responsibility/article-153515>

Companies can also not ignore the fact that new social technologies, media and networks will transform the reporting landscape. The Web 2.0 revolution is leveling the playing field by giving stakeholders the opportunity to initiate and drive conversations with (or around) companies.²⁵

The risks of inaction by property investors, failing to embrace energy efficiency, are identified in the Ceres report as:

- Expected rises in energy costs, including the cost of waste consumption
- Existing and new more stringent, legislation
- Competitive and financial risks of not responding to market demand

Those companies embracing the sustainable agenda are likely to see decreased operating costs of certified buildings and combined with increased rental levels, the cash flows of the sustainable buildings should be more attractive.

Those investors and businesses pursuing energy efficiency initiatives should position them competitively for impending national climate and energy legislation that will likely make energy consumption and waste more expensive.

Global real estate businesses, with an international property portfolio, are faced with the challenge that there are inconsistencies between currently available metrics. For example, the EPBD implementation mechanism varies across member states, hence buildings cannot be compared for energy efficiency across Europe in a consistent manner.

The impact of climate change on the ability to continue to use a building efficiently in high temperatures or withstanding storms will be important. Many properties could become unsuitable in extreme heat without appropriate climate control which will impact on tenants. Such buildings may be vulnerable to obsolescence and may need retrofitting.

For those architects and developers looking to design and build new developments, there is going to be an increasing need to be aware of the current and future legal requirements and voluntary standards which may affect the design of the development. Developers will also have to liaise with the local planning authority to understand the planning policy requirements and budget for increased costs arising from compliance with sustainability standards.

Commercial organizations that may qualify under the CRC scheme (if they have not already done so) should consider how they can resource and implement new systems to capture data and report it, in order to be compliant and also benefit from the league table ranking that will predominately be based on the first two CRC compliance years.

Conclusion

The challenge for the property market will be the creation and adoption of standardized approaches to sustainability, from the valuation process to the design and construction of buildings. The occupation of

²⁵ The GRI Learning Series (2010) 'The Transparent Economy' pg 15

buildings will also be affected by the legislative frameworks and market pressures. Creation of a universal definition of what constitutes a sustainable building and simple benchmarks will be vital.

UK Climate Change Law: securing a low carbon economy by recourse to experts and government obligations?

Mark Stallworthy

School of Law

Swansea University, Swansea SA2 8PP

United Kingdom

Email: M.Stallworthy@swansea.ac.uk

Abstract:

Appropriate policy responses to threats from climate change require consideration of radical transformations in socio-economic arrangements, with substantial consequences for both our urban and rural environments. Such tasks are more likely to be addressed (whether adequately or no) in a sustained fashion at national and regional levels of jurisdiction. This paper seeks to address a key legal element underpinning the UK's national response to climate change. Its primary focus is upon the UK Parliament's legislative intervention through the Climate Change Act 2008. Following a brief overview of statutory scheme and objectives, the paper offers both theoretical and instrumental perspectives concerning those duty bearing provisions that lie at its core as a reforming measure, and the role of scientific expertise in the process. The implications of the approaches adopted will be explored, with especially the particular policy aspirations for a low carbon economy in mind.

Keywords:

Climate Change, Legislation, Governmental Accountability

1 Introduction

The multiple implications for human society of climate change can be described as paradigmatic in two essential ways. On the one hand, they raise problems of multi-level governance relating to legal mechanisms, jurisdictions and accountabilities. And on the other, challenges arise out of the polycentric nature of ensuing impacts, upon both individual and communal lifestyle expectations. In both respects, fundamental sustainability questions arise.

Impacts upon both rural and built environments are potentially vast, and are likely directly and indirectly to affect many of the ways in which we order our forms of social organisation, our methods of production, and our patterns of consumption.

Profound questions concerning legitimacy and accountability also arise. The response thus far within international fora, including through the mechanisms set up through the 1992 UN Framework Convention on Climate Change – and the subsequent Kyoto

Protocol process – can be said at best to have plateaued. Thus the lack of multilateral, and extremely limited bilateral, consensus achieved at the Copenhagen Climate Conference (December 2009) has offered little prospect as yet of a robust successor to the Kyoto arrangements being in place from 2013.

In this light, the paper discusses the essential framing process by which the UK seeks to deliver on formal targets for the reduction of greenhouse (and in particular, carbon) emissions up to 2050, and to secure key aspects of adaptation to unavoidable impacts of climate change. The Climate Change Act 2008 is central to this project and it has placed key statutory duties, especially for the securing of phased emissions reductions, upon central government. It is argued here that such legal strategy, which suggests a departure from standard constitutional practice in the UK, requires a more rigorous consideration of its implications than has hitherto been the case.

The rationale for this significant legislative incursion, and the likely implications, are the key issues addressed in the paper. The conclusion will seek to address the prospects for such a novel approach making a substantive contribution to the delivery of effective UK policy responses to climate change.

2 A Duty Based Legislative Structure for a Low Carbon Economy

2.1 Science and Climate Change

Central to our discussions of climate change, in whatever discipline, is the considerable scientific consensus around the idea that temperature rises, that have continued since the onset of the industrial era, are predominantly accounted for by increases in greenhouse gas emissions (USNOAA, 2010), and that the predominant cause is attributable to human activities.

Despite such broad level of scientific consensus, uncertainties abound, reflecting multiple complexities in the behaviours of natural systems under stress. Yet in the face of this, climate science has increasingly been called upon to offer realistic forecasting of future events and impacts. For this task scientists are equipped with a decidedly mixed tool-kit. This ranges from the limited availability of modern-era observational records, to continuing empirical studies (more especially, in the fields of glaciology and palaeontology), the integrity of which (even where supported by sophisticated computer models) can ultimately depend upon the exercise of scientific judgment, levels of broad consensus, and the processes of peer review. In effect, climate scientists are routinely being called upon to quantify the unquantifiable.

Nevertheless, for all the uncertainties, over 98% of scientists endorse the broad consensus of view that man made emissions are responsible (Anderegg et al., 2010), both for the preponderance of warming since the onset of the industrial revolution and for its continuation today (IPCC, 2007).

2.2 Representing Problems through the Carbon Footprint

Policy discussions have shifted – at varying speeds – to questions of mitigation and adaptation – the former being broadly concerned with limiting emissions, so as to

slow (or even to reverse) increases in global temperatures, and the latter with securing changes in the ordering of our affairs so as to cope with the impacts of climate change that prove to be unavoidable. In particular, the impacts from energy production and energy use – across the whole range of our economic and social activities – are central to the possibilities for reducing carbon emissions.

The objectives for mitigating the effects of climate change can usefully be framed within the notion of a shift to what is termed a ‘low carbon economy’. It can be seen from the figure below that the idea of a ‘carbon footprint’ offers a pervasive representation of carbon impacts (both direct and indirect) from the range of socio-economic activities in which we collectively engage as legal persons, including natural and non-natural persons and other entities.

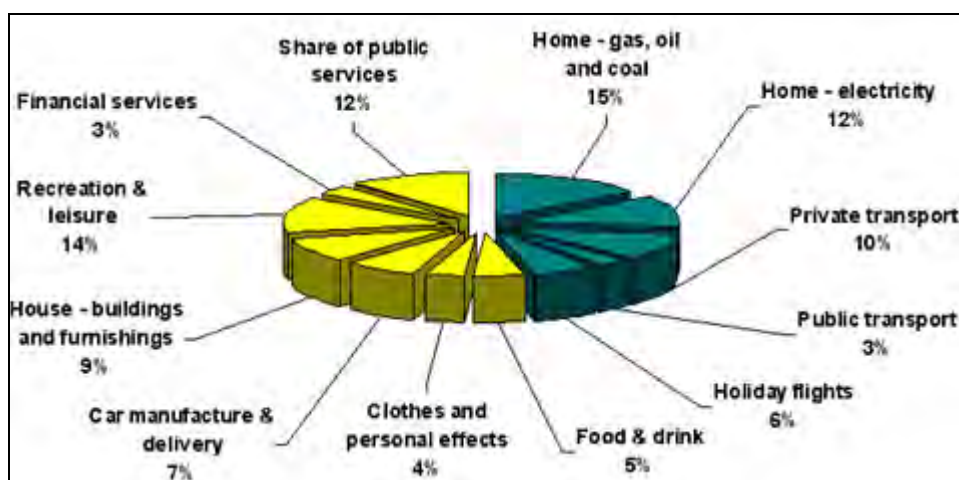


Figure 1. ‘What is a Carbon Footprint?’

Source: <http://www.carbonfootprint.com/carbonfootprint.html>

And yet, for all the growing awareness of sustainability issues, and the range of alternatives offered by environmental economists (Stern, 2008), traditional conceptions of growth, by public and private exploitation of resources, and related approaches to maximising production and consumption, continue to dominate modern economies. Hard questions arise as to how a growth ethic so conceived can be made subject to a reforming agenda, as by reconfiguring through schemes such as for ‘green’ investment, the development of green technologies, and otherwise. From a UK perspective, green product choices are still largely encouraged by at best voluntary guidance schemes. There are persisting low levels of incentive or subsidy toward the advancement of alternative energy technologies, or even towards simpler measures such as for ‘retrofitting’ older buildings for energy efficiency.

Confused signals can further be seen in the relationship between public sector investment and the macroeconomy: as in the lower relative costs of private as opposed to public forms of transport, and the development of highways and airport infrastructure on a ‘predict (demand) and provide (infrastructure)’ basis. Despite the apparent willingness of the (1997-2010) UK government to progress more explicit carbon pricing of such activities as infrastructure investment, it remained fully committed to such projects as the proposed third runway at London’s Heathrow Airport (a stance that appears likely to be reversed by the incoming administration).

2.3 Placing Duties on Central Government: the 2008 Act

From a legal perspective, given expectations of impacts from climate change, the development of suitable responses over the coming decades will require extensively reformed institutional and decision making structures and processes.

In its first attempt to create an overarching statutory framework under which it is planned further to develop national climate change policy, the UK Parliament in December 2008 passed the Climate Change Act. Importantly, the legislation articulated in mandatory terms the objective of securing realignment towards a low carbon economy. Moreover, it provided a mechanism for further development of mitigation and adaptation policy responses. That said, the Act was notably thin on policy specifics, with a few ancillary exceptions (most notable of which related to the adoption of a ‘renewable transport fuel obligation’, with minimum fuel content levels from biofuel sources).¹

The Act offers a paradigm piece of enabling legislation. Aside from the commitment (again through enabling measures) to extending the reach of emissions trading regimes within the wider domestic economy, the key provisions related to the setting of national greenhouse gas emission reduction objectives, the meeting of which was made mandatory. This central feature was supported by three further mechanisms: namely, a carbon emissions budgetary process; procedures for strategy setting; and arrangements for reporting on performance.

The most significant element in the legislation, therefore, was as a vehicle for imposing statutory duties on central government relating to defined policy outcomes. The key applications of this novel form of legislative wording can be seen in two particular instances.

In the first, section 1(1) states as follows (emphasis added):

‘It is the duty of the Secretary of State *to ensure that the net UK carbon account for the year 2050 is at least 80% lower than the 1990 baseline*’.

Then in the second, under section 4(1), it is provided that (emphasis again added):

‘It is the duty of the Secretary of State
 (a) to set for each succeeding period of five years beginning with the period 2008-2012 (‘budgetary periods’) an amount for the net UK carbon account (the ‘carbon budget’), and
 (b) *to ensure that the net UK carbon account for a budgetary period does not exceed the carbon budget*’.

2.4 Traditional and Novel Statutory Approaches Compared

The mechanism adopted accordingly poses potentially significant instrumental and constitutional questions. One explanation for its emergence could be that the application of such duties is a reflection of modern UK experience under EU law, in the field of the environment and otherwise. This extends to court rulings, not only in the European Court of Justice, but also (especially in light of the doctrine of direct

¹ Climate Change Act 2008, section 78 (and schedule 7); amending Energy Act 2004, Part 2, chapter 5, and the Renewable Transport Fuels Order 2007.

effect) in domestic courts, whereby bodies representing the state can be found in breach of obligations arising under EU law. Nevertheless, in constitutional terms, such analogy is imperfect, for it is potentially misleading to assume equivalence between obligations accepted as applying to the UK by virtue of the European Communities Act 1972,² and a purely domestic provision such as the Climate Change Act.

Therefore, whilst the legislature appears to have brought the logic of formal governmental target setting to a logical conclusion, this required statutory formulations that offer, for UK lawyers certainly, a radical departure from traditional approaches. The UK's public law system can broadly be described as one founded on notions of discretionary power, and mostly marked by a reluctance to create legal shackles on the central state. The approach can be said to reflect 'the dominant role of central government, and the historical preference in Britain for systems of political accountability over law-based systems of public administration', which arguably bring about 'a trade-off, in which political accountability triumphs over legal transparency' (Grant, 2000).

A modern illustration of the traditional approach, through provisions empowering, rather than placing duties upon, central government, can be seen in the Civil Contingencies Act 2004. This legislation provided for a structured regime applied to aspects of protection, planning, and co-ordination of responses, in the event of civil emergencies. Section 2(1) placed substantial duties on so-called 'Category 1 responders' (particularly local authorities and the emergency services) for the prevention and mitigation of risks. At the time of the Bill's passage through Parliament, a question arose as to why, despite a significant degree of central government involvement in the various processes envisaged, in this respect no equivalent duties would apply. The suggestion was decisively rejected by the sponsoring department, officials for instance made it clear that 'efforts are agreed by Ministers who are accountable for them to Parliament and are therefore not appropriate for the face of a legal instrument. This is standard practice' (Cabinet Office, 2003)

By way of contrast, aside from the Climate Change Act, the outgoing UK government had in certain limited policy fields likewise experimented with a statutory adoption of explicit duty based provisions, in support of quantitative policy target setting. Most recently, for instance, statutory duties have been placed upon the Secretary of State, not only to publish (and review) a child poverty strategy, but also to meet four UK wide income poverty reduction targets (by a deadline date of 1 April 2020). The key target appears to require that less than 10% of children must by then be living in 'low income households' (that is, below 60% – adjusted on an 'equivalised' basis to take account of larger households – of the median).³

The ground breaking measure for adopting this approach – one which is useful to consider especially because its duty basis led to litigation – was the Warm Homes and Energy Conservation Act. This provided for central governmental duties for the alleviation of fuel poverty. The Secretary of State was required to produce a strategy setting out 'policies for ensuring, by means including the taking of measures to ensure

² *R v Secretary of State for Transport, ex parte Factortame (No.2)* [1991] 1 AC 603

³ Child Poverty Act 2010, sections 1, 2.

the efficient use of energy, that as far as reasonably practicable persons (in England and Wales) do not live in fuel poverty’.⁴ The strategy had to specify a target date for achieving the above objective, the minister being further required to take steps considered necessary for its implementation,⁵ and to assess and report from time to time on progress toward ‘achieving the objectives and meeting the target dates’ and in consequence of the assessment to revise the strategy as considered to be appropriate.⁶ A Fuel Poverty Strategy was duly adopted (defining ‘fuel poverty’ as representing the need for a household to spend more than 10% of income on fuel); committing the Government ‘as far as reasonably practicable’ to ending fuel poverty for ‘vulnerable households’ (those containing older people, families with children, and householders either disabled or suffering from long term illness) by 2010, alongside a further target for eradicating all fuel poverty generally, from 2016 (BERR, 2002).

Subsequent Government projections however came to accept that large numbers would remain in fuel poverty by the identified deadline dates (BERR, 2007), in consequence of which numbers in fuel poverty since the time of the Strategy appeared broadly constant (at 1.7 million households in total), as a consequence of both income trends and real increases in fuel prices during the intervening period (HC Treasury Committee, 2008). This expected outcome prompted two leading national charities to seek judicial review,⁷ on grounds of breach of statutory duty. Their application was dismissed, on two main grounds. First, the statutory duty was held to amount to an obligation to take such steps as the Secretary of State considered necessary to implement the Strategy, in seeking or trying to end fuel poverty, which the court regarded as amounting to ‘the language of “effort” to achieve targets, rather than guarantee that targets will be reached’.⁸ Secondly, reflecting broader judicial concerns at being called upon to engage with resourcing issues, the judge ruled that the terms of the 2000 Act amounted to no more than importing a statutory duty to make efforts so far as reasonably practicable toward the desired objectives. In other words, the Secretary of State was held not to ‘assume a statutory duty whatever the cost’.⁹

Although McCombe J had described the case as presenting a difficult question of statutory construction in light of the Act’s ‘juxtaposition of policy and legal duty’,¹⁰ he had little difficulty in conforming with traditional approaches according maximum levels of administrative discretion in policy contexts related to resources, and in so doing adopted conventional reasoning: on the basis that this was not to say ‘that resources are dictating the content of the legal duty’.¹¹ It has to be remembered that the grounds of challenge in this case were limited to the question of breach of statutory duty, and that therefore the decision was made absent a rationality challenge on *Wednesbury* principles.¹² Despite the difficulties inherent in that alternative ground of challenge, it is arguable that by recourse to *Wednesbury* on such basis, the focus would also have been subtly different, and that it would have been open to the applicants to introduce questions of proportionality to perhaps greater effect. As it

⁴ Warm Homes and Energy Conservation Act 2000, section 2(1).

⁵ *Ibid.*, section 2(2)(d), 2(5).

⁶ *Ibid.*, section 2(6).

⁷ *Friends of the Earth, Help the Aged v Secretary of State for Environment, Food and Rural Affairs* [2008] EWHC 2518.

⁸ *Ibid.*, paras. 17-19.

⁹ *Ibid.*, para. 29.

¹⁰ *Ibid.*, para. 14.

¹¹ *Ibid.*, para. 30.

¹² *Associated Provincial Picture Houses v Wednesbury Corporation* [1948] 1 KB 223.

was McCombe J held that it was not open to the court ‘to review the policy decisions of the Defendants’.¹³

The above decision is instructive, in that it addresses a claim arising out of what has been suggested above, namely a novel duty-target legislative technique. Nevertheless inevitably it emerged out of its own specific policy and statutory context, and caution should be exercised before seeking any principle of broader application. In particular, the duties under the 2008 Act referred to above appear to be expressed in less contingent terms. Indeed they arguably are strict in important respects, especially in light of supplementary provisions that provide for formal publication of reasons for budget overshoot,¹⁴ and of laying before the House of proposals and policies to compensate in the succeeding budgetary period.¹⁵ That said, even the use of less conditional terminology cannot be regarded as a guarantee of successful challenge, in light of the difficulties that the courts would have in addressing issues that engage so closely with complex matters of policy. Important questions therefore remain as to how far, and on what realistic basis, the 2008 Act’s explicit duty based approach can realistically be described as open to legal challenge, and some of these are considered in the next section.

2.5 Duty Enforcement against the State

It might be expected that the adoption of the technique of imposing statutory duties of the above nature on central government would be supported by an explicit statutory enforcement regime. Yet government non-compliance is not made subject to any explicit sanctioning arrangements. This suggests that – aside from the possibilities for judicial review, the primary means of securing control over substantive breach will be through ministerial responsibility to Parliament. In support of the effectiveness of what is therefore essentially a form of political accountability, the Act contains extensive provision for mandatory ministerial reporting on performance (to Parliament),¹⁶ supported by the establishment of an overseeing reporting role, through the creation of a new statutory advisory body, the Climate Change Committee.¹⁷

In consequence, however, save through these statutory processes of Parliamentary scrutiny, it is not apparent on what basis non-compliance, especially in exceeding carbon budgets, could lead to recognised legal consequences. This places a heavy onus on securing accountability through the political process, in support of the somewhat more tenuous protection afforded by the declaratory effectiveness of the measure itself. Sponsoring ministers sought to justify the chosen approach as one rooted in the traditions of the rule of law. Thus according to Lord Rooker, in ‘putting these duties into law, we are giving them a constitutional significance which will permeate down to every level of decision making. There is no other way of achieving an equivalent effect without using the law. The duty should be looked at in this broader constitutional sense, rather than just in terms of what happens in court’.¹⁸

¹³ *ibid* para 31.

¹⁴ Climate Change Act 2008, section 18(8).

¹⁵ *Ibid.*, section 19(1).

¹⁶ *Ibid.*, sections 16(1), 18(1).

¹⁷ *Ibid.*, Part 2, sections 33-36.

¹⁸ HL Second Reading Debate, Hansard, Col. 1209, 27 November 2007.

Whilst our administrative law may indeed only exceptionally be played out in the courts, constitutional lawyers are ever suspicious at any signs of rhetorical excess used in citing the rule of law. Thus in the context of this same proposal an academic response warned that ‘governments should not pretend that they are establishing a legally enforceable regime of carbon emission reductions, thereby falsely laying claim to the credibility and legitimacy which the principle of legality, the cornerstone of the rule of law, confers’.¹⁹ Indeed, in view of the lack of explicit enforcement strategy, the enforceability of the Act’s statutory duties have been described as being based upon what amounts to ‘best endeavours’,²⁰ a term more commonly seen in transactional contexts.

Otherwise, neither does the Act contain any what might be termed value added within its procedural requirements. In other words it lacks the kind of compliance add-ons that for instance appear under the Kyoto Protocol. There, in the event of a state exceeding maximum allowable emissions, action plans must be submitted in the following commitment period, containing such elements as statements of reasons for non-compliance, proposed remedial measures, and timescales. Moreover, whilst non-compliance persists, there will be a suspension of a state’s eligibility to partake in carbon credit transfers.²¹ Further illustrations of iterative controls can be seen even within the flexible regimes contained under EU framework environmental directives. Thus the Water Framework Directive obliges Member States, upon failure to meet adopted standards for conformity with ecological, quantitative and chemical water objectives, to effect remediation measures.²² Likewise, under the Nitrates Directive, national authorities must establish targeted action programmes once waters at risk of nitrate pollution are identified and vulnerable zones designated.²³ Returning to the 2008 Act, it is noteworthy that government rejected proposals for a mandatory regime for remedial actions in the event of failure to meet net carbon account benchmarks.²⁴

It can be said therefore that the Act being silent as to enforcement, any potential for juridical interpretation of the application of the above duties can only be determined by way of recourse to judicial review, in accord with ordinary principles. This was in the mind of the Bill’s sponsors, if in a somewhat inchoate form. The head of the government’s climate change support team at the responsible department, Defra, expressed the cautious view that this being ‘a fairly new type of duty ... the courts may indeed mandate particular action, although we could not predict that’.²⁵ The minister of state went further, suggesting that whilst in most circumstances ‘courts do no more than issue a declaration’ there would also be ‘the possibility of the court making a more stringent order, such as ordering them (Government) to purchase credits’.²⁶ Yet such an order would arguably constitute an extensive judicial

¹⁹ HL/HC Joint Committee Report, *Draft Climate Change Bill*, HL 170-1, HC 542-1 (August 2007), evidence of Austin, R., para. 113.

²⁰ *Ibid.*, evidence of Forsyth, C. F., paras. 108-9.

²¹ *Kyoto Protocol to the UN Framework Convention on Climate Change*, 37 ILM (1998), 22, Part XV, 2-5.

²² Directive 2000/60/EC establishing a Framework for Community Action in the Field of Water Policy, [2000] OJ L 327/1.

²³ Directive 91/676/EEC concerning the Protection of Waters against Pollution Caused by Nitrates from Agricultural Sources [1991] OJ L375/1; also, Case C-221/03, *Commission v Belgium* [2005] ECR I-8307.

²⁴ HC Environmental Audit Committee, *Beyond Stern: from the Climate Change Programme Review to the Draft Climate Change Bill*, 2006-07, 7th Report, HC 460 (2007), para. 106; also, HL/HC Joint Committee, note 19 above, paras. 120, 123, 126.

²⁵ *Ibid.*, paras. 85-6.

²⁶ *Ibid.*, note 18, above.

interference in the specifics of policy choice, and would be unlikely to commend itself to the courts.

Climate change challenges bear the hallmarks of what have been described generically as ‘wicked problems’. These can be seen as being marked especially by high levels of uncertainty, pervasive levels of political contention, and widespread entrenchment of conflicting stakeholder and other values (Rittel and Webber, 1973). Aside from the engagement of related questions of governance, competence and legitimacy (Rajamani, 1997), therefore, judicial restraint is likely to apply *a fortiori* given the complexities of the policy fields surrounding climate change, with the courts resisting invitations to adopt more intensive approaches to review.

Yet with a view to judicial enforcement, it can on the other hand be tentatively suggested that there are aspects of the legislation that are likely to be more amenable to challenge. For instance, in so far as the Act relies upon mandated activities, across definite timelines and outcomes, possibilities for review can be envisaged. This is particularly the case in relation to the requirements concerning the production and publication of strategies;²⁷ for reporting on progress;²⁸ and for consultation (in particular through the Climate Change Committee).²⁹ In each of these respects, doubts that can be said to apply to more closely policy related questions, especially concerning outcomes, are less likely to apply.

3 Institutional Structures for a Low Carbon Economy

3.1 The Climate Change Committee

Climate change has been described as being ‘ultimately a crisis of governance’ (Hulme, 2009). This can be said to encapsulate a major problem that afflicts the development and delivery of effective climate change law and policy. Problems in producing appropriate decision making processes and outcomes are arguably reflected further in a lack of engagement across all levels of society in the potential long term effects of climate change. These are part also of wider questions concerning trust in political institutions, often laid bare in contexts where ‘sound science’ (a politicians’ term) is under scrutiny. This has been graphically put as follows: ‘what is in question is not only the limits of, or the gaps in, expert knowledge, but an inadequacy which compromises the very idea of expertise’ (Giddens, 1990).

The Climate Change Act offers a further specific response, in the emergence of new institutional arrangements for the purposes of adoption and driving through of mitigation and adaptation objectives. The establishment of the Climate Change Committee under the Act is a case in point. Although formally an advisory and reporting body, this Committee may prove to be of pivotal influence in determining whether, and how far, governments comply with the novel duty based regime previously discussed: in particular for securing phased emissions reductions up to 2050, and for meeting the 5-yearly carbon budget limits.

²⁷ Climate Change Act 2008, sections 13, 14, 19.

²⁸ *Ibid.*, sections 16, 18, 20.

²⁹ *Ibid.*, sections 7(1), 9(1), 12(3).

For its part, as referred to previously, government is subject under the Act to a number of related reporting obligations. Thus where a carbon budget is not been met then its final statement for the budgetary period must explain why this is the case.³⁰ Should the government seek to alter carbon budgets *ex post*, not only must the Secretary of State first have determined that there have been significant changes affecting the basis on which the previous decision was made,³¹ but if, following the required consultation, any amendment departs from the Committee's recommendation, then the Secretary of State must publish a statement setting out reasons.³²

An illustration of the potential importance of the Committee can be seen from its most recent progress report, produced in accordance with the requirements of section 36(1) of the 2008 Act. The Committee concluded that the implementation of policy measures accounted for only 'a small part' of total UK emissions reductions achieved in the previous year, but that the economic recession and other exogenous factors were the primary cause. In consequence the Committee recommended that over the current (five year) carbon budgetary period the government's aim should (already) be amended so as to outperform by around 6% (Climate Change Committee, 2010).

3.2 Implications of Expert Leadership in Risk Management

The emergence in such a key role of an expert body such as the Climate Change Committee arguably conforms to what has been classically described as a 'co-production' model of decision making (Weber, 1946). Such a model can be said to move beyond both more top down political determinist and technocratic approaches, by placing greater emphasis upon securing openness through negotiation in respect of such processes as setting and delivery of policy goals.

The new institutional arrangements under the 2008 Act open up opportunities for wider debate and reflection concerning optimum risk management strategies in relation to climate change. It will be important to measure the adopted approach against conformity with those essential requirements – especially, for clarity as to administrative priorities for addressing identified risks, for maximum transparency, and for participation and accountability – that have been identified as minimum demands for an effective risk-based regulatory system (Fisher, 2007).

In a closely related field, concerning infrastructure planning decisions, it is notable that (in June 2010) the incoming UK government announced its intention to bring forward a Decentralisation and Localism Bill.³³ One of its essential elements is expected to relate to the Infrastructure Planning Commission (created under the Planning Act 2008). This body, whilst having a fairly circumscribed remit, is vested with substantial executive decision making powers in relation to the process of major infrastructure planning applications. The government proposes the abolition of the Commission, and intends to place its work within the remit of its own Planning Inspectorate, with decisions made by the Secretary of State. Although major strategy

³⁰ *Ibid.*, section 18(8).

³¹ *Ibid.*, section 21(2)(3).

³² *Ibid.*, section 22(7).

³³ Hansard, HC Deb, 29 June 2010, Col 34-36WS (G Clark, Minister of State, Department for Communities and Local Government).

documents had remained within the aegis of the Secretary of State even under the Planning Act (through the issue of National Policy Statements), the above retrenchment suggests a formal return of the planning control process into the political – and away from the technical – domain.

Role distinctions as between those who are politically accountable and those with technical expertise have been widely explored: for instance the European Commission, in developing its approach to the precautionary principle, adopted a broad divide as between the roles of expert, in risk assessment, and political decision maker, in risk management (European Commission, 2000). The idea of rigidly separate roles for delivering respective assessment and management responsibilities is arguably not a realistic one; and for instance it will often be appropriate for scientists to be tasked with considering and advising on risk management options (Office of Science and Technology, 2000). Two further issues arise here. First, fuller attention to the delivery of accountability generally becomes necessary the more that decision making becomes dispersed away from those more clearly legitimated institutional arrangements.

Secondly, risk is ultimately socially constructed, and issues as to both risk perception and risk acceptability generally demand broader constituencies than can be provided by specialist bodies alone. Expert analysis is in certain respects inevitably subjectivised, for scientists too have values, and these also influence processes of interpretation, more especially in areas of uncertainty (Wynne, 1996). Beck has argued that society's reliance upon the benefits of technological progress has arguably led to an 'experimental society' (Beck, 1992), in which an unregulated, unaccountable technological and scientific establishment performs an inadequate protective role. Interestingly, in recent litigation in which judicial review was sought concerning the legality of UK Government application of the Pesticides Directive (for the purposes of providing protection to persons),³⁴ respectively the High Court,³⁵ and the Court of Appeal,³⁶ reached different conclusions in the face of conflicting approaches adopted by two expert bodies.

Accordingly, in this vital policy area of climate change (and in apparent contrast with amending proposals relating to infrastructure planning), it is likely that the state will find it increasingly either opportune or necessary to cede important areas of discretionary risk management to an expert body such as the Committee on Climate Change. The implications will require careful scrutiny, as to both the efficacy of the operation of the statutory regulatory regime and its wider consequences for the democratic polity.

4 Conclusion: Seeking Sustainable Solutions in Response to Climate Change

Some brief thoughts are offered by way of conclusion.

³⁴ EC Directive 91/414 on Plant Protection

³⁵ *Downs v Secretary of State for the Environment, Food and Rural Affairs* [2009] Env LR 19.

³⁶ *Ibid.*, [2009] EWCA 664.

The Climate Change Act 2008 is not only a unique piece of legislation by virtue of its scope and objectives for securing a transition by the UK towards a low carbon economy. It also provides a context for two particular mechanisms, which though on the face of it they are yet another mix of the substantive and procedural, suggest something more radical. First, there is the application of apparently clear, unambiguous duties upon central government, particularly to produce and to meet carbon budget commitments over defined timeframes and in light of stated net emissions reduction goals. Secondly, the creation of the Climate Change Committee, however circumscribed its executive authority may on the face of it be, could amount to a momentous institutional reform, and the outturn of the first carbon budgetary period (2008-12) will provide valuable evidence as to how far the potential trends discerned above have eventuated.

One final consequence of interest to the scientific community should be noted, as expert analysis is likely both to be called upon and more openly scrutinized in the public domain. This may lead to conflict between politicians and science advisers: as occurred in the UK in 2009, in the context of whether it was appropriate for the (academic) chair of an advisory committee on illegal drugs to continue to publish views on questions of classification at variance with government policy.

With an eye to the role of the Climate Change Committee, it must be expected that as its influence and authority becomes further established, its pronouncements will bring it increasingly into the domain of political contestation. Already, in the climate change context, academic climate scientists have found themselves increasingly under pressure, from a variety of stakeholder and ‘climate sceptic’ interests; as has been seen in the so-called ‘glacier-gate’ and ‘climate-gate’ controversies (Hickman, 2010). A body with an influential role in risk management will be an inevitable focus for challenge and controversy.

According to Jared Diamond (2006), a society’s responses to its problems ‘depend on its political, economic and social institutions and on its cultural values. Those institutions and values affect whether the society solves (or even tries to solve its problems’. Ultimately, issues of accountability, as well as policy effectiveness, lie at the core of the questions considered above. In delivering compliance with statutory duties now placed on governments, and emerging policy drivers that underpin performance, expert technical contributions will be pivotal. In light of the crisis in civil society and its institutional capacities that forms the backdrop to these developments, for all that reliance on expert inputs is a valid response, resulting systemic problems still need to be addressed. In so far as ordinary processes for engaging the public and private spheres, and civil society more generally, are proving inadequate in the face of climate change, the radical approaches ushered in by the UK’s legislation may succeed in assuaging fears that more draconian measures will prove necessary.

5 References

Anderegg W R L, Prall, J.W, Harold, J and Schneider, S H (2010) 'Expert Credibility in Climate Change', 107 (27) *Proceedings of the National Academy of Sciences of the United States of America* 12107-9.

Beck, U (1992), *Risk Society: Towards a New Modernity*, Sage, London.

BERR (2002), Department for Business, Enterprise and Regulatory Reform, *Fuel Poverty Strategy* (November 2001, revised December 2002), <http://www.berr.gov.uk/files/file16495.pdf>.

BERR (2007), *5th Annual Progress Report*, 2007, <http://www.berr.gov.uk/files/file42720.pdf>.

Cabinet Office (2003), Civil Contingencies Secretariat, response before Houses of Lords/Commons Joint Committee (cited in Walker, C and Broderick, J (2006) *The Civil Contingencies Act 2004: Risk, Resilience and the Law in the United Kingdom*, OUP, Oxford, para. 8.27).

Climate Change Commission (2010), *Meeting Carbon Budgets – Ensuring a Low-Carbon Recovery*, 2nd Progress Report to Parliament (June 2010).

Diamond, J (2006), *Collapse – How Societies Choose to Fail or Survive*, London, Penguin, 14-15.

European Commission (2000), *Communication on the Precautionary Principle*, COM(2000) 1.

Fisher, E (2007), *Risk, Regulation and Administrative Constitutionalism*, Hart, Oxford.

Giddens, A (1990), *The Consequences of Modernity*, Polity, Cambridge, 131

Grant, M (2000), *Environmental Court Project: Final Report*, Department for Environment, Transport, and the Regions, London, para. 12.10.12.

HC Treasury Committee (June 2008), *Budget Measures and Low-Income Households*, 13th Report, 2007–08, HC 326, para. 151.

Hickman, L (2010), 'US Climate Scientists Receive Hate Mail Barrage in Wake of UEA Scandal', *Guardian*, 2 July.

Hulme, M (2009), *Why We disagree about Climate Change – Understanding Controversy, Inaction and Opportunity*, CUP, Cambridge.

IPCC (2007), International Panel on Climate Change, Expert Working Group I, *Fourth Assessment Report, Climate Change 2007 – The Physical Science Basis*, CUP, Cambridge.

Office of Science and Technology (2000), *Review of Risk Procedures Used by the Government's Advisory Committees Dealing with Food Safety*, HMSO, London.

Rajamani, L (2007), 'Public Interest Environmental Litigation in India: Exploring Issues of Access, Participation, Equity, Effectiveness and Sustainability', 19 *Journal of Environmental Law* 293.

Rittel, H and Webber, M (1973), 'Dilemmas in General Theory of Planning', 4 *Policy Sciences* 155.

Stern, N (2008), *Key Elements of a Global Deal on Climate Change*, London, LSE.

USNOAA (2010), US National Oceanic and Atmospheric Administration, Earth System Research Laboratory, Mauna Loa, Hawaii, *Climate Report*:
<http://www.esrl.noaa.gov/gmd/ccgg/trends/>.

Weber, M (1946), *Essays in Sociology* (eds. Gerth, H H and Mills, C W), OUP, Oxford.

Wynne, B G (1996), 'May Sheep Safely Graze? A Reflexive View of the Expert-Lay Knowledge Divide', in Lash, S, Szerszynski, B and Wynne, *Environment and Modernity*, Sage, London, 59.

COBRA 2010 Legal Research Symposium

Understanding property rights in carbon: A methodological inquiry

Jason Prior^{1,2} and Spike Boydell²

¹Institute for Sustainable Futures,
University of Technology, Sydney, PO Box 123, Broadway NSW 2007
Australia

²Asia-Pacific Centre for Complex Real Property Rights,
University of Technology, Sydney, PO Box 123, Broadway NSW 2007
Australia

Email: Jason.Prior@uts.edu.au; Spike.Boydell@uts.edu.au.

Abstract:

Climate change threatens to have wide-ranging impacts on the sustainability of ecosystems and presents enormous challenges for conventional modes of socio-economic governance. Against this backdrop, there have been a range of responses to put a price on carbon such as the EU Emissions Trading Scheme (ETS) and New Zealand ETS, and the unsuccessful Carbon Pollution Reduction Scheme Bill 2009 in Australia. These ETS models do not clearly articulate the underlying asset, the carbon property right, upon which the price of commoditised carbon is secured. As a result of such initiatives, a clear understanding of what comprised a property right in carbon is emerging as a foundation component in the arsenal of mechanisms that are being brought together to mitigate and adapt to climate change at the international level.

Drawing on the work of various researchers this paper develops a framework for understanding disciplinary, interdisciplinary and transdisciplinary research in the context of carbon property rights. In developing this framework, the paper engages with the inevitable task of exploring the diverse approaches that have been applied to analyse real property rights. We outline how recent research on emergent property rights in natural resources, such as water and carbon, have challenged the way in which related rights, obligations and restrictions have been conceived.

In undertaking this exploration, we take a step back from the immediacy of the issues of carbon property rights to explore the way they are framed through institutions and disciplines. We present the findings from this exploration in two parts. The first explores the state of research and practice into carbon property rights from within disciplines, between disciplines and across disciplines. The second part extends this

discussion by exploring some of the challenges that face such research. The paper is intentionally broad in scope, although our research focuses attention on the challenges facing research within, between and across the ‘normative’ disciplines such as law, economics, planning, and politics.

Keywords:

Carbon property rights, disciplinary, interdisciplinary, transdisciplinary, planning, law, economics, sociology

1 Introduction

In recent times the issue of climate change has moved to the forefront of scientific and policy agendas. Climate change threatens to have wide-ranging impacts on the sustainability of ecosystems and presents enormous challenges for conventional modes of socio-economic governance. Against this backdrop, there have been a range of responses to put a price on carbon (e.g. EU Emissions Trading Scheme, NZ Emissions Trading Scheme). However, like the unsuccessful Carbon Pollution Reduction Scheme Bill 2009 in Australia, these ETS models do not clearly articulate the underlying asset, the carbon property right, upon which the price of carbon is secured. Achieving a coherent international understanding of what comprises an emergent property right in carbon is a core component in the arsenal of mechanisms / techniques that are being put together to mitigate and adapt to climate change.

Drawing on the work of various researchers this paper develops a framework for understanding disciplinary, interdisciplinary and transdisciplinary research (Max-Neef, 2005, Klein, 2004, Wickson et al., 2006, Riedy, 2003, Lawrence and Despres, 2004, Nicolescu, 2006) in the context of carbon property rights. In developing this framework, the paper engages with the inevitable task of exploring the diverse disciplinary (e.g. Arnold, 2002, Zellmer and Harder, 2007), pluridisciplinarity, and interdisciplinary (e.g. von Benda-Beckmann et al., 2006a) approaches that have been applied to analyse real property rights. In particular, these approaches investigate how emergent property rights in natural resources, such as water and carbon, have begun to question the way in which related rights, obligations and restrictions have been conceived. Such an exploration provides insights into the values, ethics and epistemology but also the institutional frameworks that one may need to engage in such practices (Snow, 1964, Gieryn, 1999, Pickering, 1992)

In undertaking this exploration, we inevitably take a step back from the immediacy of the issues of carbon property rights to explore the way they are framed through institutions and disciplines. We present the findings from this exploration in two parts within this paper. The first explores the state of research and practice into carbon property rights from within disciplines, between disciplines and across (or more correctly, transcending) disciplines. The second part extends this discussion by exploring some (and by no means all) of the challenges that face such research, ranging from disciplinary participation through the ability of disciplines to support the necessary and rapid emergence of regulatory frameworks for carbon property rights. The paper is intentionally broad in scope, although our research in the second stage of the paper focuses attention on the challenges facing research within, between and across the ‘normative’ disciplines such as law, economics, planning and politics,

disciplines which are of particular interest to this COBRA conference and to our own ongoing research agenda.

2 Within, between and transcending disciplinary research and practice

2.1 Disciplinary research and practice

Disciplinarity is specialisation in isolation, which is the level of most academic knowledge production at present (Miller et al., 2008, Max-Neef, 2005, 6). Over the past century, property rights research, and more recently property rights of natural resources such as carbon, has emerged as a foci within a broad range of disciplines including archaeology (Earle, 2000), anthropology, ethics, sociology (Sorensen, 2000, Hann, 2007, Hann, 1998a), psychology (Beaglehole, 1934-35, Beaglehole, 1931), law (Arnold, 2002, Sheehan and Kanas, 2008, Sheehan, 2008, Boyd, 1995), geography (Bromley, 2008, Kedar, 2003, Waters et al., 2003), history, philosophy, economics (Furubotn and Pejovich, 1974, Gordon, 1954, Alchian, 1965), biology (Hardin, 1968) and planning (Jacobs and Paulsen, 2009, Krueckeberg, 1995, Birch and Silver, 2009, Webster and Lai, 2003) to name but a few.

The recent attention on carbon property rights from within these diverse disciplines highlights a range of concerns and uncertainties. Soil science, for example, questions if there is adequate existing scientific knowledge to support the idea of a sequestered carbon property right in biota, soil or water (Sheehan and Kanas, 2008). Technological disciplines, such as forestry and geology, question whether techniques such as bio-sequestration or geo-sequestration are effective, efficient and sustainable enough to support a carbon management based on carbon property rights. Concerns within such normative disciplines as law, economics, planning and politics challenge what a carbon property right is and how it is best constituted for the future benefit of society, providing a stable foundation for exchange markets and good governance (Boydell et al., 2009b, Boydell et al., 2009a, Hepburn, 2007, Hepburn, 2009). And finally, concerns from within disciplines, such as philosophy, as to the values and ethics that underpin carbon management systems supported by property rights.

The questions raised within these disciplines not only relate to the emergent property rights associated with these natural resources, but also challenge the way in which property rights have hitherto been understood. A key example here can be drawn from recent research within the discipline of law which has sought to question the very foundation of property rights as conceived through the 'bundle of rights' and has argued for its complete replacement (Arnold, 2002, Zellmer and Harder, 2007). For more than a century the 'bundle of rights' identified by Henry Maine in *Ancient Law* (1861), has provided some common ground for dialogue on property within law (but also within a range of other disciplines).

Various approaches have been proposed as to how to define the components of such bundles. Schlager and Ostrom (1992) offer: *use rights*, such as the *right to access* the resource, *withdraw* from a resource, or *exploit* a resource for economic benefit; and *control*, or *decision-making* rights, such as the rights to *management* (plant a crop), *exclusion* (prevent others from accessing the field), and *alienation* (rent, sell, or give away the rights) (for a more detailed exposition of these issues, see Boydell and Searle, 2010, at this conference).

Some scholars have argued for the need to develop more complex metaphors and models for understanding contemporary property rights arguing that the ‘bundle of rights’ is conceptually limited (e.g. Zellmer and Harder, 2007, Arnold, 2002). These critiques have sought to question the existing conceptual effectiveness of the bundle to address the complex nature of the inter-relationships and interests that are involved in property rights associated with natural resources. Instead, they have presented new metaphors and models that establish new building blocks from which to understand and problematise contemporary property. For example, Arnold (2002, 281 & 283) argues that the ‘bundle of rights’ is unable to address two of the essential principles that underline modern property rights of natural resources such as: the interconnectedness of people and their physical environment and, the importance of the unique characteristic of each object.

In the face of these seemingly insurmountable limitations to the bundle of rights, Arnold called for the formulation of a ‘web of interests’ as a metaphor for property that addresses the “concepts of interconnection, ‘thingness’ (object-regard), and the uniqueness of the object of property” (Arnold, 2002, 283, see also Zellmer and Harder, 2007). Considering property through a web of interest suggests a more heterogeneous approach to the management of real property, intended to capture a wide range of relationships between people, and between people and property. Within the context of the web of interest, “the thing [such as water] considered property is at the centre [of a concentric web] and relationships with the thing ... form the internal strands of the web and the surrounding framework” (Zellmer and Harder, 2007, 7). In devising this broad-based metaphor for property, scholars are clear to point out that “it must encompass all aspects of property, not merely those that involve the natural environment” (Arnold, 2002, 283).

These critiques from within law share the view that articulating property through the bundle of rights is too narrowly conceived, that it brings up the image of exclusivity and separation. The bundle approach does not adequately reflect the increasing sense of interconnection and co-existence that marks contemporary property rights such as those associated with carbon (Boydell et al., 2009a, Arvanitakis and Boydell, 2010) or water (Zellmer and Harder, 2007). These new models and metaphors proposed by such researchers as Arnold promote the use of such terms as ‘interests’ rather than just ‘rights’, as a way to promote a more fluid articulation and understanding of contemporary property in natural resources. Such approaches to the reconceptualisation and formalisation of property rights have not been without criticism (Meinzen-Dick and Mwangi, 2009, Hann, 2007).

2.2 At the boundaries of disciplinary research and practice

The ability to shift beyond the realms of disciplinarily to interdisciplinary and transdisciplinary research into carbon property rights is complex. Interdisciplinary work is troublesome because the search for knowledge in different disciplinary fields entails different interests, and thereby different values too; and the different possibilities of knowledge about different subjects also lead to different epistemologies. Thus, differences among those who practice the various disciplines are pervasive and have been aptly described as cultural ones (Snow, 1964, Snow, 1962). By seeing disciplines as cultures, one recognises that knowledge, methods, and theoretical approaches to a particular cannot be separated from practice.

The intractability of working effectively across disciplinary cultures is apparent through the emergence of disciplinary ‘boundary work’ (Gieryn, 1999, Pickering, 1992). One purpose of boundary work between disciplines involves an ongoing effort to delineate the domain of disciplines. The process of maintaining disciplinary boundaries involves delineation of their borders and related resistance to threats from outside (see Solomon and Hackett, 1996) as well as threats to its cognitive authority from within (e.g. pseudo-science). It is to societal advantage to protect such strong disciplinary based research, as long as it does not preclude the transdisciplinary and boundary work that is essential to realising new and meaningful interpretations of complexity. Put another way, for example we would not want to stifle the powerful research that such a discipline as soil science could contribute to our ability to understand and conceive of carbon sequestration in soils (Sheehan and Kanas, 2008). Such disciplinary knowledge has the ability to bring a sense of security to a carbon management system that inevitably we hope will play a key role in our intergenerational sustainability. However, such knowledge only serves society if it can be operationally integrated and applied within other disciplines.

Just as it is important for us to support such disciplinary research, it is also important for us to have the opportunity to explore possibilities outside of these disciplinary structures and between them. Initially, ‘boundary work’ was formulated to maintain the boundaries of disciplinary communities, however, there is also increasing recognition from within disciplines that work at the boundary / between disciplines can be fruitful. Such work is not seen as an unbridgeable chasm, rather one that must often be approached carefully through the establishment of boundary objects and organisations. This allows disciplines to explore the flexibility of boundary work without threatening the core values and interests of those disciplines. That is, boundary objects and organisations which sit between different disciplinary cultures, can be used by individuals within each for specific purposes without losing their own cultural identity (Star and Griesemer, 1989). A clear example of a boundary object that is truly interdisciplinary is the formation of bio-chemistry; the formation of such an interdisciplinary field must overcome the ‘epistemic and methodological sovereignty’ of those disciplines that formed it. The idea of boundary organisations has proved particularly useful within the context of environmental management (Guston, 1999). This is because a need has emerged for the development of entities that contain both scientific and political elements. As Miller (2001, 480) identifies, these “institutions that are neither laboratories nor conventional political organizations are increasingly prevalent features of the institutional landscape of modern society and play key roles in managing the interactions between science and politics, economics, and cultures”.

These new boundary institutions have emerged in the context of “mobilizing science in support of global policy making [to address climate change] ... turning to science in a wide array of decisions such as apportioning blame for greenhouse gas emissions, assessing the need for regulatory intervention in the global economy, finding policy and technological solutions, and compensating victims of climate change currently underway” (Miller, 2001, 479, see also Pickering and Guzik, 2008)

The call for such boundary work has arisen within the domain of property rights, particularly within the emergent rights in water and carbon. Such boundary work is a particular focus of our ongoing research. We have been influenced in this boundary

work by research undertaken by the Legal Pluralism Research Group at the Max-Planck Institute for Social Anthropology, notably by the thinking of Franz and Keebet von Benda-Beckmann (see von Benda-Beckmann et al., 2006a). They question the strong foundation of property rights from within particular disciplines such as economics and law (Cole and Grossman, 2002, Hoppe, 2004, FIABCI, 2002, Coase, 1960, Furubotn and Pejovich, 1974, Economist, 2006, Devlin and Grafton, 1998, McEwin, 1993). For example the dominance of the standard liberal model and related economic (mis)understandings over property rights can be seen in the Coasian foundation of such instruments as the Kyoto Protocol. Within this context, putting a price on carbon, and the use of carbon offsetting and carbon trading as mechanisms to abate carbon emissions is underpinned by the Coasian solution for the ‘tragedy of the commons’ – the clear assignment of individualised property rights by “privatizing the commons and trading the resulting property rights” (Hepburn, 2007, 376).

The Coasian solution puts faith in private bargaining and is supported by the ‘new institutional economics’ of North (1990) and others (see Hann, 2007, 293, for a critique of the ‘standard liberal model’). Let us contextualise this in Hardin’s parable of the Tragedy of the Commons. If the property rights to carbon are divided among individuals, then (according to the standard liberal model) each individual will have proper economic incentives to manage those property rights wisely. Questioning the dominance of economic and legal disciplinary perspectives over our understanding of property rights, von Benda-Beckmann’s contest that property is too significant for the functioning and reproduction of social systems to be confined to the domain of economy, pointing out that property regimes “cannot easily be captured in one-dimensional political, economic or legal models” (von Benda-Beckmann et al., 2006a, 2). In place of these one-dimensional economic and legal models, von Benda-Beckmann et al. (2006a) argue for a more interdisciplinary exploration of property rights, in particularly one that takes into consideration the social and anthropological understanding of property.

A similar call for a more interdisciplinary approach to property rights can be found in the work of Elinor Ostrom (see Ostrom et al., 1994, Ostrom and ebrary Inc., 2002, Ostrom, 2005, Ostrom, 1990). She advocates the analysis of common property through the Institutional Analysis and Development framework (for Ostrom’s own critique of the Coasian solution and Hardin’s parable, see Ostrom, 1990, 1-28, Hess and Ostrom, 2007, 10-12). Such a perspective seeks to draw on the long tradition outside of law and economics, which has sought to understand property rights through such disciplines as sociology and anthropology amongst other social sciences (see Gluckman and University of Zambia. Institute for Social Research., 1967, Gluckman and University of Zambia. Institute for African Studies., 1972, Gluckman, 1965, Gluckman, 1955, von Benda-Beckmann, 1995).

Real property rights, obligations and restrictions can be found in and change across the full range of human societies, both in time and space (e.g. Hann, 1998b, Beaglehole, 1931, Emigh, 1999). The social disciplines seek to argue that different societies give different answers to the question of what can be owned? The inclusion of new objects or the exclusion of old ones is a process variably shaped by political, cultural, economic, and technological factors - why some objects can be owned and others cannot depends on culture and formal law. The set of potential owners varies across societies (e.g. different societies recognise ownership by households, lineages,

villages, kin groups, or other collectives). Gluckman (1965) was driven by concerns that the bundle of rights approach would privilege western colonial powers. He developed another influential property metaphor with his concept of ‘estates of administration’, which grew out of his social anthropological fieldwork in southern Africa in the late colonial period (early to mid 20th century). Gluckman’s theorisation was intended in part to create property rights that recognised indigenous relations (Gluckman, 1965). The sociological and anthropological perspectives seek to recognise that as society’s demands on real property have evolved, so have the use rights thereon (Polanyi and Prosch, 1975). Important counsel can be drawn from Hann (2007, 310), “...we should not exaggerate the impact of the present fetishization of private property. The trend to propertization, in the sense of extending individual ownership and converting more and more ‘non-real’ objects into market or quasi-market commodities often appears unilineal and irresistible”.

Using boundary work in property rights as proposed by the von Benda-Beckmann’s (2006a), and arguably through our interpretation of researchers such as Ostrom (see Ostrom, 1990), we explore property rights through interdisciplinary and transdisciplinary means in the next section of this paper. This contrasts with specific delimited boundary work that may be just multidisciplinary, a situation where there may be many practitioners who are looking at their individual pieces of the puzzle from their respective disciplinary perspectives without evident cooperation (see Caulfield, 1998). It also contrasts with pluridisciplinarity, which is demonstrated through the level of cooperation between practitioners from different disciplines without clear coordination or stimulus to promote this cooperation (Max-Neef, 2005).

2.3 Transcending disciplinary boundaries in research and practice

Both interdisciplinary and transdisciplinary approaches require collaborators to transcend unconscious habits of disciplines that are not only epistemic but also often institutional (for instance, the modern research university is quite clearly tied together as institutions, in addition to their disciplinary components; the two in effect are often blurred). The more those unconscious habits are explicated and the more we understand how the disparate characteristics of the various intellectual disciplinary cultures are related to the necessarily different interests, values, and epistemologies, the more feasible becomes the goal of transcending disciplinary thought habits and engaging in interdisciplinary and ultimately transdisciplinary research. ‘Transdisciplinarity concerns itself with what is between the disciplines, across the different disciplines, and beyond all disciplines’ (Nicolescu, 2006, 143).

Nicolescu highlights the importance of recognising that there is a theoretical transdisciplinarity (implying a well-defined methodology), phenomenological transdisciplinarity (implying building models to connect theory to observed reality), and experimental transdisciplinarity (allowing experimentation using a procedure that will enable research design to be replicated). Accepting simultaneous and rigorous consideration of these three aspects will allow the huge potential of transdisciplinarity to be realised. Max-Neef (2005) provides insight into this explication by identifying ‘hierarchical levels’ for disciplines that need to be taken into consideration when approaching interdisciplinary and transdisciplinary research. We expanded this model (see Figure 1) in an earlier paper (McDermott and Boydell, 2010, 4), by integrating

several interpretations of transdisciplinary hierarchies (Stewart and Cohen, 1997, Max-Neef, 2005, von Benda-Beckmann et al., 2006b).

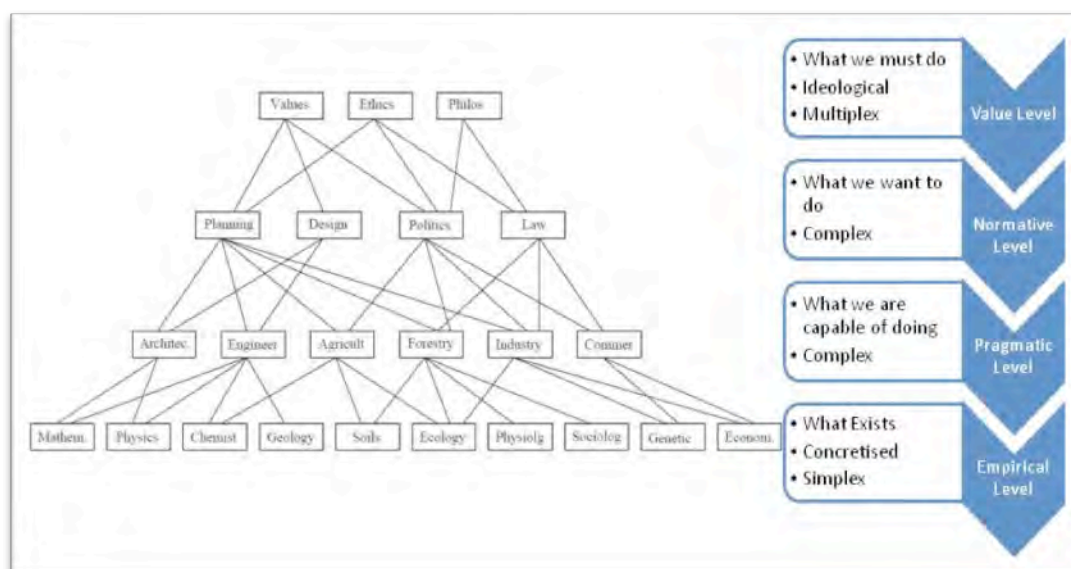


Figure 1: Hierarchy of Transdisciplinarity

(Source: McDermott and Boydell, 2010)

For Max-Neef (2005, 7), “interdisciplinarity is organized at two hierarchical levels. It thus connotes coordination of a lower level from a higher one”. A natural example of interdisciplinary is the relationship between the normative discipline of planning and those lower in the hierarchy “which defines the purpose of technical disciplines such as engineering, agriculture, forestry and architecture” (i.e. on the second and third hierarchical levels) (Max-Neef, 2005, 7). Climate change, as Max Neef has noted, is one of the “list of ... problematiques that are defining the ... [21st century that can’t] ... be tackled from the sphere of individual disciplines” and which clearly represents a transdisciplinary challenge (Max-Neef, 2005, 5). What is true for carbon is also a reality for carbon property rights.

Integrating these concepts with the work of researchers such as von Benda-Beckmann et al. (2006b), it is interesting to note the similarity between the transdisciplinary practice framework put forward by Max-Neef and the systematically rigorous framework for the study of property put forward by von Benda-Beckmann et al., not to mention other researchers within the field of property commons such as Elinor Ostrom (Schlager and Ostrom, 1992, Ostrom, 1990, Ostrom, 2005). In its systematic rigour, this framework is an advancement on the loose notion of what Hann (1998a) refers to as the *embeddedness* of property rights within such disciplines as economics and law. Real property must be analysed at four layers which approximate those proposed in Max-Neef’s in his hierarchy of disciplines. Firstly, as the von Benda-Beckmann model suggests, the notion of property rights is an important element in ideological or cultural understanding; that is to say that it has a value, ethic and philosophy (i.e. how should we do what we want to do?). Next they argue that property rights receives extensive attention within the normative disciplines such as law (i.e. what is it we want to do?). They also point to the fact that property rights manifest themselves at the level of concrete practice, (i.e. what are we capable of doing?), and finally that it also finds itself within everyday practices (i.e. what

exists?). The von Benda-Beckmann model is inherently transdisciplinary, at least by Max-Neef's criteria.

In place of disciplinary models, von Benda-Beckmann et al. put forward a transdisciplinary model that distinguishes the need to engage with property rights through the norms and values of a cultural tradition – the ideological layer. The second layer consists of political and legal regulations models. Layer three consists of the social relations of property (such as particular land uses and the way that these may be tied to particular forms of kinship) and the implications for different forms of social inequality. Finally, at the layer of 'practices', the actors may reinforce the patterns of the other layers or they may initiate changes. Such a transdisciplinary model for property rights, integrating notions of legal pluralism, provides an opportunity to question deeply engrained beliefs about property.

3 Key challenges

The following section of the paper posits a series of challenges that we face in the future development of research and practice relating to property rights in carbon through disciplines, between disciplines and across disciplines. The goal here is to point the way towards the broad challenges facing scholarship and practice that lie on the horizon, a very close horizon at that. As noted earlier when discussing these broad conceptual challenges, we will draw mostly on the normative disciplines (law, economics and planning) as examples. All of these challenges are faced in unique ways by all disciplines and their ongoing ability to research and engage in policy development with government institutions.

3.1 The challenge of disciplinary participation

The first set of challenges focuses on issues of *adequate* participation by different disciplines. To make our point, we will draw on the strong body of research that has emerged into property rights from law and economics, but has remained comparatively weak and is underrepresented as a body of research within the planning discipline. We are not suggesting that research from within such disciplines as law is stifled, rather that the research within other disciplines such as planning is strengthened to provide the opportunity for developing more complex and multidimensional insights into property rights.

In recent years, the disciplinary area of planning in its narrow sense of land-use planning (see Scott, 2010, Shaw and Lord, 2007, Stead and Meijers, 2009, Wilson, 2007, Wilson, 2009, Woltjer, 2005), and in its broadest sense (see Carlsson-Kanyama et al., 2008, Chermack and van der Merwe, 2003, Chermack et al., 2006, Cole, 1997, Grossmann, 2006), has increasingly been engaging with the breadth of the potential impacts of global warming as have a broad spectrum of other disciplines such as law and economics. Yet within this broader spectrum of research, there has been a conspicuous absence of research from within planning into the complexity of property rights.

We contest that this absence is principally due to the study of property rights in natural resources being part of the solution to global issues, issues that have been limited for the most part to a range of disciplines such as law and economics

(McEwin, 1993, Cole and Grossman, 2002). There are increasing calls for and a nascent body of research into property rights from within planning (Jacobs and Paulsen, 2009, Webster and Lai, 2003, Needham, 2006, Zellner, 2008). As Jacobs and Paulsen (2009) have recently noted, property rights is very much the neglected theme of 20th century planning within the US, and has the potential to remain as neglected an issue in planning in the 21st century. We support their argument that this is a significant oversight given “the role private property rights have played in shaping public land use planning [during the 20th century]” (Jacobs and Paulsen, 2009, 134).

We would take this call a stage further by arguing that research into property rights from within planning is not only important for the way in which it shapes land use planning but also for the way in which it “confronts the art and artfulness of [the diverse domains of planning such as] social planning” (Denman, 1978, see preface) and environmental planning (Zellner, 2008). At present, planning relies heavily on law and economics for its understanding of property rights (see Needham, 2006). Whilst these links are relevant and should not be denied, it is important to recognise that planning extends far beyond such intersecting domains. If planners are to provide a more powerful contribution to such ideas as property rights, they need, as Jacobs and Paulsen argue, to begin to “tell their own story about how property rights relates to planning” (Jacobs and Paulsen, 2009, 141). We see this type of story emerging within the recent work of such researchers as Needham (2006) as opposed to earlier research (Denman, 1978). Needham (2006, 1) asks fundamental questions such as how “land-use planning affects property rights”.

The benefits, for example, of studying property rights more fully from within planning are broad and could range from its epistemic make-up through to its objects of practice. As Max-Neef (2005) has noted, like philosophy, planning is one of the most integrative disciplines. As such, research led by planning has the potential to develop more complex understanding of property rights than many other disciplines. As Jacobs and Paulsen have noted planning practice has the ability to engage more fully with property rights through the development of such practical instruments as “property rights impact statements as part of the planning process” (Jacobs and Paulsen, 2009, 141). Such instruments could play a powerful role in exploring and testing the broad impacts of property rights and ultimately verifying if the intended moral and ethical foundations which inform the ideologies behind particular types of property rights such as those in carbon are being attended.

3.2 The challenge of internationalisation of disciplinary research and practice

The second set of challenges are those of internationalisation associated with the global extent of the climate change problem and the difficulties on studying property rights not only from within diverse disciplines but within disciplines that have epistemic differences both in and between countries. This is particularly important given that climate change does not operate between but across national boundaries. We have to look no further than the authors’ base – Australia, being a federation of six states and two territories, suffers the challenge of having eight similar yet discrete sets of legislation relating to property. All Australian States now have legislation to define *Carbon Sequestration Rights* (CSRs), recognising sequestered carbon as a property right that can be bought and sold, and protected through registration on

freehold land title if desired. The challenge of the Australian situation is compounded by multiple interpretations of carbon property rights – that is, variously, forestry rights, forestry carbon rights, profit à prendre, and carbon property rights (for a broader discussion on this see Boydell et al., 2009a, Hepburn, 2009).

We acknowledge the efficacy of the EU ETS, which manages to transcend a diversity of legal frameworks under the umbrella of the open border mandate of the European Union. This is despite the point that the epistemic base of disciplines does not only change across disciplines but between countries (Browaeys, 2004, 2-3), and in particular for this paper, between Anglo-Saxon and Continental legal traditions. The Anglo-Saxon tradition is grounded in the feudal ‘doctrine of estates’, grounded in the ultimate ownership of the monarch – i.e. fee simple absolute in possession there by grace of the Crown. Conversely the Continental legal tradition, through its engagement with Napoleonic and Germanic legal systems, starts with the concept of ‘full ownership’ (Needham, 2006, 34), albeit that the term ‘ownership’ has a range of contested interpretations given that what is actually ‘owned’ is one or more property rights, obligations or restrictions. “The most basic clement in the anthropologist’s approach to property (and other key concepts) is to question whether the understanding that has emerged in Western intellectual traditions can provide an adequate base for understanding the whole of humanity” (Hann, 2007, 289). Moreover, as Hann questions, are the currently dominant understandings of property relations (i.e. the ‘standard liberal model’) a “highly distorted representation of how contemporary Euro-American property systems actually function”? (Hann, 2007, 289-230) The answer to this question necessarily triggers a range of philosophic reflections on the theory of knowledge (surrounding the qualities of our beliefs and values, now and into the future) that are beyond the scope of the current paper.

3.3 The Challenge of Integration

The third set of challenges arises from the imperative of integrating governance systems (for example, international, federal and state) and disciplinary knowledge (for example, law, planning and science) in the creation of property rights in carbon. The breadth of the potential impacts of global warming invites such integration. The logic of integrated management in the climate change context is unquestioned given the borderless nature of the commons. As is apparent from our discussion of boundary work previously, responses to environmental management generally requires integration across a number of dimensions, including across disciplines, information sources, ecological components over time and space, as well as governance and institutional structures. The clear fault lines that currently exist in the current structures which we use to support and understanding property rights are beginning to be identified by such researchers as Arnold (2002) and von Benda-Beckmann et al. (2006a) as outlined above. These insights provide clear reasoning as to why such integrated governance and understanding is required in terms of property rights in natural resources (e.g. water and carbon).

As we mentioned above, some of the research thinking that we have drawn on (such as Max-Neef, 2005, Nicolescu, 2006, von Benda-Beckmann et al., 2006a) provides useful insights that we can apply to the issue of carbon property rights to achieve a more integrative perspective. As Max-Neef (2005) has made clear, the obvious starting point for promoting such integration rests with those disciplines, such as

philosophy, planning, law and design, which by their very nature are fertile domains for integration. For example, planning by its very nature is *multidisciplinary* involving the natural sciences, social sciences and humanities, and also incorporates action-oriented practices aimed at transforming the built and natural environment. As von Benda-Beckmann et al. (2006a) have noted, we need to find ways of linking up and understanding property rights in natural resources across all four of the disciplinary / professional levels. Max-Neef (2005, 7) argues that these can be reduced down to four simple questions: How should we do what we want to do (with carbon property rights)? What is it we want to do (with carbon property rights)? What are we capable of doing (with carbon property rights)? What exists (that is, what scientific evidence exists that allows us to understand what a carbon property right is)?

The challenges to such integration are multiplex. Our capacity to develop an integrated understanding and approach to carbon property rights is contingent on our ability to comprehend its complexity. The compartmentalisation of scientific and professional knowledge through disciplines is compounded by the sectoral division of responsibilities in contemporary society around responsibility (see discussion immediately above around multi-governmental directions within Australia). Such challenges to integration are not only apparent across disciplines and sectoral divisions, but also *within* disciplines and sectoral divisions. Perhaps the most complex challenge that faces the creation of effective carbon property rights in the immediate future is how we integrate (or otherwise) significant new disciplinary and sectoral ‘bodies of knowledge’ about carbon property rights into other pre-existing disciplinary and sectoral domains. How, for example, do we integrate the significant new body of ideals about what a carbon property right should embody into legal principles, and how do we integrate these with other existing general environmental laws and property law, and with the broader legal framework.

As a recent / emergent commodity, carbon property rights are a creature of the common law. As we highlighted in a recent paper (Boydell et al., 2009a, 113), “... such common law will attempt to characterise this emergent tenurial interest in terms of the characteristics of other established tenures in the Australian property law system. ... there is an inherent susceptibility within the “bundle of rights” for some legal rights to be favoured over other rights. As the common and statute law currently stand, this susceptibility in the case of carbon is arguably a fatal limitation to the prospect of establishing a true freestanding property right in carbon”. Given the shortcomings of these issues in Part 10 of the Australian Government *Carbon Pollution Reduction Scheme Bill 2009*, we believe the Minister for Climate Change, Penny Wong, was fortunate that the Bill did not receive assent. Had it become legislation, the lack of clearly defined property rights in carbon would have quickly rendered it unworkable.

4 Conclusion

Several important implications about the conception of carbon property rights, carbon reduction and research arise from this paper.

First, the challenge caused by multiple diverse disciplinary perspectives on the subject of carbon property rights. In exploring carbon property rights, we noted how several disciplines have explicit goals and boundaries. Within these respective contexts,

carbon property rights are seen as an extension of each particular discipline. As we note, these disciplinary perspectives can provide key insights into what constitutes a carbon property right and should not be negated. The answer (or way forward) lies in contextualising these fragmented and disparate insights.

Second, there is increasing recognition that key questions about the natural and human-made environment need to be explored using innovative concepts and methods that cannot be addressed solely from within individual disciplines with their respective epistemic and theoretical frames. As we discussed, this stems from the fact that the capacity of human societies to deal with environmental questions such as climate change increasingly require research and thinking that extends across the boundaries of disciplines. The study of property rights in natural resources, such as carbon, does not escape this, as we note there is an increasing body of research calling for property rights in natural resources to be explored from more integrated perspectives - pluridisciplinarity, interdisciplinary, and transdisciplinary. Such boundary work requires disciplinary practitioners to subject their epistemological and methodological rules to reflexivity, argumentation and negotiation. This can be complex given that researchers within disciplines do not always agree on their own disciplinary rules and standards. Thus, more integrative practice and research must take a spirit of generosity and goodwill.

Finally, we highlighted some of the key challenges to the study of carbon property rights from these diverse positions. Our focus here was on the challenges facing research within, between and across the 'normative' disciplines such as law, economics, and planning. As Hann (2007, 307) identified, "...the economist who considers only rational maximising behaviour is blind to the relationships and values which underpin property and hold society together". Just as the standard liberal model is an inappropriate way of solving the global financial crisis, existing legal precedent or market based thinking is not necessarily going to assist in addressing the impacts of climate change. Adaptation is part of the solution. Likewise, transdisciplinary and boundary approaches are essential practices to find new and innovative ways to articulate workable property rights in carbon.

5 References

- ALCHIAN, A. A. (1965) Some Economics of Property Rights. *Il Politico*, 30 (4): 816-29.
- ARNOLD, C. A. (2002) The Reconstitution of Property: Property as a Web of Interests. *Harvard Environmental Law Review*, 26, pp.281-364.
- ARVANITAKIS, J. & BOYDELL, S. (2010) The Miner and the Activist : an Australian parable for our carbon constrained world. *Journal of Political Ecology : Case Studies in History and Society*, 17, 55-67.
- BEAGLEHOLE, E. (1931) *Property - A Study in Social Psychology*, George Allen & Unwin.
- BEAGLEHOLE, E. (1934-35) Onwership and Inheritance in an American Indian Tribe. *Iowa law Review*, 20, 304-316.
- BIRCH, E. N. L. & SILVER, C. (2009) One Hundred Years of City Planning's Enduring and Evolving Connections. *Journal of the American Planning Association*, 75, 113-122.

- BOYD, T. (1995) The Valuation of Land of Indigenous People - the New Zealand Experience. *Australian Land Economics Review*, 1, 9-15.
- BOYDELL, S. & SEARLE, G. (2010) The contemporary urban commons - a case study of Darling Harbour, Sydney. IN ADSHEAD, J. (Ed.) *COBRA RICS Research Conference, Law and Dispute Resolution Special Session*. Paris, RICS.
- BOYDELL, S., SHEEHAN, J. & PRIOR, J. (2009a) Carbon Property Rights in Context. *Environmental Practice*, 11, 105-114.
- BOYDELL, S., SHEEHAN, J., PRIOR, J. & HENDY, S. (2009b) Carbon Property Rights, Cities and Climate Change. *Fifth Urban Research Symposium 2009*. Marseille, France.
- BROMLEY, D. (2008) Formalised property relations in the developing world: The wrong prescription of the wrong malady. *Land use policy*, 26, 20-27.
- BROWAEYS, M.-J. (2004) Complexity of epistemology: Theory of knowledge or philosophy of science? *Fourth Annual Meeting of the European Chaos and Complexity in Organisations Network (ECCON)*, Driebergen, Netherlands, NYENRODE University, The Netherlands Business School.
- CARLSSON-KANYAMA, A., DREBORG, K. H., MOLL, H. C. & PADOVAN, D. (2008) Participative backcasting: A tool for involving stakeholders in local sustainability planning. *Futures*, 40, 34-46.
- CAULFIELD, C. (1998) *Masters of illusion: the World Bank and the poverty of nations*, London, Pan Books.
- CHERMACK, T. J., LYNHAM, S. A. & VAN DER MERWE, L. (2006) Exploring the relationship between scenario planning and perceptions of learning organization characteristics. *Futures*, 38, 767-777.
- CHERMACK, T. J. & VAN DER MERWE, L. (2003) The role of constructivist learning in scenario planning. *Futures*, 35, 445-460.
- COASE, R. H. (1960) The Problem of Social Cost. *Journal of Law and Economics*, 3, 1-69.
- COLE, D. H. & GROSSMAN, P. Z. (2002) The Meaning of Property Rights: Law versus Economics? *Land Economics*, 78, 317-330.
- COLE, S. (1997) Futures in global space. *Futures*, 29, 393-418.
- DENMAN, D. R. (1978) *The Place of Property: A New Recognition of the Function and Form of Property Rights in Land*, Cambridge, Geographical Publications Limited.
- DEVLIN, R. A. & GRAFTON, R. Q. (1998) *Economic rights and environmental wrongs : property rights for the common good*, Northampton, Mass., Edward Elgar.
- EARLE, T. (2000) Archaeology, Property, and Prehistory. *Annual Review of Anthropology*, 29, 39-60.
- ECONOMIST (2006) Property Rights. *Economics A-Z*. Economist.com.
- EMIGH, R. J. (1999) Means and Measures: Property Rights, Political Economy, and productivity in Fifteenth-Century Tuscany. *Social Forces*, 78, 461-490.
- FIABCI (2002) Property Rights and Economic Prosperity. *Issue 100 November 2002*. FIABCI Press.
- FURUBOTN, E. G. & PEJOVICH, S. (Eds.) (1974) *The Economics of Property Rights*, Cambridge, Mass., Ballinger Publishing Company.
- GIERYN, T. F. (1999) *Cultural boundaries of science : credibility on the line*, Chicago, University of Chicago Press.

- GLUCKMAN, M. (1955) *Judicial process among the Barotse of Northern Rhodesia ; by Max Gluckman.*
- GLUCKMAN, M. (1965) *Politics, law and ritual in tribal society*, Oxford, Blackwell.
- GLUCKMAN, M. & UNIVERSITY OF ZAMBIA. INSTITUTE FOR AFRICAN STUDIES. (1972) *The ideas in Barotse jurisprudence*, [Manchester, Eng.], published on behalf of the Institute for African Studies.
- GLUCKMAN, M. & UNIVERSITY OF ZAMBIA. INSTITUTE FOR SOCIAL RESEARCH. (1967) *The judicial process among the Barotse of Northern Rhodesia*, Manchester,, published on behalf of the Institute for Social Research, University of Zambia, by Manchester U.P.
- GORDON, H. S. (1954) The Economics of a Common Property Resource: The Fishery. *Journal of Political Economy*, 62, 124-42.
- GROSSMANN, I. (2006) Three scenarios for the greater Hamburg region. *Futures*, 38, 31-49.
- GUSTON, D. H. (1999) Stabilizing the Boundary between US Politics and Science: The Role of the Office of Technology Transfer as a Boundary Organization. *Social Studies of Science*, 29, 87-111.
- HANN, C. (2007) The State of the Art: A new double movement? Anthropological perspectives on property in the age of neoliberalism. *Socio-Economic Review*, 5, 287-318.
- HANN, C. M. (1998a) Introduction: the embeddedness of property. IN HANN, C. M. (Ed.) *Property Relations - Renewing the anthropological tradition*. Cambridge, UK, Cambridge University Press.
- HANN, C. M. (1998b) *Property relations : renewing the anthropological tradition*, Cambridge ; New York, Cambridge University Press.
- HARDIN, G. (1968) The Tragedy of the Commons. *Science*, 162, pp.1243-1248.
- HEPBURN, C. (2007) Carbon Trading: A Review of the Kyoto Mechanisms. *Annual Review of Environment & Resources*, 32, 375-393.
- HEPBURN, S. (2009) Carbon Rights as New Property: The benefits of statutory verification. *Sydney Law Review*, 21, 239-271.
- HESS, C. & OSTROM, E. (2007) *Understanding knowledge as a commons : from theory to practice*, Cambridge, Mass., MIT Press.
- HOPPE, H.-H. (2004) The Ethics and Economics of Private Property. Mises Institute.
- JACOBS, H. M. & PAULSEN, K. (2009) Property Rights: The Neglected Theme of 20th-Century American Planning. *Journal of the American Planning Association*, 75, 134 - 143.
- KEDAR, A. (2003) On the legal geography of ethnocratic settler states: notes towards a research agenda. IN HOLDER, J. & HARRISON, C. (Eds.) *Law and Geography*. Oxford, Oxford University Press.
- KLEIN, J. T. (2004) Prospects for Transdisciplinarity. *Futures*, 36, 515-526.
- KRUECKEBERG, D. A. (1995) The Difficult Character of Property: To Whom Do Things Belong? *Journal of the American Planning Association*, 61, 301-309.
- LAWRENCE, R. J. & DESPRES, C. (2004) Introduction: Futures of Transdisciplinarity. *Futures*, 36, 397-405.
- MAINE, H. S. (1861) *Ancient Law: its Connection with the Early History of Society, and its Relation to Modern Ideas*, New York, Charles Scribner.
- MAX-NEEF, M. A. (2005) Foundations of transdisciplinarity. *Ecological Economics*, 53, 5-16.

- MCDERMOTT, M. & BOYDELL, S. (2010) Complexity Epistemology and Real Property Rights. *FIG Congress 2010*. Sydney, International Federation of Surveyors.
- MCEWIN, R. I. (1993) Legal Versus Economic Perceptions of Real Property Rights. *The Australian Economic Review*, 3rd Quarter 1993, 35-40.
- MEINZEN-DICK, R. & MWANGI, E. (2009) Cutting the web of interests: Pitfalls of formalizing property rights. *Land Use Policy*, 26, 36-43.
- MILLER, C. (2001) Hybrid Management: Boudnary Organizations, Science Policy, and Environmental Governance in the Climate Regime. *Science, Technology, & Human Values*, 26, 478-500.
- MILLER, T., BAIRD, T., LITTLEFIELD, C., KOFINAS, G., CHAPIN, S. & REDMAN, C. (2008) Epistemological Pluralism: Reorganizing Interdisciplinary Research. *Ecology and Society*, 13.
- NEEDHAM, B. (2006) *Planning, law and economics : an investigation of the rules we make for using land*, London ; New York, Routledge.
- NICOLESCU, B. (2006) Transdisciplinarity - past, present and future. IN HAVERKORT, B. & REIJNTJES, C. (Eds.) *Moving Worldviews: Reshaping sciences, policies and practices for endogenous sustainable development*. Leusden, ETC/COMPAS.
- NORTH, D. C. (1990) *Institutions, Institutional Change and Economic Performance*, Cambridge, Cambridge University Press.
- OSTROM, E. (1990) *Governing the Commons: the Evolution of Institutions for Collective Action*, New York, Cambridge University Press.
- OSTROM, E. (2005) *Understanding institutional diversity*, Princeton, Princeton University Press.
- OSTROM, E. & EBRARY INC. (2002) *The drama of the commons*. Washington, DC, National Academy Press.
- OSTROM, E., GARDNER, R. & WALKER, J. (1994) *Rules, games, and common-pool resources*. Ann Arbor, University of Michigan Press.
- PICKERING, A. (1992) *Science as practice and culture*, Chicago, University of Chicago Press.
- PICKERING, A. & GUZIK, K. (2008) *The mangle in practice : science, society, and becoming*, Durham, Duke University Press.
- POLANYI, M. & PROSCH, H. (1975) *Meaning*, Chicago, University of Chicago Press.
- RIEDY, C. (2003) A Deeper and Wider Understanding of Sustainable Development. Melbourne, Victoria, Ecopolitics XIV Conference: Greening Sustainability.
- SCHLAGER, E. & OSTROM, E. (1992) Property Rights Regimes and Natural Resources: A Conceptual Analysis. *Land Economics*, 68, 249-262.
- SCOTT, M. (2010) Introduction to Rural Planning. *Planning Theory & Practice*, 11, 134 - 136.
- SHAW, D. & LORD, A. (2007) The cultural turn? Culture change and what it means for spatial planning in England. *Planning Practice and Research*, 22, 63 - 78.
- SHEEHAN, J. (2008) Carbon Property Rights and Emissions Trading: The Australian Experience. *Environmental Law Update Conference in Wellington*. Wellington, New Zealand.
- SHEEHAN, J. & KANAS, O. (2008) Carbon Property Rights in Soil. *14th Annual Pacific Rim Real Estate Society Conference (PRRES)*. Kuala Lumpur, PRRES.

- SNOW, C. P. (1962) *Science and government : the Godkin lectures at Harvard University, with a new appendix*, New York, New American Library.
- SNOW, C. P. (1964) *The two cultures: and a second look*, Cambridge, University Press.
- SOLOMON, S. M. & HACKETT, E. J. (1996) Setting Boundaries between Science and Law: Lessons from Daubert v. Merrell Dow Pharmaceuticals, Inc. *Science, Technology, & Human Values*, 21, 131-156.
- SORENSEN, A. B. (2000) Toward a Sounder Basis for Class Analysis. *The American Journal of Sociology*, 105, 1523-1558.
- STAR, S. L. & GRIESEMER, J. R. (1989) Institutional Ecology, 'Translations' and Boundary Objects: Amateurs and Professionals in Berkeley's Museum of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19, 387-420.
- STEAD, D. & MEIJERS, E. (2009) Spatial Planning and Policy Integration: Concepts, Facilitators and Inhibitors. *Planning Theory & Practice*, 10, 317 - 332.
- STEWART, I. & COHEN, J. (1997) *Figments of Reality: The Evolution of the Curious Mind*, Cambridge, Cambridge University Press.
- VON BENDA-BECKMANN, F. (1995) Anthropological approaches to property law and economics. *European Journal of Law and Economics*, 2, 309-336.
- VON BENDA-BECKMANN, F., VON BENDA-BECKMANN, K. & WIBER, M. (2006a) *Changing properties of property*, New York, Berghahn Books.
- VON BENDA-BECKMANN, F., VON BENDA-BECKMANN, K. & WIBER, M. G. (2006b) The Properties of Property. IN VON BENDA-BECKMANN, F., VON BENDA-BECKMANN, K. & WIBER, M. G. (Eds.) *Changing Properties of Property*. New York, Berghahn Books.
- WATERS, R., WOODSFORD, P. & HOSKINS, R. (2003) Getting it together - the geography jigsaw. London, RICS Geomatics Research, RICS.
- WEBSTER, C. & LAI, L. W.-C. (2003) *Property Rights, Planning and Markets: Managing Spontaneous Cities*, Cheltenham, UK, Edward Elgar.
- WICKSON, F., CAREW, A. L. & RUSSELL, A. W. (2006) Transdisciplinary Research: Characteristics, Quandaries and Quality. *Futures*, 38, 1046-1059.
- WILSON, E. (2007) Comment: Response to Planning Theory and Practice 2006, 7(2), Interface: Is the Issue of Climate Change too Big for Spatial Planning? *Planning Theory & Practice*, 8, 125 - 127.
- WILSON, E. (2009) Multiple Scales for Environmental Intervention: Spatial Planning and the Environment under New Labour. *Planning Practice and Research*, 24, 119 - 138.
- WOLTJER, J. (2005) Comment: The Multidimensional Nature of Public Participation in Planning: Comment on Innes and Booher. *Planning Theory & Practice*, 6, 273 - 276.
- ZELLMER, S. B. & HARDER, J. (2007) Unbundling Property in Water. Tuscaloosa, Alabama, Alabama Law Review.
- ZELLNER, M. L. (2008) Embracing Complexity and Uncertainty: The Potential of Agent-Based Modeling for Environmental Planning and Policy. *Planning Theory & Practice*, 9, 437 - 457.

Building regulations and control in the face of climate change

Henk Visscher,¹ Frits Meijer and Laure Itard¹

¹OTB Research Institute for the Built Environment,
Delft University of Technology,
The Netherlands

Email: H.J.Visscher@tudelft.nl; F.M.Meijer@tudelft.nl

Abstract:

Since a few years the attention for building regulatory systems and enforcement procedures is growing. Various developments in society, politics and the construction industry have influenced changes in the systems of building control in the last 20 years. The influence of climate change and the related demands on buildings will have a very strong impact of further transformations in this field. This was a reason to initiate a new CIB taskgroup (TG79) focussing on building regulations and control in the face of climate change. In this paper we present some research findings about the relation between regulations and actual performances of buildings. These experiences make clear that the function of building control might change in the next years.

Keywords: Building Regulations, Building Control, Energy Performance, Energy Saving, Climate Change

1 Introduction

The European building sector is responsible for about 40% of the total primary energy consumption. To reduce this share, the European Commission (EC) has introduced the Energy Performance of Buildings Directive, the EPBD (2002/91/EC). This framework has led to energy performance certificates for buildings. The EC has also highlighted that future adaptations of the EPBD may be extended to include 'low energy or Passive Houses' as a requirement, setting a target date of 2020. For newly built houses the national building regulations prescribe increasing levels of energy performances. More and more countries, but also regions or municipalities, formulate ambitions for net zero energy or carbon neutral houses also on the shorter term.

Formulating ambitions and sharpening regulations is relatively easy to do. Technical solutions are currently available to realize the passive house standard in building projects. There is quite some evidence however that the mainstream of building processes do not lead to the pre-defined quality. Traditionally the municipal departments of building control in most countries had an important role in assuring that building plans and construction processes would lead to buildings that meet the minimum required quality levels. There is a tendency to put more emphasis on the responsibilities of owners and private parties to ensure quality. This means that the

private parties will have to improve their working process and will have to learn to handle performance guarantees. Owners will require guarantees from the designers and building companies for the quality of their property. Certification and accreditation of parties, processes and products will become more important for building processes in general.

For the realization of high energy performance standards, a reliable quality assurance system will be very important. In most countries that have some experiences with passive houses some form of performance guarantee and associated quality assurance scheme exists. It is crucially important to study these examples.

This paper continues in section 2 with an elaboration on the trends in regulations and building control. Section 3 presents some signs of failures in current building process in the Netherlands. In section 4 the results of a study on the relation between the levels of energy performance regulations and the actual energy use in the built houses are presented. In section 5 we draw conclusions and finally in section 6 the TG79 group with its purpose and intentions is presented.

2 Developments in systems of building regulations and building control

Building regulations are the subject of an ongoing debate between, on the one hand, those in favour of deregulation and reducing the administrative burden and, on the other hand, new quality demands that require government intervention. Currently in the Netherlands, both sides of this debate appear to be gaining in importance. Deregulation, as well as high targets for energy conservation, save guarding structural safety and a reliable government, are high on the politicians' agenda. The desire for deregulation is leading to the opinion that greater emphasis should be placed on the responsibility of property owners, which could lead to less government intervention. However, the existing forms of quality control for private actors in the Dutch building industry seem to be of quite a low standard. Accidents occur and physical quality does not appear to be sufficiently important. As the CO₂ and energy targets increase, stronger regulations and accurate building control become a priority. In the past ten years, it has become increasingly clear that the quantity and quality of assessments carried out by many municipal authorities leave something to be desired (VROM Inspectorate, 2007).

In this context we should remember that the client and the parties who engage for the design and construction stages have primary responsibility for complying with regulations. When a building permit is granted, this suggests that the plan has been shown to comply with all the regulations. But this is not the case. In practice, a permit is granted because, during the checking process, the plan was not found to deviate from the regulations.

We will now return to the continuing call by politicians for greater deregulation and easing of the administrative burden. In 1997 we contributed to the building-regulations project as part of the MDW (Market Forces, Deregulation & Legislative Quality) programme of the Ministry of Economic Affairs. The purpose of our research was to formulate deregulation proposals on the basis of examples from other European

countries (Visser, 1997). Notably, in those countries, many private-sector parties are involved in assessment and inspection. We have studied (Visser, 2000) how the responsibility for these tasks could be transferred to the private sector in the Netherlands too, primarily through the certification instrument. The Ministry of Housing, Spatial Planning and the Environment (VROM) also took up this idea. Since the end of the 1990s, it has been developing a process certificate for assessing building plans against the requirements of the Building Decree.

The current cabinet is aiming to reduce the administrative burden by 25%. Again, the field of building regulation is seen to have a great deal of potential in this regard. The Ministry of Economic Affairs and the Ministry of VROM appointed the Construction Sector Fundamental Review Committee (Commissie Fundamentele Verkenning Bouw) chaired by Sybilla Dekker, the former Minister for VROM, to draw up proposals for the far-reaching simplification of building regulations. The committee recommended the abolition of preventive assessment of building plans by local authorities. The client should be responsible for complying with the regulations and should also ensure that sufficient checks are in place. It can engage a certified body to do this, but there may be alternatives. The role of the municipal authorities will shift towards that of process auditing, i.e. supervising the checks. The question is then: how this can be put into practice?

In many countries there are problems with a lack of compliance with building regulations, and this often serves as a stimulus for reviewing and improving the system of building control. The considerable pressure to deregulate in the Netherlands has parallels in other countries. There is a clear trend towards increasing the role of private parties. In many countries, the role of local authorities in carrying out assessments and implementation inspections has virtually disappeared.

Therefore it is interesting to study innovative ways in which quality is guaranteed by private parties. The certification of passive houses is a field that requires building actors to transform the usual building process into a performance based approach and to learn by doing. In the next section we illustrate how the building process can be impacted when the client requires a passive house.

Besides the conditions described in the previous section, the poor performances of the building industry in the mainstream building projects in combination with a withdrawing government from building regulations and actual building control is perhaps the most important reason to develop reliable certification schemes.

3 Failures in the Dutch building industry

The cost of failures in the Dutch building industry amounts to more than 10% of turnover (USP marketing consult, 2007). Total investment costs (including maintenance) in homes were € 46 billion in 2005, which means annual wastage of € 4.6 billion in this part of the building industry. Vereniging Eigen Huis, a consumer organization for homeowners, carries out final inspections on many new homes. In 2005 it was reported that construction companies are gradually improving their standards. The average number of deficiencies in more than 1,400 homes examined at new build

housing areas was 17.5 per home. However, some homes had as many as 71 deficiencies. There are also many problems with aspects of building physics, as revealed in a study of 78 housing projects by the VROM Inspectorate (Kuindersma et al., 2007). The researchers observed acute health risks, reduced living comfort and, above all, poor energy performance. In order to comply with EPC regulations, a system whereby heat is recovered from the ventilation system (balanced ventilation) is often installed. In the past few years, this system has been installed in approximately 400,000 Dutch homes. Problems with the system in the Vathorst area of Amersfoort have featured regularly in the news (Duijm et al., 2007). An analysis of the problems has shown that they are not necessarily due to the ventilation system itself, but that poor quality management throughout the construction chain can lead to an accumulation of faults.

We suspect that the Dutch situation is not unique. At a meeting of the European Consortium of Building Control in Riga in 2008, representatives from many countries reported on problems in the individual countries. Although the problems are very diverse, it is apparent that in many countries there is a discussion about the organization of building control in the context of quality problems.

There are major challenges in terms of realizing and maintaining the physical performance of homes. Requirements will become much more stringent than is currently the case, particularly with regard to energy conservation, the indoor environment and integral environmental quality. Quality management and, above all, quality assurance are becoming more and more important. In the future, responsibility for these aspects will be increasingly transferred to parties in the building sector.

4 The realisation of required energy performances in building processes

2.1 Introduction

In the Netherlands the energy consumption of new buildings is subject to performance based legislation. The energy performance is based on the Energy Performance Coefficient (EPC). The lower the EPC, the more energy efficient the building should be. Ever since its introduction in 1995, the EPC-value has become stricter. At that time, the compulsory EPC-level was 1.4. Since then it has been lowered regularly down to a value of 0.8. It is likely that the level will be lowered again in the near future. In the framework of the Energy Performance of Buildings Directive, 30 European EN-norms are being developed to establish the methods on which the energy performance is calculated. In the Netherlands, the norm has been compulsory for almost 15 years now and the evaluation of its effects on the actual energy used in dwellings was useful to determine the efficiency of this type of regulation. In this section the effect of the Dutch EPC regulation on the energy use for heating in dwellings built after 1995 is described.

2.2 Methods and data

This section is based on a study (Itard e.a., 2009) for a group of stakeholders (Government, real estate developers, the construction sector and the umbrella organisation for housing associations) that signed a covenant (The Spring Agreement) for the realisation of new housing estates in the coming years with 50% lower energy

consumption than the current level of regulations. Therefore it is useful to get more insight in how the energy performance regulation (EPC) actually works out in practice. The EPC-value accounts for space heating, space cooling, tap water heating, and electricity needed for mechanical ventilation and lighting. Energy use for cooking and electricity consumption for white and brown goods is excluded, because they are not related to the building itself. It is evident in all cases that there is no direct match between the measured consumption and the components of the EPC. Monitoring the EPC is therefore not an easy task because it cannot be achieved by conventional monitoring (monitoring of the energy used at the meter box, getting the data either from the occupants either from energy companies).

There were only three surveys conducted during the past fifteen years to assess the effectiveness of the energy performance regulation against the actual energy used for heating. The first one by Jeeninga e.a. (2001), on a sample of 146 dwellings, the second one by PRC (2004) and Uitzinger (2004), with a sample of 649 dwellings, and the last one in 2008, by Itard e.a. (2009), on a sample of 217 dwellings. The scarcity of data and the relatively small size of the samples may be related to the difficulties of monitoring as was explained in previous section.

In the OTB sample, two sources of data were used: a survey among households in two districts in the Netherlands, and Energy Performance calculation files from municipalities and architect firms. The obtained Energy Performance files were those from the dwellings where the survey was conducted, therefore allowing to match the responds from the household survey with data on building characteristics. The survey was carried out in autumn 2008 in two districts in the Netherlands built after the introduction of the EPC. The final sample size was 217 households. The energy reported by the respondents was from the last available energy bill.

However, none of these samples may be assumed to be representative for the Dutch building stock built after 1995. There are many historical data about energy use in dwelling, but from these databases, the WoON survey is the only one accounting for a performance comparable to the EPC: the Energy Index (EI). Instead of being based on design data like in the EPC, the EI is based on inspections and the formula used for the calculation is comparable to the EPC formula. The WoON survey was carried out in 2005 by the Ministry of Housing in the Netherlands. It consisted on two questionnaires applied to the occupants, and a dwelling inspection. The energy data included in the database refers to the real gas used during one year in the dwelling. The sample of buildings built after 1995 contains 584 cases.

2.3 Results

In the OTB sample the dwellings were characterized according to EPC categories ‘no EPC’, ‘ $1.4 \geq \text{EPC} > 1.2$ ’, ‘ $1.2 \geq \text{EPC} > 1.0$ ’, ‘ $1.0 \geq \text{EPC} > 0.8$ ’ and ‘ $\text{EPC} \leq 0.8$ ’. These categories were chosen because the maximum allowed EPC value was set at 1.4 in 1996, and tightened to 1.2 in 1998, 1.0 in 2000 and 0.8 in 2006. No statistical significant differences were found between the actual energy consumption of houses with different EPC categories (for detailed statistics, see Guerra Santin e.a. 2009a., 2009b. and 2009c.). There is only a statistical difference between dwellings in the category ‘no EPC’ and dwellings with an EPC, meaning that the introduction of the EPC regulation had an effect on the actual energy use, but not the further tightening of the EPC-values. There were also no significant differences when the actual energy use per square meter was used instead of the actual energy use of the whole dwelling. Figure 1

shows the mean energy used per type of dwelling and per EPC category in the OTB sample. With the exception of corner houses, for which the decrease is significant, no

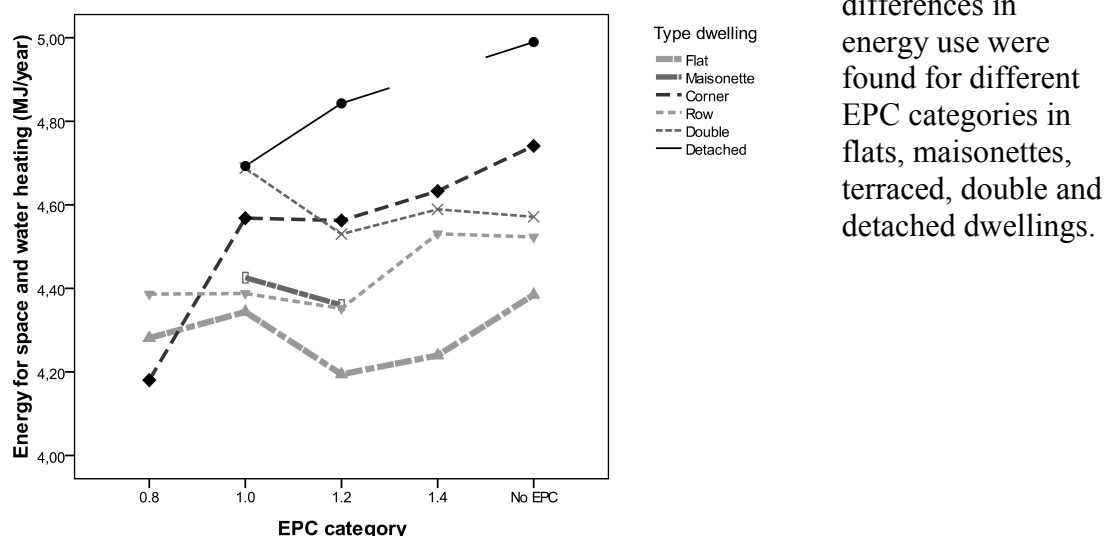


Figure 1: Energy used (mean LOG) per dwelling type and EPC level. For reasons of clarity, the 95% confidence interval is not plotted.

Similarly, analysis of the WoON database (for dwellings built after 1995) also shows the absence of correlation between the actual energy use and the Energy Index.

Furthermore, an analysis of the gas consumption of dwellings built during different construction periods seems to indicate as well that there was no significant reduction of the energy used for heating in buildings built in the periods 1996-1997, 1998-1999 and 2000-2006 (see Figure 2).

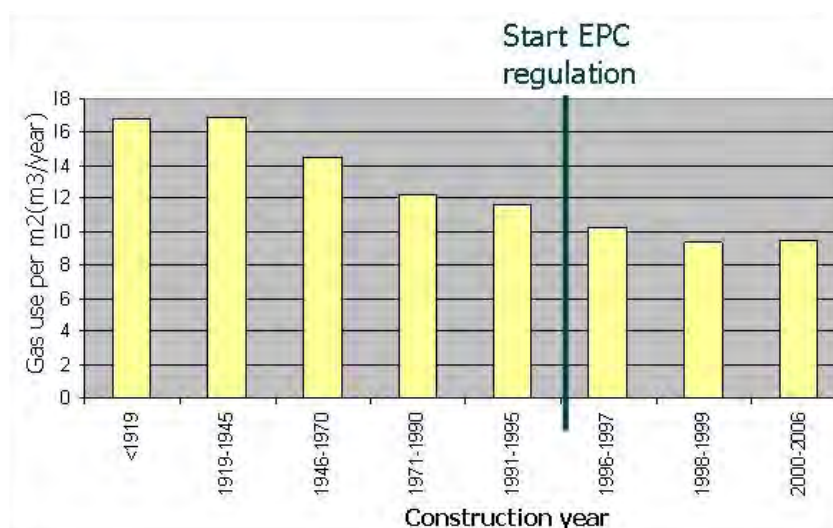


Figure 2: Average gas use per m² dwelling in 2004 [10]

In the OTB survey a correlation was found between the actual and the theoretical energy consumption, although the real energy consumption is mostly lower than the theoretical one.

Therefore we can state that although the EPC itself is not related to the actual energy use, some of its constituents are. This indicates that the formula used to calculate the

EPC, summing up space cooling, ventilation, lighting and auxiliary energy, and using neutralization coefficients introduces some kind of noise.

In a previous study using the KWR survey, which is a the predecessor of the WoON survey and contains less detailed data on the building characteristics and occupant behaviour than the WoON, we already found that 42% of the variation in energy consumption could be attributed to building characteristics, the type of dwelling being by far the most important parameter, followed by the useful living area, the construction year, the insulation degree, the presence of a thermostat, the number of heated rooms and the hours of presence home. It was found that 4.2% of the variation on energy consumption could be attributed to household characteristics and occupant behaviour. However, this study encompassed the whole Dutch building stock and not only the recent one. Because the recent dwelling stock may be assumed to be more homogenous, it can be expected that building characteristics will explain less variation and occupant behaviour more variation. The study was therefore repeated with the WoON survey with the dwellings built after 1995. In a regression analysis, taking into account only the significant parameters, 23% of the variation in energy used could be explained by the building characteristics, the most important parameters being the heat transfer surfaces and the number of bedrooms.

In the OTB survey detailed data on the dwelling characteristics were gathered through municipalities and architects, and a statistical regression model showed that the thermal properties of the dwellings could predict 19 % of the variation in energy used, heat transfer surfaces and number of bedrooms being again the most important parameters. The OTB survey was also designed to allow for a better screening of possibly relevant households characteristics and occupant behaviour and showed indeed that these were responsible for 11.5% of the variance in energy used for heating.

5 Conclusions

After almost 15 years of EPC regulation in the Netherlands, only three statistical studies were conducted to assess the effect of the regulation on the actual energy use. The samples were of limited size. In two of these samples, no statistical correlation was found between EPC-level and actual energy use per dwelling or per square meter. In the analysis of the WoON survey, representative for the Netherlands, no correlation was found between the EI and the actual energy use per dwelling and per square meter. We found that building characteristics (including heating and ventilation equipment) were responsible for 19 to 23% of the variation in energy used in the recent building stock. Household characteristics and occupant behaviour seem to be responsible for 3 to 15% of the total variance. Neither our study nor the studies found in the literature allow to state that building characteristics, household characteristics and occupant behaviour altogether are responsible for more than 38% of the variation on energy consumption of dwellings built after 1995. Therefore at least 62% of the variation in energy use is unexplained yet.

There are indications from literature that the explanation for this unexplained part could be related to buildings being realized differently than written in official documents and to HVAC services running under very different conditions than assumed on paper. A report by Nieman (2007) showed that in a sample of 154 dwellings, 25% did not meet the EPC requirements: the EPC was incorrectly calculated; nevertheless the building permit was issued. In 50% of the dwellings, the realization was not in accordance with

the data used to calculate the EPC. Gommans (2007) monitored for 17 years the energy performances of energy efficient buildings. 40% of solar boilers appeared to function poorly. Only 25% of the heat pumps reached the expected efficiency. This was essentially due to realization faults, lack of control and lack of continuous monitoring. Another study by Elkhuisen e.a. (2006) in office buildings showed that up to 28% energy could be saved by better monitoring.

Taking into account the fact that tightening the EPC did not lead to less energy use for heating and that 62% of the variation in energy use is still unexplained, it seems legitimate to be careful about a further tightening of the EPC and to search if there are more efficient means to really decrease the energy consumption of new built. This could be done by ensuring a correct realization and monitoring of the calculated performances, putting attention on the knowledge needed by contractors and on an effective building control process.

In the last 20 years there has been a clear trend in reducing the role of governments and more freedom for the private sector. This paper has emphasized the need for accurate building control. This can either be conducted by public authorities, but will more likely be based on compulsory forms of complete and accurate certificated self-control.

6 CIB Taskgroup 79

Since a few years the attention for building regulatory systems and enforcement procedures is growing. Various developments in society, politics and the construction industry have influenced changes in the systems of building control in the last 20 years. The influence of climate change and the related demands on buildings will have a very strong impact of further transformations in this field. Therefore CIB has recently established Task group 97, Building Regulations and Control in the Face Climate Change. It will be useful to compare developments of the systems of building control in countries worldwide. How do the various countries deal with the demand of deregulation and privatization of enforcement tasks? What are the effectiveness and efficiency of the systems? Could green promotion incentives be incorporated to improve the systems? How do the public regulations and enforcement systems contribute to the basic quality of buildings? Are the current systems suitable to ensure that buildings will have a very low energy demand and are comfortable, healthy and safe at the same time? What are the ideas to improve or change the systems? And what will be the impact on these systems from climate adaptation measures in the longer run?

The proposed task group will make an original contribution to the CIB Sustainable Construction priority theme. The majority of research on sustainable construction addresses the development of innovative techniques. However, there is now an awareness that it is equally important that mechanisms are developed which can effectively implement the new techniques on a large scale. These will include building regulations which incorporate appropriate incentives for the promotion of the green agenda and which therefore stimulate the use of new techniques. Equally, systems of building control will also be required which are capable of monitoring the effective and accurate design and construction of the buildings. These and similar issues will be addressed by the proposed task group.

The proposed task group will work closely with CIB working commissions W113 (Law & Dispute Resolution) and W108 (Climate Change and the Built Environment) as well as with CIB task group TG69 (Green Buildings and the Law).

Working programme

In the first phase an international research group will be established, specializing in the development of systems of building regulations and building control in the face of climate change. This will be achieved by publicizing the establishment of the task group through W113 membership and other Working Commissions, BEL-NET, CEBC and IRCA and other national and international regulation and building process groupings. Expressions of interest will then be invited from task group members and an appropriate mix of academics and practitioners from across the world will be selected to form the working core of the group.

Within 12 months: Initial review of the law and literature across key countries will be completed. Contributors for a published work will be selected, which will detail and analyze the systems in a range of countries worldwide and form a basis for comparative analysis. The exercise leading to this will inevitably be an interdisciplinary one, drawing upon science, building processes and disparate regulatory disciplines. It is the intention, not just to examine the body of topic specific literature, but also to review the literature relating to underpinning theories and principles. The result will be a valuable and readily accessible resource that can be made available to support future research activity in building regulations and building control research.

Within 18 months: Contributions for the published work, detailing and analyzing the situation across various countries will be completed. Editors work on comparative analysis to accompany these. It is important that detailed analysis of the individual contributions is undertaken and this will be carried out by the proposers and published along with the selected outputs.

Within 20 months: Comparative work will be published. The ultimate goal of the task group will be to produce a CIB publication containing a selection of research outputs authored by prominent building regulatory scholars across the world. This will outline the state of developments of systems of building regulations and enforcement procedures in the face of climate change. It is anticipated that important lessons can be drawn from this work by legislators, actors in the building processes and governments and that individual nations will be able to learn from the experience of others who are at different stages in the process of developments.

7 References

- Elkhuizen P., Scholten J.E., Rooiackers E., 2006, Quality Control of HVAC services: evaluation of existing instruments and a vision for the future, in Dutch, TNO bouw/Halmos report for Senter Novem.
- Duijm, F., Hady, M., Ginkel, J. van, Bolscher, G.H. ten, 2007, Gezondheid en ventilatie in woningen in vathorst; onderzoek naar de relatie tussen gezondheidsklachten, binnenmilieu en woningkenmerken, Amersfoort, GGD Eemland.
- Gommans L.J., 2007, Energy performances of energy efficient buildings, in Dutch, TVVL magazine, September 2008, pp.18-24

- Guerra Santin O., Itard L. & Visscher H. 2009a., Effect of energy performance regulations on the energy use for space and water heating in residential building in the Netherlands, submitted to Building Research & Information, November 2009.
- Guerra Santin O., Itard L. & Visscher H, 2009b. 'The Effect of Occupancy and Building Characteristics on Energy Use for Space and Water Heating in Dutch Residential Stock' Energy and Buildings 41 (1223–1232)
- Guerra Santin O., Itard L., 2009c. Occupant behaviour in residential buildings in the Netherlands: Determinants and effects on energy consumption for heating, submitted to Building Research & Information in November 2009.
- Itard L., Meijer A., Guerra Santin O., 2009, Consumentenonderzoek Lenteakkord, in Dutch, OTB/TU Delft report, Delft.
- Heijden, J.J., Visscher H.J., Meijer, F.M., 2007, Problems in enforcing Dutch Building Regulations, Structural Survey, Special Issue (mid 2007).
- Jeeninga H., Uytendinck M. and Uitzinger J., (2001). Energy Use of Energy Efficient Residences, in Dutch, Report ECN & IVAM.
- Kuindersma, P., Ruiter, C.J.W., 2007, Onderzoek naar de woonkwaliteit van het binnenmilieu van nieuwe woningen, Utrecht, Adviesbureau Nieman.
- Meijer, F.M., Visscher, H.J., Sheridan, L., 2002, Building Regulations in Europe, part I: A Comparison of systems of building control in eight European countries, Housing and Urban Policy Studies 23, Delft (Delft University Press).
- Meijer, F.M., Visscher, H.J., 2006, Deregulation and privatisation of European building-control systems Environment and Planning B: Planning and Design, Volume 33 – 4, pp. 491–501.
- Meijer F and Visscher H (2008), Building regulations from an European perspective, in Proceedings of COBRA 2008 - The construction and building research conference of the Royal Institution of Chartered Surveyors, RICS, London.
- Nieman, 2007, Final report Housing quality indoor environment in new built dwellings, in Dutch, report Wu060315aaA4.PK, Vrom inspectie Regio Oost, Arnhem.
- Pedro J B, Meijer F and Visscher H (2009), Comparison of tasks and responsibilities in the building control systems of European Union countries, in Proceedings of COBRA 2008 - The construction and building research conference of the Royal Institution of Chartered Surveyors, RICS, London, p. 1339-1355.
- PRC Bouwcentrum, 2004, EPC and energy consumption of new built dwellings, Novem report, Kompas, Utrecht 2004.
- Visscher, H.J., Meijer, F.M., 2007, Certification of building control in the Netherlands, Building Research Journal, volume 55 – 1/2.
- Uitzinger J., 2004, Analysis of EPC and energy use in dwellings, in Dutch, IVAM/SenterNovem report.
- WoON Energie, 2009, Woononderzoek Nederland, module energie, VROM.
- USP Marketing Consult, 2007, Faalkosten in de bouw.

Managing urbanisation and environmental protection in Australian cities – approaches for integrating biodiversity and urban growth in Sydney

Peter Williams¹

¹Faculty of the Built Environment,
The University of New South Wales,
Sydney, NSW, 2052,
Australia

Email: p.williams@unsw.edu.au

Abstract:

This paper examines the approaches taken to incorporate biodiversity conservation in the management of urban growth in Sydney and more broadly in New South Wales, Australia's most populous state. Problems associated with managing Sydney's growth – particularly from the intersection of dealing with perceived property rights and the protection of natural resources such as biodiversity – are identified. The issues are illustrated through significant State Government development decisions relating to the retention of biodiversity in the new growth areas of Sydney. The singular reliance on traditional 'command and control' regulatory approaches as both a cause and ineffectual solution to the problems faced in biodiversity conservation is highlighted. Newer 'market based' mechanisms which are being introduced (e.g. biobanking) or should be adopted (e.g. transferable development rights), as well as management at the strategic level (e.g. biodiversity certification), are considered. This paper argues that to better integrate biodiversity conservation in Australian cities a mixed approach be adopted in which a number of tools are utilised – and that this needs to occur in the context of a sound overarching strategic planning framework. This constitutes a hybrid approach involving a 'fixed' strategic spatial plan informing statutory-based regulation primarily through zoning and other development controls, augmented by a range of market based tools implemented through statute and common law measures such as conservation covenants.

Keywords:

Biodiversity, planning tools, property rights, Sydney's Growth Centres, urban growth management.

1 Introduction

As a country with a federal system of government, jurisdictional responsibility in Australia for planning, environmental protection and land management issues generally resides with its constituent states and territories (Farrier and Stein, 2006, p.12).

Australian states and territories have relied primarily on a regulatory-based statutory planning framework, derived from the traditional British system, to implement planning and land use policy. Australian planning systems are statutory-based in that they are prescribed by various state and territory legislation, including a dominant *planning act* (in New South Wales the *Environmental Planning and Assessment Act 1979*). Such statutes in turn impose a series of controls or regulation through a system of land use zoning, development standards, building regulations, subdivision controls, heritage provisions etc. The dominant source of government planning regulatory power is thus statutory law in the form of overriding planning legislation and subsequent delegated or subordinate legislation such as planning instruments, ordinances and regulations. In addition case law – especially in the form of judicial interpretation of relevant legislation and judicial review of planning decisions – also plays a key role, given Australia's legal tradition as a common law country.

Although the statutory and regulatory emphases remain in Australia – particularly in the form of land use zoning – more recently planning approaches influenced by United States systems of financial and planning incentives have emerged to complement this traditional 'command and control' regulation hegemony. Part of a self-styled 'smart regulation' package, these seek to give Australian planning systems greater flexibility through the use of market-based mechanisms and financial incentives. These tools include planning bonuses, green offsets, and the acquisition of development 'rights' pertaining to realty. Often these also involve the utilization of traditional common law mechanisms such as covenants and easements. One sphere of application of this hybrid mix of planning approaches and tools is the protection of biodiversity and other natural resource values in areas subject to urbanization.

Problems associated with managing the urban growth of Sydney – particularly from the intersection of dealing with perceived property rights and the protection of natural resources such as biodiversity – are identified in this paper. The singular reliance on traditional 'command and control' regulatory approaches as both a cause and ineffectual solution to the problems faced in biodiversity conservation is highlighted and the broader issues are illustrated through several significant case studies.

Expanding this theme, the paper examines the approaches taken to incorporate biodiversity conservation in the management of urban growth in Sydney and more broadly in New South Wales (NSW), Australia's most populous state. These approaches include better informed strategic planning for the identification of areas of high biodiversity value, the implementation of protection mechanisms through statutory and regulatory means, green belts and urban growth boundaries, land acquisition, transferable development rights, green offsets in the specific form of *biobanking*, and at a strategic level *biodiversity certification*.

A critique of the role of these approaches as applied to the integration of biodiversity and urban growth management within the framework of the NSW planning system and the context of entrenched perceived property rights is presented in this paper. Market-based approaches which seek to work with property rights and the property market – transferable development rights, biobanking and biodiversity certification – in addition to 'fixed' strategic plans informing traditional regulatory mechanisms such as zoning

and other development controls, augmented by common law tools such as covenants and easements, are seen as pivotal and are advocated in this paper.

2 Context – *City of Cities* and Sydney’s Growth Centres Formatting of Main Headings

The context of this analysis of the need for the utilization of an appropriate suite of tools to better integrate biodiversity conservation with continuing urbanization is the current metropolitan strategy for Sydney, *City of Cities – A Plan for Sydney’s Future*, released in December 2005 (DoP, 2005a). *City of Cities*, is “probably the most comprehensive planning strategy that Sydney has had since its first strategy” in 1951 (Searle, 2006, p.553). However, it was arguably developer-led, being devised in response to heavy lobbying by the property industry, particularly the Property Council of Australia, for a new strategy (PCA, 2004). It certainly bears the hallmarks of a blueprint to satiate developer desire for the further – and possibly the complete – urbanization of the Sydney basin, with the developer vision “writ large across the *City of Cities* strategy and its supporting documents” (Searle, 2006, p.533). All major Australian cities have released metropolitan strategic plans in recent years, and so the challenges examined in this paper are not unique to Sydney, and its findings have implications beyond Sydney.

The *City of Cities* strategy is intended as the plan for Sydney over the 25 years between 2006 and 2031. During this time the city’s population is forecast to increase by 1.1 million people, from 4.2 million to 5.3 million. To accommodate this predicted population growth and an anticipated fall in average household size, it is estimated that 640,000 new homes will be required. This target translates to 445,000 new dwellings projected for the existing areas of Sydney, consistent with an ongoing policy of urban consolidation, and 195,000 in greenfield areas, 135,000 of which were to be located in the North West (NW) and South West (SW) Growth Centres (that is, two new urban release areas), and 60,000 in other greenfield areas (DoP, 2005a). An additional 25,000 dwellings were originally projected to be built in the Growth Centres between 2032 and 2041, giving a total dwelling capacity in the Growth Centres of 160,000 dwellings; this was subsequently adjusted to 181,000 dwellings following the abandonment of proposed ‘green zones’ by the State Government in 2006 (further discussed below).

A pertinent action of the Metropolitan Strategy in relation to housing is the application of sustainability criteria for new urban development (*Action C1.2*). Here, proposed land release areas are to be assessed against sustainability criteria and infrastructure funding. As a precursor, a general qualitative assessment of all land identified for release within the Growth Centres was provided by the then NSW Sustainability Commissioner, Professor Peter Newman in November 2004 (Newman, 2004). This assessment was based on eight sustainability criteria, with a ranking assigned to each criteria ranging from ‘Poor’, ‘OK’, ‘Good to Best’, depending on how well the Sustainability Commissioner believed that the criteria “have been addressed in terms of global best practice for land development and also in terms of accepted practice in Sydney” (Newman, 2004, p.1).

Criteria 2: Environmental Protection, aims “to protect and enhance biodiversity, air, water and agricultural land.” Within this criterion the biodiversity benchmark is to

“save core biodiversity values and enhance natural ecosystem of the bioregion”; for water quality it is to ‘maintain and improve waterway health’; and for agricultural land to “ensure important agricultural land is conserved.” For new land release areas this criterion rates as “‘good’ to ‘best’ practice as one of the major features of the area is the new ways that the environment will be protected however air and water quality limits are approaching so any development has to be very clean” (Newman, 2004, p.4).

Despite the positive ratings of the sustainability criteria of natural resources and environmental protection, two points are of concern. First, in his report the Sustainability Commissioner “raises the question of whether the Land Release areas are needed at all. Is it possible to somehow stop Sydney growing or at least prevent any further fringe growth?” (Newman, 2004, p.4). Second, it appears that *additional* urban development outside the identified Growth Centres may be approved if it meets the eight sustainability criteria, (Searle, 2006, p.556), which would undermine the reason for the Growth Centres in terms of the objectives of the sustainability criteria such as minimizing Sydney’s ecological footprint and to protect and enhance biodiversity, water and agricultural land. Evidence of developer pressure for further land releases outside the designated Growth Centres release areas includes recent preliminary investigations by the Department of Planning into new releases, such as Macarthur South (Keneally, 2009), and Badgerys Creek (DoP, 2009). This is a significant development and places heavy pressure to extend the apparently inexorable growth of Sydney.

3 Policy failure, the demise of Sydney’s green zones and smart regulation

From the perspective of protection of areas of biodiversity value in and around Sydney’s growth centres, the current Metropolitan Strategy and the tools drafted into its implementation arguably represent a dramatic example of policy failure in the face of perceived property rights. Failure of measured strategic spatial planning, and in its stead the formulation of ad hoc, reactionary ‘policy on the run’, was evident in the response by the NSW Government to the politics of property rights exerted by land owners and developers in the Growth Centres.

On 3 November 2005, a media release issued by the office of the New South Wales Minister for Planning, the Hon Frank Sartor MP announced the scrapping of two proposed ‘green zones’ in the NW and SW urban growth centres (Sartor, 2005). This ‘green overlay’, designed to preserve existing non-urban land for aesthetic, biodiversity conservation, recreation and agricultural purposes, covered 8,400 hectares in the land release areas, and a further 14,000 hectares outside the growth centres boundary. The decision to abandon these green zones or areas – formally described as Landscape and Rural Lifestyle Zone (LRLZ) under the Sydney metropolitan strategy – was taken, stated the media release, “following widespread public consultation”. Reasons given for the decision were basically two-fold. First, the Department of Planning had received more than 3,000 submissions on the Growth Centres plans over a four-month exhibition period. It was clear, stated the Minister, that “the proposed LRLZ caused widespread concern and confusion, with nine in ten written submissions objecting to the new zone, which affected more than 7,000 properties” (Sartor, 2005). Many landowners complained about a perceived loss of property values and development rights (DoP,

2005c). Second, it was argued that the environmental benefits of the proposed green zones were limited, because 45% of the land identified for the LRLZ zone had already been cleared (Sartor, 2005).

Putting aside the issue of the poor quality of Departmental mapping and lack of 'ground-truthing' resulting in the misidentification of appropriate quality green space, the clear message was that public objection to the green zones was the primary reason for their demise, as both cleared and uncleared green areas were abandoned. This public objection rested on the expectation (whether reasonable or otherwise) that landowners' land – whose current zoning was not residential – in and around the NW and SW Growth Centres would be urbanised, with the windfall gain accruing to the property owners that this land use conversion process entails. As described in the news media at the time, the 'dumping' of the green zone on Sydney's fringe occurred after "a backlash from landowners angry their properties would not be considered for housing subdivisions" (Goodsir, 2005a, p.3), with fears that "land values in some areas will plummet as a result..." (Goodsir, 2005b, p.9).

One clear message from this episode is the role played by property rights and concomitant development expectations or rights in opposing – and ultimately determining – public policy designed to protect the environmental and natural resource values of the south-west and north-west fringe of Sydney. This role was admitted by the Minister in an earlier media release (9 September 2005) when he announced a review of the LRLZ and stated that "the green zones were never intended to change people's existing land use rights" (DoP, 2005b). It should be pointed out however, that the green zone landowners were expecting more than their existing use rights. Rather, they wanted a right to develop or use their land in way that they were not presently entitled, that is, for residential purposes. This has two significant implications. First, this right that was perceived to pertain to non-urban land does not exist even in land already zoned residential, since development consent is first required before residential subdivision and development can proceed. Second, landowner insistence on, and State Government accedence to, such rights, can only lead to speculation in areas in and around the growth areas not zoned residential, which has indeed since occurred (Keneally, 2009). Recognition of these implications was acknowledged in the *Sydney Morning Herald* the next day when it reported:

Developers and groups representing thousands of aggrieved landholders yesterday applauded the State Government's decision to walk away from a green zoning proposal that had denied property owners the right to cash in on future housing estates.

(Goodsir, 2005c, p.11)

A further aspect of the State Government's decision in regard to the abandoned green zones was the announcement that it would attempt to retain some environmental aspirations by focusing on protecting the best sections of vegetation and waterways in the two Growth Centres. This new approach, developed in consultation with the (then) Department of Environment and Conservation, created four new zones into which land was classified: flood-prone, urban-capable, urban edge and conservation (Sartor, 2005). Significantly, this approach focuses on biodiversity certification and relies on a new

environmental offsets or biobanking scheme, under which developer payments are used to conserve areas of bushland, further discussed below.

Planning implementation of the growth centres component of the Sydney metropolitan strategy was deferred to the making of a specific statutory plan in the form of a state environmental planning policy (SEPP), which finalised the release area boundaries and the constituent land zones and controls. Work on a draft of this SEPP progressed throughout 2005, and major changes were made to its envisaged land use zones following the State Government's decision to abandon the proposed green zones. A 'final' version of the draft SEPP – minus the now moribund green zones – was released in January 2006 for public exhibition and comment. Subsequently, on 28 July 2006 the Minister for Planning gazetted *State Environmental Planning Policy (Sydney Region Growth Centres)* 2006. Yet the SEPP as gazetted was modified from that placed on public exhibition, following 750 submissions from members of the public, industry and State Government agencies (DoP, 2006a).

Opinions on the nature of the modifications to the SEPP greatly diverged. On the one hand an update on the Sydney metropolitan strategy issued by the Department of Planning was emblazoned with the heading 'More Green Spaces for Western Sydney' (DoP, 2006b), whilst on the other the *Sydney Morning Herald* more pessimistically – and accurately – reported that 'Housing eats our next park' (Davies, 2006). These two contrasting views are discussed in turn below.

The increase in green infrastructure in Western Sydney announced by the Minister involved the rezoning under the SEPP of a long-standing undeveloped 560 ha parcel of land owned by Airservices Australia, a Commonwealth Government agency. Located at Shane's Park, in Sydney's north-west, the land is the site of a former international radio transmission station and about 80% was already listed on the Commonwealth's Register of the National Estate. However the site was to be merely rezoned by the State Government from one open space designation to another. Conversely, the SEPP also sought to reduce by over one-half, the amount of land proposed for the Rouse Hill Regional Park. The park, which surrounds Rouse Hill House, one of Sydney's earliest properties, had been planned since 1989 and a further commitment to create the 115-hectare park was given in 1997. In a decision that "outraged heritage experts, local councils and even the Government's own advisers on its metropolitan strategy", the State Government decided not to acquire Stage 2 of the park, with the Planning Minister arguing that the \$120 million to be paid for the 81-hectare site "was too much to pay for views" (Davies, 2006). Instead, the land was designated under the SEPP to be released for housing. Public opposition to this proposal was so intense however that the Minister quickly reversed his decision, announcing the entire area of the park would be protected by restoring its earlier open space zoning (DoP, 2006d).

Since the publication of *City of Cities* further potential for undermining of this plan has been evident in the form of developer pressure for land to be released for urban development *outside* the designated growth centres. In particular, one major development company with extensive land holdings in the Macarthur South/Appin area south of Sydney has actively lobbied the NSW State Government for this land to be added to Sydney's release areas (Frew and Snow, 2008). This is part of an area

investigated for urbanisation under the 1988 Sydney metropolitan strategy *Sydney Into Its Third Century* (DEP, 1988) but subsequently deferred due to water and air pollution problems (DoP, 1993; Holliday, 1998; Vipond, 2001). Clearly, such a release would also undermine a fundamental component of *City of Cities* of limiting Sydney's urban expansion to 2031 to the designated growth centres and hasten the urbanisation of the Sydney basin. Following initial consideration, the Government recently announced the deferral of further investigation of Macarthur South as its development is presently unviable due to prohibitive infrastructure costs (Keneally, 2009). It should be pointed out that similar challenges are facing other Australian cities. For example, the recent and controversial expansion of Melbourne's urban growth boundary into its 'green wedge' areas (Buxton and Goodman, 2002; Buxton and Scheurer, 2007), has strong parallels with the Sydney situation.

This brief vignette reveals a number of key factors that must be taken into consideration in contemporary growth management on the rural-urban fringe of Australian cities and towns. First is the deficiency, on their own, of traditional command and control mechanisms such as land use zoning and planning restrictions to guarantee the protection of non-urban land. Second is the apparent inevitability of continued urban sprawl unless appropriate growth management policy responses can be crafted and implemented to counter this apparent biodiversity-damaging form of urbanisation. Third is the role – rightly or wrongly – that claims to property rights play in land use planning and development decisions. Fourth is the reluctance of government to rely, otherwise than de minimus, on the public purse to protect non-urban land (for example, through land acquisition for the provision of green infrastructure). The fifth factor – argued here to be an inevitable conclusion given the previous four considerations – is the role that newer alternative mechanisms such as smart regulation and market-based instruments that operate within the context of property rights can play, particularly in the context of seeking to ensure that natural resource and environmental values are protected in the face of the pressure and expectations of continued urban expansion. It is in the context of these factors that some complementary mechanisms for planning policy implementation – specifically the integration of biodiversity conservation into managing the growth of Australian cities – are now considered with Sydney as a case study.

4 Transferable development rights

The concept of property rights is integral to schemes such as transferable development rights (TDR). TDR is described as a property rights-based tool since it recognizes, and seeks to work with, a development right which is perceived as one of a number of rights accruing from ownership or other interest in property (Wiggins, 1988). Fundamental differences however, can be identified in the practical application and consequences of the concept of property rights. Specifically property rights-based tools have, depending on the approach taken, been ascribed as constituting an example of a market-based or economic instrument, or alternatively a manifestation of free market environmentalism (Gunningham and Grabosky, 1998).

Yet in essence, TDR is one type of planning tool that seeks to compensate landowners whose development rights have been restricted by regulation. Compensation is

achieved by allocating to those owners an amount of development that may be transferred from the restricted site to another site (Bindon, 1992). Fundamentally, development rights are severed from a parcel designated for protection ('sending area'), and the severed rights are transferred to a parcel in an area where additional development is permitted ('receiving area') (Johnston and Madison, 1997, p.365). The scheme thus allows more development on the receiving parcel while reducing or preventing development on the donor parcel. Under such a program, the development rights of the sending parcel may be either sold by that owner to the owner of the recipient parcel, or transferred directly from the donor to the receiving site if they are under common ownership. The number of development rights that can be transferred depends on how many development-rights 'credits' a planning authority allocates and how much it allows in areas designated for growth (Daniels, 1999, p.224).

Within the modern system of formalised land tenure the bundle of rights that constitute land ownership are often consolidated in the hands of a single 'owner'. By allowing voluntary acquisition and conveyance of specific rights for specific uses, trading in partial interests offer this more flexible and refined alternative to a strictly regulatory approach or acquiring full ownership rights (Wiebe and Meinzen-Dick, 1998). Development rights have been viewed as one of a number of rights embodied in the ownership interest in property. These development rights have been classified as a real property interest, which entitles the owner of a fee simple interest to deal with the land as the owner wishes, subject only to government regulation, principally through zoning (Arnold, 1992). However, the right to transfer development rights is not ordinarily part of the bundle of rights that comes with land ownership: because in Australia at least there is no right to develop land except within the terms of planning instruments. Government must therefore enact specific legislation to legalise the sending of a building right from one parcel to another (Daniels and Bowers, 1997). Once legislatively sanctioned, an owner may separate and transfer one of the rights incidental to ownership whilst retaining the other rights (Arnold, 1992). In the US the acquisition and conveyance of partial interests to land has proven to be a popular, flexible and effective tool for land use and conservation policy.

TDR has several attractions to commend it – which revolve around its 'respect' for property rights. TDR is a (hybrid) market based mechanism under which developers pay for preservation in return for additional development potential. Where a TDR scheme is in place, a developer buys development rights, with zoning provisions identifying the number of additional units allowed in designated receiving areas. TDR is therefore effective when the TDR option is more profitable than the non-TDR option for landowners and developers. The motivation for utilizing this scheme is the ability to sell and transfer development rights – thereby increasing residential densities in targeted sites – and yet retain land and appropriate uses in receiving areas. Schemes such as the purchase or transfer of these interests or rights have allowed public agencies and private non-profit conservation groups to influence the use of public and private land without incurring the political costs of land regulation or the full financial costs of outright land acquisition (Wiebe and Meinzen-Dick, 1998). It is a voluntary approach to influencing land use, by offering landowners and farmers financial incentives for environmental conservation, restoration, and preservation.

In Australia there is no such inherent right to develop land; rather a property owner may have the right to seek development consent, after the granting of which, development for the specific purpose approved can legally commence before the consent lapses after a prescribed period. Nevertheless, in practice the Australian experience is that a landowner may have certain development expectations based on the applicable statutory planning controls. Implicit in the controls is a perceived probability of gaining approval for a certain type and quantity of development (Bindon, 1992). As a consequence, the fundamental principles behind the US model has been recognised and adopted by several local councils in Australia that have established TDR systems (Williams, 2004). These include heritage conservation in Sydney, Adelaide, Melbourne and Brisbane (Ryan, 2004), protection of the Mount Lofty Ranges near Adelaide, provision of open space and conservation reserves in Gosford (NSW), urban growth management in Wellington (NSW) and protection of the Illawarra Escarpment near Wollongong, south of Sydney (Williams, 2004).

Yet the apparent reluctance for more widespread use of TDR as a planning tool remains. This is despite TDR being identified as a tool worthy of consideration, for example, in the NSW *Plan First* planning system reforms a decade ago (DUAP, 2001). Three reasons can be advanced to explain this situation. First, the utilization of market based tools is still relatively recent in Australia. There has been a tradition of reliance on ‘command and control’ regulation in Australia, which is quite different to the history of market based tools in the US and bargaining for planning gain/negotiated planning agreements in Britain. Second, there is a lack of understanding of the TDR mechanism by planning decision-makers (both politicians and planners). Third, there is ongoing legal uncertainty and impediments surrounding TDR. Evidence of the present legal impediments to the more widespread adoption of a TDR scheme in NSW include expression of doubt by the NSW Land Environment Court about the legality of TDR schemes (see for example *Leighton Properties Pty Limited v North Sydney Council* [1998] NSWLEC 39), concerns raised by a Commission of Inquiry regarding the transparency of Wollongong City Council’s TDR scheme for the Illawarra Escarpment (Commission of Inquiry, 1999), and ongoing reluctance by the NSW Parliamentary Counsels Office to support draft statutory plans produced by local councils that seek to include TDR provisions.

5 Biodiversity certification and biobanking

Biodiversity certification and biobanking are two relatively new planning tools in NSW. Biodiversity certification was introduced with the passage of the *Threatened Species Legislation Amendment Act 2004* (NSW). It is designed to integrate threatened species management into the strategic planning stage through the auspices of the formulation of environmental planning instruments (EPIs), rather than being mired within the ‘trench warfare’ of assessment under the development application process. Under this scheme, environmental biodiversity studies are required to be prepared up front, and planning options for development and conservation methods identified and evaluated. A draft EPI is then prepared and forwarded to the Minister for Climate Change and the Environment who then decides whether to grant biodiversity certification. The outcome of the process is that a developer does not need to prepare a species impact statement, and biodiversity does not comprise part of the environmental assessment of the

proposed development. Significantly, the first EPI to receive biodiversity certification in NSW was *State Environmental Planning Policy (Sydney Region Growth Centres) 2006* which contains the main statutory planning controls for Sydney's Growth Centres.

Biobanking is an example of an offsets scheme. Under an offset arrangement, industries or resource users are given the choice of either offsetting the damage they cause or paying an authority to do it on their behalf. The provision of an offset is a mandatory requirement or condition of the granting of approval to undertake development with potentially adverse environmental impacts. The arrangements operate partly through regulatory mechanisms such as permits or approvals, and partly through a market-based system, which allows one property owner who undertakes some form of environmental restoration to sell offset credits to another owner or industry seeking approval to undertake development. Thus an offset arrangement is clearly a hybrid of regulatory and market-based instruments.

The *NSW Biodiversity Banking and Offset Scheme* was introduced under the *Threatened Species Conservation Amendment (Biodiversity Banking) Act 2006* (NSW). The scheme seeks to provide an additional, market-based, mechanism to assist in conservation management. It aims to achieve more predictable development and conservation outcomes by guiding development to appropriate places, and to promote private land conservation through income generating opportunities for landowners who provide biobank sites. Landowners create credits by establishing biobank sites and earn income from managing land for conservation. The scheme aims to be comprehensive – the biobanking provisions include requirements for biobanking statements, creation of biodiversity credits (calculated using published assessment methodology), detailed regulations (including cost recovery), establishment of a BioBanking Trust Fund, BioBanking public registers and enforcement provisions (DECC, 2007).

The NSW BioBanking and Offsets Scheme seeks to address the loss of biodiversity by enabling landowners to establish biobank sites to secure conservation outcomes and offset impacts on biodiversity caused by development. Conceptually, this is achieved through the use of an 'improve or maintain' test for biodiversity values, which means avoiding significant biodiversity conservation areas and offsetting impacts in other areas (DECC, 2007, p.4). The offsets are measured in terms of credits, using the published BioBanking Assessment Methodology (DECC, 2009), and developers participating in the scheme are required to meet this improve or maintain test based on the impact of their proposed project.

The BioBanking Scheme has four key components (DECC, 2009b):

1. Establishing biobank sites on land through biobanking agreements between the Minister for Climate Change and the Environment and participating landowners. A biobanking agreement is similar to a covenant and is attached to the land title. It runs with the land, and generally will have effect in perpetuity so as to offset the impacts of development on biodiversity values.
2. Creating biodiversity credits for management actions that are carried out, or proposed to be carried out, to improve or maintain biodiversity values on

biobank sites. The biobanking assessment methodology is the tool used to determine the number of biodiversity credits that may be created for these management actions.

3. The trading of credits, once they are created and registered.
4. Enabling the credits to be used to offset the impact of development on biodiversity values. The assessment methodology is the tool that is used to determine the number and class of credits that must be retired to offset the impact of a development and ensure that the development improves or maintains biodiversity values.

BioBanking in NSW is still in its infancy – indeed at the time of writing no biobanking agreements had been listed publicly (DECCW, 2010a) although several have been shortlisted as either biobank sites, development sites or joint biobank/development sites – and biodiversity certification has presently not extended beyond the Sydney Growth Centres (DECCW, 2010b). Nonetheless, some observations may be made in relation to both these biodiversity conservation tools. Both approaches require an ‘improve or maintain’ outcome for biodiversity values, which is difficult – if not impossible – to achieve given the high conservation value of the remaining biodiversity and ecological communities in the Sydney Region. Disagreement exists in relation to the identification of appropriate offset ratios – i.e. the ratio of conservation land to offset developed land – with this generally well in excess of a simple 1:1 ratio. Problems have also arisen with using biobanking and biodiversity certification to justify the loss of areas of high biodiversity – something which DECC seeks to avoid and the Land and Environment Court has had to adjudicate in the case of specific residential developments in Sydney (see for example *Sanctuary Investments Pty Ltd & Ors v Baulkham Hills Shire Council* [2006] NSWLEC 733). Concerns have also been expressed by some local councils in the Sydney Region about the location of biodiversity offsets. Such councils have argued that the offset sites should be located in the same local government area that the development is occurring, whereas some State Government agencies believe that the funds generated under the BioBanking Scheme could be better used to conserve larger areas of land outside the Sydney Region, where land acquisition costs are cheaper. Further, problems have arisen in situations where developers have sought to offset the loss native vegetation on development sites with biobank sites containing ecological communities of inferior conservation status, contrary to the principles of the BioBank Scheme (DECC, 2009c).

Finally, the attraction of offering biodiversity offsets – generally as individual land dedication agreements outside the formal Biobank Scheme – has been used as a bargaining chip by developers to persuade State Government to give favourable consideration to unscheduled urban releases, particularly in the Lower Hunter region of NSW (Creagh and Munro, 2007). One such land dedication agreement which attracted significant media attention involved the brokering of the agreement, and subsequent granting of planning approvals, by a former Minister for Planning for proposed residential developments at Catherine Hill Bay and Gwandalan in the Lower Hunter Valley, situated to the north of Sydney. The Ministerial approvals were subsequently quashed by the Land and Environment Court on the grounds of apprehended bias and reliance on irrelevant considerations (*Gwandalan Summerland Point Action Group Inc v*

Minister for Planning [2009] NSWLEC 140). Pertinently also, the invalidated approvals relied on land dedication agreements that were outside the framework of the Government's own BioBanking and Offsets Scheme and were inconsistent with the Department of Planning's strategic spatial plan for the region, the *Draft Lower Hunter Regional Strategy* (Williams, 2009).

6 Conclusion

Faced with the rising influence of the property rights movement (along with the contemporary problem of 'down-zoning' land as witnessed in the proposed green zones in the Sydney Growth Centres), the challenge to land use managers and planners has been to devise planning mechanisms which respect the integrity of private property on the one hand, and yet still achieve planning policy objectives on the other. It is in this context that more creative mechanisms such as TDR and biodiversity certification and biobanking need to be considered as urban growth management tools. In the case of TDR the issue of the lack of understanding of this mechanism by decision-makers (including planners) and the present legal uncertainty surrounding its application need to be resolved – the latter by legislative action. In the case of biodiversity certification and biobanking the main issue relates to the unintended consequences of these tools. This includes their use to gain approval for inappropriate development in terms of undermining both the promotion of orderly land releases and the protection of areas of high biodiversity value within the Sydney Region.

Part of the reason for these problems lie in the fact that market based tools are still very much in their infancy in the NSW planning system. To better integrate biodiversity conservation with managing the growth of Sydney, a hybrid framework which utilizes sound metropolitan planning, strategic or planning phase biodiversity certification, statutory-based regulation such as zoning, and market-based mechanisms including TDR and biobanking reinforced by conservation covenants or agreements, should be adopted.

7 References

- Arnold, C. (1992) Transferable Development Rights – A Planning Tool for the Preservation of Heritage Buildings. *Environmental and Planning Law Journal*, 9(6), pp 458-474.
- Bindon, J. (1992) Transferable development rights: A review. *Australian Planner* 30(3), pp 136-141.
- Buxton, M. and Goodman, R. (2002) *Maintaining Melbourne's Green Wedges. Planning policy and the future of Melbourne's green belt*. School of Social Science and Planning, RMIT University, December 2002.
- Buxton, M. and Scheurer, J. (2007) Density and Outer Urban Development in Melbourne. *Urban Policy and Research* 25(1), pp 91-111.
- Commission of Inquiry. (1999) *The Long Term Planning and Management of the Illawarra Escarpment, Wollongong Local Government Area*. Office of the Commissioners of Inquiry for Environment and Planning, Sydney.
- Creagh, S. and Munro, C. (2007) Sartor overrides conservation advice. *Sydney Morning Herald*, 19 March 2007, p.8.

- Daniels, T. (1999) *When City and Country Collide: Managing Growth in the Metropolitan Fringe*. Island Press, Washington, D.C.
- Daniels, T. and Bowers, D. (1997) *Holding Our Ground: Protecting America's Farms and Farmland*. Island Press, Washington, D.C.
- Davies, A. (2006) Housing eats our next park. *Sydney Morning Herald*, 2 August 2006, <http://www.smh.com.au/text/articles/200/08/01/1154198137553.html>, viewed: 03/08/2006.
- DECC (2007) *BioBanking. Biodiversity Banking and Offsets Scheme. Scheme Overview*. NSW Department of Environment and Climate Change, Sydney.
- DECC (2009a) *BioBanking. BioBanking Assessment Methodology and Credit Calculator Operational Manual*. NSW Department of Environment and Climate Change, Sydney.
- DECC (2009b) *The BioBanking framework*. NSW Department of Environment and Climate Change, Sydney, <http://www.environment.nsw.gov.au/biobanking/biobankframework.htm>, viewed: 11/08/2009.
- DECC (2009c) *Principles for the use of biodiversity offsets in NSW*. NSW Department of Environment and Climate Change, Sydney, <http://www.environment.nsw.gov.au/biocertification/offsets.htm>, viewed: 11/08/2009.
- DECCW (2010a) *Biobanking agreements*. NSW Department of Environment, Climate Change and Water, Sydney, <http://www.environment.nsw.gov.au/bimsprapp/ExportAgreement.aspx>, viewed: 14/08/2010.
- DECCW (2010b) *Notices of biodiversity certification*. NSW Department of Environment, Climate Change and Water, Sydney, <http://www.environment.nsw.gov.au/biocertification/notcert.htm>, viewed: 14/08/2010.
- DEP (1988) *Sydney into its Third Century: Metropolitan Strategy for the Sydney Region*, NSW Department of Environment and Planning, Sydney.
- DoP (1993) *Sydney's Future. A Discussion Paper on Planning the Greater Metropolitan Region*. NSW Department of Planning, Sydney.
- DoP (2005a) *City of Cities – A Plan for Sydney's Future*, NSW Department of Planning, Sydney, <http://www.metrostrategy.nsw.gov.au/dev/uploads/paper/governance/index.html>, viewed: 12/06/2006.
- DoP (2005b) *Statement on Landscape and Rural Lifestyle Zone*. Media Release, NSW Department of Planning, Sydney, 9 September 2005.
- DoP (2005c) *New ground rules for the north-west and south-west land release areas*. Fact Sheet, NSW Department of Planning, Sydney, http://www.metrostrategy.nsw.gov.au/dev/digitalAssets/1350_1130983357783_051103%20fact%20sheet%201.pdf, viewed: 11/08/2009.
- DoP (2006a) *Managing Sydney's Growth Centres Fact Sheet – New State Environmental Planning Policy (Sydney Region Growth Centres) 2006*, NSW Department of Planning, Sydney, http://www.metrostrategy.nsw.gov.au/dev/digitalAssets/1794_1154575936782_FACT%20SHEET%20-%20FINAL%20SEPP.pdf, viewed: 09/09/2006.

- DoP (2006b) More Green Spaces for Western Sydney. *Metropolitan Strategy – What’s New*, NSW Department of Planning, Sydney, Friday 28 July 2006, <http://www.metrostrategy.nsw.gov.au/dev/ViewPage.action?siteNodeId=59&languageId=1&contentId=528>, viewed: 09/09/2006.
- DoP (2006c) *North West Growth Centre – Transitional Land Information Presentation*, 20 February 2006, NSW Department of Planning, Sydney, http://www.metrostrategy.nsw.gov.au/dev/digitalAssets/1610_1141791228574_Transitional%20Land.pdf, viewed: 11/11/2006.
- DoP (2006d) *Statement on Rouse Hill Regional Park*, Media Release, Minister’s Office, 3 August 2006, NSW Department of Planning, Sydney, http://www.planning.nsw.gov.au/mediareplan/fs20060803_393.html, viewed: 11/11/2006.
- DoP (2009) *New South Wales Department of Planning Submission to the General Purpose Standing Committee No.4. New South Wales Upper House Inquiry into Badgerys Creek Land Dealings and Property Decisions*, 25 September 2009, NSW Department of Planning, Sydney, <http://www.planning.nsw.gov.au/LinkClick.aspx?fileticket=S5yrQdsPCwY%3d&tabid=366>, viewed: 14/08/2010.
- DUAP (2001) *Plan First: Review of plan making in New South Wales*, NSW Department of Urban Affairs and Planning, Sydney.
- Farrier, D. And Stein, P. (2006) *The Environmental Law Handbook*, 4th edn, Redfern Legal Centre Publishing, Sydney.
- Frew, W. and Snow, D. (2008) Developers push into fringeville. *Sydney Morning Herald*, 29 November 2008, p.11.
- Goodsir, D. (2005a) Angry landowners winning the war on green zone. *Sydney Morning Herald*, 9 September 2005, p 3.
- Goodsir, D. (2005b) Landholders rush zoning help line. *Sydney Morning Herald*, 6 September 2005, p 9.
- Goodsir, D. (2005c) Owners’ joy as state dumps green zone plan, *Sydney Morning Herald*, 10 September 2005, p 11.
- Gunningham, N. and Grabosky, P. (1998) *Smart regulation: designing environmental policy*, Oxford University Press, Oxford.
- Holliday, S. (1998) Metropolitan Planning and Demography: Sydney as a Case Study, in R. Freestone (ed) *The Twentieth Century Urban Planning Experience*, Faculty of the Built Environment, University of New South Wales, Sydney.
- Johnston, R. and Madison, M. (1997) From Landmarks to Landscapes: A Review of Current Practices in the Transfer of Development Rights. *Journal of the American Planning Association* 63(3), pp 365-378.
- Keneally, The Hon Kristina, MP (2009) *Planning Starts For New Housing And Employment Land in Sydney*, Media Release, 22 July 2009.
- Newman, P. (2004) “New Land release Areas and Sustainability Criteria”, letter from the NSW Sustainability Commissioner to DIPNR, 11 November 2004, http://www.metrostrategy.nsw.gov.au/dev/digitalAssets/844_1102557381520_Sustainability%20Commissioner%20on%20Land%20Release.pdf, viewed: 13/12/2008.
- PCA. (2004) *Metro Strategy: A Property Industry Perspective*, Property Council of Australia, Sydney.
- Sartor, The Hon Frank, MP (2005) *New Ground Rules for Green Space in Growth Centres*, Media Release, 3 November 2005.

- Searle, G. (2006) Is the City of Cities Metropolitan Strategy the Answer for Sydney? *Urban Policy and Research*, 24(4), pp 553-566.
- Ryan, S. (2004) Conservation through Development: The Potential for Transferable Development Rights in Queensland. Paper presented at *2004 QELA Conference – Carrots, Sticks and Toolkits*, Cairns, 12-14 May.
- Wiebe, K.D. and Meinzen-Dick, R. (1998) “Property rights as policy tools for sustainable development. *Land Use Policy* 15(3), pp 203-215.
- Wiggins, D. (1988) The limits of planning: What’s in a name? Paper presented at the *Planning Law and Practice Short Course*, School of Town Planning, The University of New South Wales, Sydney.
- Williams, P. (2004) Use of transferable development rights as a growth management tool. *Environmental and Planning Law Journal* 21(2), pp 105-123.
- Williams, P. (2009) Land bribes’ – a cause for alarm in wider planning practice?, *New Planner*, 81, p.25
- Vipond, J. (2001). Regional Planning in NSW. *Australian Planner* 38(3/4), pp 121-127.

Sustainable development and legal ownership in Danish urban planning law

Helle Ina Elmer, Ph.D.¹

¹ Agency for Spatial and Environmental Planning,
Ministry of the Environment, Denmark

Email: hinel@blst.dk

Abstract:

Sustainability is as a concept incorporated in Danish planning law. The theories concerning sustainability as an environmental concept presupposes that the incorporation of the concept in Danish planning law will cause a change in the concept of ownership in Danish planning law. The correctness of this bold statement depends on how the two involved concepts are perceived. In this paper the basis of the two concepts in the Danish jurisprudence are analyzed and the legal consequences are drawn.

Keywords:

Danish Planning Law, Ownership, Sustainable Development.

1 Introduction

Both the concept of sustainable development and the concept of ownership is in Danish jurisprudence contested concepts and when sustainable development is said to have an impact on the way the property owner is exercising his or her rights of ownership (Tuomela 1995; Basse 1995; Pagh 1998), then it challenges the conceptual legal system. The question of what influences what and how – if at all – hasn't yet been analysed.

For this paper, I have chosen to focus on the Danish Planning Act and only the part of spatial planning concerned with urban planning. The Danish Planning Act¹ states the national planning authority of the Minister of Environment, regional spatial development planning, municipal planning, municipal local planning, zoning and rural land management, supervision and administrative provisions, and appeals and legal proceedings. The choice of perspective is motivated by the fact that the Planning Act is low rated, when it comes to looking at sustainable development and high rated, when it comes to the legal impact on ownership.²

¹ Act no. 388 of 6 June 1991 cf. consolidated Act no. 937 of 24 September 2009. Throughout the paper the abbreviation Planning Act will refer to this consolidated Act unless otherwise stated. The Danish – and other Scandinavian countries' – legal systems are in classic comparative law-approach characterized as a mixed legal system: Denmark has a constitution and an extensive amount of legislation but some areas of daily legal life are left unregulated such as the basic conception(s) of ownership. The Scandinavian countries did not follow the Codification-movement in Europe, and that is why Danish law is not bound by certain principles to the extent shown by the German, French, Dutch, Austrian and Swiss Civil Codes. Please note that by Danish law compulsory acquisition is subject to full compensation.

² Please note that this paper only states my personal opinion.

2 Ownership in Denmark

2.1 Ownership in Danish Jurisprudence

In Danish jurisprudential account of ownership³ is defined as an exclusive negative bounded possession that is a possession in every respect, which is not limited by law or private act of will. Some scholars describe the possibility of dividing the powers of the owner subject as a flexible feature of ownership, by which they refer to that a temporary reduction in the owner's power over the object etc. returns to the owner after completion of absences i.e. letting a flat or pledging an asset. The emphasis on ownership elasticity may seem superfluous if it is used only in conjunction with a temporary absence of certain powers. However, as a principle of keeping the rights ownership together – for whatever reason - elasticity is attributed an independent status in the following historical explanation and conceptual clarification. At the same time the patient reader will be introduced to the basic ideas behind the property in planning terms.

Until late 18th century Danish law like many other legal systems is based on natural law (Nørregaard 1784). There were also drawn a distinction between the king's overall ownership and the ownership of his or her loyal subjects. Ownership was not only ascribed to physical persons but also to trust and foundations, whereby large sums of money were placed under 'the dead hand'. In early 19th century the exclusiveness of ownership was questioned probably for many reasons but also because quite a few trust and foundations had lost their purpose, which they were aimed to support. So, the king and later the government would have the right to dissolve such futile trust and foundations (Ørsted 1815). During the 18th century a shift in the legal structure of ownership occurs, so natural law in the classical sense is replaced by a philosophically justified conceptual structure. Even the king lost the exclusive ownership and the state got sovereignty instead. In addition, most descriptions of ownership use Roman law terminology and conceptual design (Aagesen 1881; Björne 1995; Björne 1998; Tamm & Vogt 2009).

Ownership was viewed as a unified concept but at what point in time does this unified right transfer from say a seller to a buyer? This question led to several interesting constructions of implicit and tacit nature (Tamm 1976; Björne 2002), but one scholar, Carl Torp, went the opposite direction, suggesting that ownership consisted of relations, and some relations may transfer at one point, others at another point in time (Torp 1892). It all depends on the legal nature of transfer and to what has been agreed upon by the persons involved. With the fragmentation of concept ownership there was no longer any affiliation of rights, so how was it possible to uphold the concept of ownership in a legal sense? Two prominent figures in Danish jurisprudence, Frederik Vinding Kruse and Alf Ross, disputed about this for almost thirty years (Björne 2007).

Frederik Vinding Kruse, an empiricist, based his theory on a very broad definition of ownership/property. He maintained that four core rights have to be present in order to accept the presence of ownership: 1) the inwardly command and use of an object, 2) the outwardly rights of subrogation, 3) the right of posting the object as collateral, and 4)

³ Hohfeld has shown that a right of ownership may be subdivided into different legal modalities. However, there are diverse notions of rights and ownership, so a common approach of rights is required. The term 'right' is then the generic term used in this paper.

the right of bequeath. Limited rights may also be categorized as ownership, when either one of the following combinations are used: (1,3,4), (1,2,4), (1,2,3) but not (1,4). It is not quite clear, what the substance of 1) is, but 2)-4) were also used in the 19th century description of ownership (Vinding Kruse 1929, 143ff). Being an empiricist Vinding Kruse did not intend to describe ownership as a substantial right by the four core rights. Rather, he wanted to provide a qualitative based criterion for the presence of ownership (in case of transfer), and thus, his concept of ownership may be described as being an inherently qualitative structure.

Alf Ross, positivist, understands in his early writings ownership as functionalistic, which makes him of largely the same kind as Carl Torp, the early spokesman for a relative ownership. Ross finds that the presence of ownership is solely based on a quantified structure: If the owner has transferred all but one right, then the owner is no longer the owner no matter what sort of right he would have left to himself.

Later on in his career – in the early 1950's – where the enthusiasm for logical positivism had already cooled off, because of its heavy demands on scientific language and the demarcation of science/pseudoscience, Ross chose to extend the quantified structure of ownership and concluded that ownership is without semantic reference, and it may be dissolved into a list of alternative elements of law and cumulative legal consequences. Ross argues at the same time for elasticity as part of the concept of ownership, not as a description of a possible temporarily bilateral agreement but as a defence against the disintegration of the concept.⁴ Ross never said much about elasticity but it renders an explanation that takes into account his presumption of ownership – at all. Alf Ross's concept of ownership may then be described as being an inherently quantified structure.

To complete the picture it is also possible to encounter a nominalistic approach to legal ownership in Danish jurisprudence. These proponents i.e. Knud Illum don't accept any inherent structure but they can't totally disregard the concept of ownership either. Instead, they use ad hoc explanations when necessary to link a right to the concept of ownership. The ad hoc explanations may be the specific agreement in question, company/corporation law, theories about juristic persons or pure and simple the statute in question.

What commenced as a quest for dismantling the unified concept of ownership to solve the problem of transfer of ownership, ends up leaving us with two possible ways of talking about ownership – again.⁵ The determination of ownership will often be resolved by using the conditions for the transfer of ownership.

⁴ Alf Ross concept of ownership was challenged a few years later, when the Icelandic government wanted part of the collection of Icelandic sagas back to Iceland from a Danish research trust in Copenhagen. Ross didn't find that the trust met the requirements of an owner, but the Supreme Court did cf. *Ugeskrift for Retsvæsen* 1967.22 H.

⁵ Having analyzed the theories of ownership by Frederik Vinding Kruse and Alf Ross Peter Pagh concludes, that in any way ownership is still presupposed (1998, 140).

Inherent	structure	Nominalistic structure
Quantitative criteria	Alf Ross	Knud Illum
Qualitative criteria	Frederik Vinding Kruse	

Figure 1. Distinction of inherent and nominalistic structure.

(Source: Elmer, 2008)

The transfer of ownership may, when viewed as a qualitative/quantitative transfer of rights, be sufficient to determine the legal owner and hence the right to exclusive control and use apart from limitations imposed by law, legal doctrines etc. However, the distinction between a nominalistic and an inherent structure of ownership is not sufficient to clarify possible limitations on the rights of the owner. The relation between uncompensated regulation and compulsory acquisition⁶ will be of interest no matter how the issue of ownership is approached, because it presupposes a specific problem or conflict. It should be noted, though, that the use of theories about the inherent structure of ownership may preclude some areas of arguments otherwise available for the proponent of nominalistic ownership.

2.1.1 Ownership in Danish spatial planning – a historical overview

Factors like migration from rural to urban areas, stagnation in the construction industry, poor economic situation in the wake of the first World War accentuated the need for better spatial planning for towns and cities and their surroundings and not just a regulation of building construction in order to prevent health hazards like unsuitable constructions, fire and epidemics. Experts in various fields of work had since 1916 (Madsen 2009) tried to persuade the politicians to act on this urgent need for spatial urban planning, and in 1924 a bill⁷ was introduced leaving it to the municipalities to decide whether a local plan was needed or not.

During the reading of the bill a change was made to it at a crucial point: the grounds for compensation. The passed and enacted bill had become almost worthless as an instrument of planning. The Town Planning Act of 1925⁸ made it possible for every land- and house owner to claim compensation, if they were affected by a local plan proposal, a adopted final local plan, an interim injunction and dispensations from and changes of local plans. There was no deadline in the Town Planning Act for submitting the claim. Finally, the owner may reject the compensation offered and instead demand a compulsory purchase of the property.⁹

The Town Planning Act did not recognize the possibility of non-compensational regulation. Frederik Vinding Kruse was a strong proponent for compulsory urban planning, and he found the Town Planning Act to be a discontinuation of Danish jurisprudence (1929, 235 and 306). Thus, the Town Planning Act received a lot of

⁶ By Danish law compulsory acquisition is subject to full compensation.

⁷ Bill of 20 November 1924 concerning town planning read in Landstinget (the Upper House) of the bicameral system of that time.

⁸ Town Planning Act no. 122 of 18 April 1925.

⁹ Section 9 and 10 of Town Planning Act.

criticism, because the municipalities couldn't estimate the cost of adopting a local plan for an area.¹⁰

Through the 1930's Frederik Vinding Kruse developed a more specific view on planning and ownership in connection with non-compensational regulation and compulsory acquisition. The key feature here is, whether the local plan presupposes a transfer of ownership rights or just limitations imposed by law or a local plan. Almost needless to say that transfer of ownership rights would result in compensation, while limitations imposed by a local plan wouldn't. When the Town Planning Act was amended in 1938¹¹ it was Frederik Vinding Kruse's described relationship of ownership and compensation that constituted the system used in the Act (Blomquist 1991 cited Forchhammer 1940; Meyer 1952). At the same time it became compulsory to adopt local plans for boroughs and towns with more than 1000 inhabitants.

However, far from all statutes and their amendments within the area of environment and planning has had the benefit of a discussion concerning the concept of ownership and the distinction between non-compensational regulation and compulsory acquisition. In some cases due to by a rather unclear approach to the presupposed idea of legal ownership; in other cases due to pragmatic approaches to compensation and political negotiations, when debating the specific bills. It should also be taken in to account that the upgrading and expansion of the administration during the construction of the welfare state has resulted in the transfer of tasks from traditional legal regulation and the judiciary and so changing from reactive individual defenses against the state to the administrations proactive protection of the citizens (Blomquist 1991). The specific ways to view the concept of ownership may also vary depending on the subject to which it is applied within the area of environmental law (Pagh 2006).

2.1.1.1. The concept of ownership in the Planning Act

The Planning Law of 1992 incorporated several statutes and in terms of ownership and compensation the Act re-enacted the principle of non-compensational regulation (Anker 2006). Spatial planning according to the Planning Act is limitations imposed by a regional development plan, a municipal plan or a local plan. However, the limitations may be so severe and close to a transfer of rights that the owner may claim compensation on the grounds that the specific planning is compulsory acquisition. Such a claim would be warranted in the Danish Constitution,¹² but if the municipality can refer to adequate and legal grounds for the specific plan, the claim will probably not stand up in court (Anker 2006; Andersen & Christiansen 1989; Friis Jensen 1978).¹³

The Planning Act also warrants compulsory acquisition for the execution of a local plan, when the acquisition is appropriate, necessary and reasonable.¹⁴ Until August 2009 it was also possible to warrant compulsory acquisition based on a municipal plan for

¹⁰ General explanatory notes to bill of 9 February 1938 concerning town planning.

¹¹ The amended Town Planning Act no. 181 of 29 April 1938.

¹² Section 73.

¹³ This point of view has been tried by the courts (Ugeskrift for Retsvæsen 1975.743 H, Ugeskrift for Retsvæsen 1979.938 Ø, Ugeskrift for Retsvæsen 1987.849 H and Ugeskrift for Retsvæsen 2009.188 H) but so far only one unpublished decision (14 November 1984) from the Western Division of the Danish High Court held that a local development plan imposed too many restrictions on the owner.

¹⁴ Planning Act Section 47.

urban development i.e. to acquire rural land for future urban development. The compensation would be calculated on the basis of rural land instead of urban zone.¹⁵ Proponents of this warrant for acquisition pointed out that the municipality could acquire rural land at a low cost some time before the urban development would commence.¹⁶ Opponents of the warrant for acquisition referred to the exclusive right of the owner and compulsory acquisition was for the execution of public interest i.e. roads, schools etc., but not for the interest of the municipality, which would be the receiver of the surplus, when reselling the acquired land to urban developers.¹⁷

The eligibility to make compulsory acquisition based on a municipal plan, however, were removed by an amendment to the Planning Act in 2009¹⁸ on the grounds that it was a violation of the right of the owner to execute the future plan for urban development and hence receive the surplus. At the same time municipalities were urged to make greater use of urban regeneration and densification of existing settlements, rather than to plan for projects involving rural land – planning principles, which falls well in line with considerations of sustainability in urban development.¹⁹

2.1.2 Concluding remarks

Compulsory acquisition was a part of the municipalities' toolbox in the 1970's and the beginning of the 1980's (Revsbech 1995). It was a time with focus on the collective and general interest in spatial planning and not the individual interests and rights of the land- and house owners. By 2002 focus has shifted; now the ownership of the individual owner of property requires the utmost attention and protection. Spatial planning is still for the benefit of the community, but it can't be done at the expense of the individual property owner. Ideologically, we have a conception of ownership that equals a qualitative inherent structure – but even more so, because the argument seems to be that ownership is a substantial right. The inherent structure in Figure 1 is then exceeded as to the presupposed structure of the concept. This may be done by letting the rhetorical speech flourish, so for now it still seems probable to place the concept of ownership within figure 1. It is beyond doubt, though, that the ownership of the property owner has received a higher priority since the change of government in 2001.

3 Sustainable Development in Denmark

3.1. Jurisprudential and ethical considerations

The thought of having ethics forming the basis for sustainable development in a jurisprudential context is by itself a controversial subject in Danish jurisprudence. It is due to the fact that Danish jurisprudence has a longstanding influence by logical positivism, which as a scientific paradigm leaves no room for ethical judgements. However, some scholars have argued that ethics should have a more predominant position, when trying to understand, how the concept of sustainable development may

¹⁵ Compulsory Acquisition for Urban Development Act no. 123 of 1 April 1980.

¹⁶ The Danish Parliament's first reading of bill no. 94 of 21 December 1979 concerning Compulsory Acquisition for Urban Development.

¹⁷ Op.cit.

¹⁸ Amendment of the Planning Act no. 512 of 12 June 2009.

¹⁹ General explanatory notes to bill no. 135 of 19 February 2009.

help us securing the future for us and generations to come. The concept of sustainable development is considered to hold four core ethical issues:

1. Intergenerationality, where equality for each generation is sought and respect for each generation's own need (Bugge 1995; Bugge 2008; Pagh 2006).
2. Intragenerationality, where the distribution of wealth should be equal between the rich and the poor countries and between the rich and poor citizens in each nation (Bugge 1995; Bugge 2008; Pagh 2006).
3. The relation between human beings and nature and the consumption of non-renewable resources must be precautious and we have to think in alternative measures (Bugge 1995; Bugge 2008; Pagh 2006).
4. The question of uncertainty about values and priorities, when the future level for protection is predetermined (Bugge 1995; Pagh 2006).

Hans Christian Bugge²⁰ (1993) refers to the first and second principle as a double claim for justice, which strengthens the ethical value of and set the standards for the two principles in competition against the two other principles. It also seems to work as a way to emphasize and ensure the effectiveness of the principles, when the ethical principles are used as a guideline for future transformation into legal regulation. In (1995, 35f) Hans Christian Bugge points to second generation rights as a way to ensure the right to education, health, social services, housing etc. and he also discusses the possibility of using third generation rights in order to ensure that nature and perhaps also future generations are taken into account.

In her earlier work Ellen Margrethe Basse (1995; 2001) elaborates on Hans Christian Bugge's conceptual description and specific argues for a substantive side to the concept of sustainable development. She draws a parallel to the concept of justice, which purports a right for the citizens to an environment of a certain quality and a right to be heard. However, in her later work Ellen Margrethe Basse (2006) seems to concur with the mainstream scholarly opinion that sees the exercise of discretion in national administrative legal systems emphasizing environmental considerations and a weakening of the considerations governing the legal division of the administrative fields of responsibility (Pagh 2006; Christensen 1997; Andersen 2002; Garde 2004).

Veit Koester finds that the concept of sustainable development is difficult to apply in a jurisprudential context (1999). Veit Koester regards the concept as a broader term compared to other more substantial principles, by which it is implemented. Similar views are expressed by Philippe Sands (2003) and Michael Bowmann (2000).

Henrik Zahle (1995) contemplates the implications of incorporating a legal protection of nature in the Danish Constitution. Henrik Zahle was far from being an anti-environmentalist, but he did found that some difficulties had to be overcome first: how do we delimit the concept of nature, how do we delimit it to the concept of environment or do the two concepts overlap each other, who or what is the right-holder and last but not least shouldn't it be a responsibility instead of a right? Unlike the Norwegian Constitution the Danish Constitution has not been amended, but Ole Kristian Fauchald

²⁰ Hans Christian Bugge is a Norwegian professor of law but he has had a tremendous impact on the theoretical development in Danish environmental law.

shows in (2007) that the legal impact of introducing a protection of the environment in a national constitutional document has had little bearing on the judicially outcome.

Finally, it should be noted that although the theory of John Rawls may play an important part, when it comes to understand the priorities of the principles in the concept of sustainable development, Danish environmental legal scholars haven't picked up the idea of reflective equilibrium as a way to describe the more substantial side of sustainable development. It may be due to fact that substantial analysis of concepts isn't part of the jurisprudential paradigm in Denmark.

3.1.1. Concluding remarks

The concept of sustainable development is not a legal concept, from which we may derive rights and obligations. Rather, it is a general or background concept, which works its way into the legal system by challenging the way we interpret the statutes (Tuomela 1995). However, the materialization of the concept through more operational concepts has more focus on the statutes securing the public access to information.

3.2. Sustainable development in Danish Planning Law

The WCED's definition of sustainable development was incorporated in the Planning Act as part of an environmental reform; a reform that has been criticized for being unnecessary, not political motivated and more to the benefit of the administrative system (Garde 1993) and Ellen Margrethe Basses comments on this (1993). Nevertheless, as a result of the reform the concept of sustainable development became a constituent part of the purpose of the Planning Act and other environmental laws. Accordingly, the Act states²¹ that

the overall planning synthesizes the interests of society with respect to land use and contributes to protecting the country's nature and environment, so that sustainable development of society with respect for people's living conditions and for the conservation of wildlife and vegetation is secured.²²

Further, the Planning Act²³ aims especially towards 1) appropriate development in the whole country and in the individual administrative regions and municipalities, based on overall planning and economic considerations; 2) creating and conserving valuable buildings, settlements, urban environments and landscapes; 3) that the open coasts shall continue to comprise an important natural and landscape resource; 4) preventing pollution of air, water and soil and noise nuisance; and 5) involving the public in the planning process as much as possible.

As a result of the UN Conference on Environment and Development (UNED) of 1992 the Ministry of Environment agreed with LGDK²⁴ and CCDK²⁵ to launch a joint

²¹ Planning Act, Section 1, subsection 1.

²² David Breuer has translated The Planning Act in Denmark, Consolidated Act No. 813 of 21 June 2007 <http://www.blst.dk/NR/ronlyres/4BDB0226-256C-4999-91EC-1FF90AC148BF/50763/planlovenpengelsk2007.pdf>

²³ Planning Act, Section 1, subsection 2.

²⁴ Local Government Denmark (LGDK) is the interest group and member authority of Danish municipalities.

Agenda 21-campaign aimed at local and regional authorities, citizens, companies, institutions etc. The actual work of Agenda 21 was and still is based on commitment, partnerships and voluntariness, but Folketinget (the Danish Parliament) passed an amendment to the Planning Act in order to supervise the implementation of the Agenda 21-strategies. Regional councils and municipal councils are obliged to publish a report on their strategy for sustainable development with information on how this work will be carried out in a holistic, interdisciplinary and farsighted manner and how the public, enterprises, organizations and associations will be involved in this work.²⁶ The reports are also submitted to the Minister for the Environment. Every four years, the Minister for the Environment submits a report²⁷ on Agenda 21 work to a committee established by the Folketing. The report is produced in cooperation with the associations of regional councils and municipal councils.²⁸

The concept of sustainable development in the Danish Planning Act is based on the WCED definition, which retains the anthropocentric approach, but what about the intergenerational perspective? One may say that spatial planning - being a future-related activity - it presupposes the generations to come, and by doing so spatial planning provides the settings for the lives of future generations.²⁹ It is, however, difficult to accept this point of view, because if we are to respect the fulfilment of future generations own needs, it may be limited by historic settings. The intragenerational perspective has very little to do with spatial planning in the Planning Act-sense in so far that the problems associated with the intragenerational question are seen as allocation of resources based on a position on legal ownership, leasing etc. In Danish planning law it is not possible to determine the legal ownership neither in regional development plans, municipal plans nor in local plans. If, however, it is possible to describe the regulation of legal ownership as taking social sustainable development into consideration, then it has been among the top priorities of the political agenda for the past decade.

3.2.1. *Governmental Action Plans and a national strategy*

The method of choice for initiating sustainable development in different areas of interest has in Denmark up till 2002 been the so-called action plans. The terms of an action plan may be very abstract ideas, but it may also be followed up by very specific proposals. As a source of law action plans are political documents subjected to a parliamentary reading in order to resolve political differences.³⁰ Thus, action plans aren't binding on administrative decisions (acts), they are politically binding on future legislative decisions made by ministers of the government, when the action plan states specific proposals. Helle Tegner Anker (1993) has put forward the idea that action plans are in some ways similar to preparatory work like explanatory statements and so they may be regarded as binding on administrative decisions. There are, however, no authorities in this area, so in regard to action plans as a source of law the traditional approach to is

²⁵ County Council Denmark (CCDK) was an interest group and member authority of Danish county councils. The county council-structure has been changed in 2005 from 14 counties to five regions.

²⁶ Planning Act, Section 33a.

²⁷ The latest Ministerial account of Local Agenda 21-strategies is published in November 2009.

²⁸ Planning Act, Section 33b.

²⁹ Peter Pagh (2006) holds that the concept of sustainable development operates directly and independent within the temporal dimension, but the scope of the opinion is not clear in regard to Planning Act.

³⁰ This is due to the political situation in Denmark. In Folketinget (the Parliament) the 179 members are among 13 political parties and none of the parties have absolute majority.

upheld, see also Henrichsen (1997, citing additional litteratur). This approach does rule out the possibility that bills passed by Folketinget but not yet received the Royal Assent in some cases may be regarded as a binding source of law (Blume 1994) in a court of law.

3.2.1.1. Specific Plans of Action

The understanding of the sustainable development in the Danish Planning Act used the Action Plan for Environment and Development (Handlingsplan for miljø og udvikling 1988) as a starting point. This action plan did not apply a different understand of sustainable development than the one laid down in WCED. Cities are still the resource intensive and polluting organism and the focus of urban planning is restricting and preventing excessive use of fossil energy, water and non-recyclable materials (Handlingsplan for miljø og udvikling 1988, 124).

Following the EU-Commission's plan of action for sustainable urban development in the European Union³¹ the Danish Urban and Housing Ministry³² launched a Action Plan for the City of Tomorrow.³³ Following this plan of action the concept of sustainable development was in an urban planning context delimited but in a more proactive manner. Urban planning was to be much more than damage control. Five focus areas were highlighted: 1) a favourable trade environment, 2) increased democracy and participation, 3) residential social work and mixed forms of ownership, 4) the city's spatial layout, and 5) the framework and organization of urban policies. These five focus areas were analyzed and elaborated by a expert committee – Trade and Urban Policy Commission (Erhvervs- og Bypolitisk Udvalg 2001).³⁴ In the final report submitted by the Commission the development of the city is split up into the diverse city and the sustainable city. The diverse or multiple city holds different initiative to facilitate mixed forms of ownership and the sustainable city holds ideas concerning recycling, urban regeneration and urban renewal.

The distinction between the diverse and the sustainable city reflects legal possibilities for imposing changes of far-reaching implications. This became clear, when in 2001 a government was formed by the Right Wing Party (Venstre) and The Conservatives (Konservativt Folkeparti) with the Danish People's Party as the supporting party to ensure majority in Folketinget. A national twenty years strategy for sustainable development was presented in 2002 "Common Future – Balanced Development" (Fælles fremtid – udvikling i balance, 2002). The strategy had adopted quite a few of the previous objectives from the Committee's report but the role of commerce had been given a noticable priority – even mixed ownership was now prioritized as self-ownership in terms of homeowners and resident owned flats. This neo-liberalistic³⁵ key issue was especially intended for the independent social housing organizations and not unlike the

³¹ (COM/98/605)

³² The Ministry was in 2001 merged with the Ministry of Social Affairs

³³ The Action Plan was also given as a Ministerial account of urban policies to Folketinget, 23 February 1999.

³⁴ Commission Report no.1397/2201.

³⁵ The former Prime Minister Anders Fogh Rasmussen was influenced by Robert Nozick's 'Anarchy, State and Utopia' and in 1993 he published the book 'Minimal State' – a neoliberal political manifesto at a time, when the Social Democrats Party, Central Democrats, The Liberal Left Party and the Christian People's Party had formed a governmental coalition. Anders Fogh Rasmussen was Prime Minister from 2001-2009 is currently Secretary General of NATO.

British model for selling local government housing to sitting tenants in the 1980's a bills was passed³⁶ to ensure that Danish social housing stock was put up for sale on favourable terms to sitting tenants. The Social Housing Organizations claimed the Act to be illegal compulsory acquisition, because they had not been offered any compensation as the purchase price, when selling a flat or house, was not paid to the organization. The claim was brought before the Danish Supreme Court,³⁷ where five judges out of nine ruled in favour of the Act being noncompulsory acquisition. Four judges gave a dissenting judgment in which they disagreed with both the result and the justification of the majority.³⁸

3.2.1.2. Action Plan motivated Amendments to the Planning Act

The present national strategy for sustainable development has also resulted in two amendments of the Planning Act:

In 2003³⁹ was the focus on the possibilities for urban regeneration of older industrial- and harbour districts into mixed urban business- and housing areas. In relation to the impact on private ownership it should be noted that the municipality may decide to hand responsibility for maintenance of the common area to an association of homeowners. The terms and delimitation of the responsibility will be outlined in a local plan for the specific area.

The second amendment was passed in 2007.⁴⁰ Property owners were now given the opportunity to enter into voluntary development agreements with the municipality on contributing to the physical infrastructure, such as squares, streets and paths through planning for urban development or urban regeneration.

3.2.2. Concluding remarks

The concept of sustainable development is a very broad concept i.e. the social housing-case verifies that. The amendment of the Planning Act, which took away the possibility of compulsory acquisition for future urban planning, wasn't motivated by an action plan and it wasn't motivated by the idea of sustainable development. The latter was more of a side-effect of the principle idea: to respect the ownership rights of the property owners.

4 A Brief Conclusion

Implementation of the concept of sustainable development in the Planning Act of 1992 hasn't directly delimited or even broadened the rights of property owner in any significant way. Nor has the specific local planning had any impact on the rights of the property owners due to sustainable development regulation. The change of government in 2001 did, however, have an impact on the way in which ownership rights and sustainable development were defined. The emphasis on social sustainable development has widened the scope for reforms in other areas of society and a deregulation of spatial planning.

³⁶ Amendment to Social Housing Act no. 485 of 2004 of 9 June 2004.

³⁷ The 7 November 2007, published in Ugeskrift for Retsvæsen 2008.378 H.

³⁸ A review of the Supreme Court's judgment, see (Jensen 2008; Elmer 2008).

³⁹ Amendment to the Planning Act no. 440 of 10 June 2003.

⁴⁰ Amendment to the Planning Act no. 537 of 6 June 2007.

Acknowledgements

I wish to thank the Agency of Spatial and Environmental Planning for making it possible for me to write this paper, and former librarian Liselotte Møller, Agency of Spatial and Environmental Planning, for her enthusiasm and competent help.

5 References

- Andersen, B, Christiansen, O, (1989), *Kommenteret kommuneplanlov*, Jurist- og Økonomforbundets Forlag, København.
- Andersen, J., (2002), *Udfyldning*, In: Gammeltoft-Hansen, H., et al, Forvaltningsret, Jurist- og Økonomforbundets Forlag, København.
- Anker, H.T., (1993), Handlingsplaner, konsekvensvurderinger og kvalitetsnormer på miljøområdet, *Juristen*, vol. 75, no. 10, pp. 426-433.
- Anker, H.T., (2006), *Planlovgivning*, In: Basse, E.M., Miljøretten, Jurist- og Økonomforbundets Forlag, København.
- Basse, E.M., (1993), Lovbemærkningers sandhedsværdi - om miljøreformen 1990-91/Johan Garde., *Juristen*, vol. 75, no. 9, pp. 383-387.
- Basse, E.M., (1995), *Bæredygtighedsbegrebet – et nyt retligt begreb*, In: Basse, E.M. (ed.), *Bæredygtighed – en retsteoretisk begrebsanalyse*, GADJura, København.
- Basse, E.M., (2001), *Miljørettens udvikling og grundlag*, In: Basse, E.M., Miljøretten, Jurist- og Økonomforbundets Forlag, København.
- Basse, E.M., (2006), *Miljørettens udvikling og grundlag*, In: Basse, E.M., Miljøretten, Jurist- og Økonomforbundets Forlag, København.
- Björne, L., (1995), *Den nordiska rättsvetenskapens historia, vol. I, Tiden före år 1815, Patrioter och institutionalister*, Rättshistoriskt Bibliotek LII, Lund.
- Björne, L., (1998), *Den nordiska rättsvetenskapens historia, vol. II, 1815-1870, Brytningstiden*, Rättshistoriskt Bibliotek LVIII, Lund.
- Björne, L., (2002), *Den nordiska rättsvetenskapens historia, vol. III, 1871-1910, Den konstruktiva riktningen*, Rättshistoriskt Bibliotek LX, Lund.
- Björne, L., (2007), *Den nordiska rättsvetenskapens historia, vol. IV, 1911-1950, Realism och skandinavisk realism*, Rättshistoriskt Bibliotek LXII, Lund.
- Blomquist, H., (1991), Jura og jurister mellem politik, eksperter og ideologi – Den danske byplanlov 1916-1940, *Retfærd*, vol. 14, no. 53, pp. 47-60.
- Blume, P., (1994), Retspolitik og lovfortolkning, *Juristen*, vol. 76, no. 6, pp. 231-235.
- Bowmann, M., (2000), *The Effectiveness of International Nature Conservation Agreements*, In: Anker, H.T., and Basse, E.M. (eds.), *Land Use and Nature protection – Emerging Legal Aspects*, Jurist- og Økonomforbundets Forlag, København.
- Brennan, A., Lo, Y.-S., (2008), Environmental Ethics, *The Stanford Encyclopedia of Philosophy* (Summer 2010 Edition), Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/entries/ethics-environmental/>, viewed 29.06.2010.
- Bugge, H.C., (1993), Bærekraftig udvikling og andre aktuelle perspektiver i miljøretten, *Lov og Rett*, pp. 485-498.

- Bugge, H.C., (1995), *The Ethics of Sustainable Development – a challenge to the legal system*, In: Basse, E.M. (ed.), *Bæredygtighed – en retsteoretisk begrebsanalyse*, GADJura, København.
- Bugge, H.C., (2008), *1987-2007: “Our Common Future” Revisited*, In: Bugge, H.C., and Voigt, C. (eds.), *Sustainable Development in International and National Law*, Europa Law Publishing, Gronningen.
- Christensen, B., (1997), *Forvaltningsret – opgaver, hjemmel, organisation*, Jurist- og Økonomforbundets Forlag, København.
- Elmer, H.I., (2008), Grænser i dansk ejendomsret?, *Tidsskrift for Rettsvitenskap*, vol. 121, no. 4-5, pp. 508-531.
- Fauchald, O.K., (2007), Forfatning og miljøvern – en analyse av grunnloven § 110 B, *Tidsskrift for Rettsvitenskap*, vol. 120, no. 1-2, pp. 1-84.
- Friis Jensen, O. (1978), *Ekspropriation og andre emner*, In: W.E. von Eyben (ed.), *Dansk Miljøret*, Akademisk Forlag, København.
- Garde, J. (1993), *Lovbemærkningers sandhedsværdi - om miljøreformen 1990-91*, GADs Forlag, København.
- Garde, J., (2004), *Saglige krav*, In: Garde, J. et al., *Forvaltningsret. Almindelige emner*, Jurist- og Økonomforbundets Forlag, København.
- Henrichsen, C. (1997), *Retssikkerhed og moderne forvaltning*. En retspolitisk studie i samspillet mellem stat, forvaltning og borger, Akademisk Forlag, København.
- Jensen, M.H. (2008), Højesterets dom i sagen om salg af almene familieboliger, *Ugeskrift for Retsvæsen*, 2008 B. pp. 214-223.
- Koester, V., (1999), Århus-konventionen om ”borgerlige rettigheder” på miljøområdet – især i et menneskerettighedsperspektiv, *Juristen*, vol. 81, no. 3, pp. 87-102.
- Koester, V., (2006), *International beskyttelse af biodiversitet*, In Basse, E.M., *Miljøretten*, Jurist- og Økonomforbundets Forlag, København.
- Madsen, H.H., (2009), *Skæv og national, Dansk byplanlægning 1830 til 1938*, Bogværket, København.
- Nørregaard, L., (1784), *Natur-rettens Første Grunde*, Gyldendals Forlag, Kjøbenhavn.
- Pagh, P., (1998), *Miljøansvar – en ret for hvem?*, Jurist- og Økonomforbundets Forlag, København.
- Pagh, P., (2006), *Lærebog i miljørettens almindelige del*, Jurist- og Økonomforbundets Forlag, København.
- Revsbech, Karsten, (1995), *Planlægningsloven – reguleringsformål eller retssikkerhed*, In: Christiansen, O. (ed.), Christensen, B., and Koester, V., *Retligt set – synsvinkler på arealanvendelseslovgivningen*, Miljø- og Energiministeriet, København.
- Ross, A., (1951), *Tû-Tû*, In: O.A. Borum and Knud Illum (eds.), *Festskrift til Henry Ussing*, Nordisk Forlag Arnold Busck, København.
- Ross, A., (1966), *Om ret og retfærdighed*, Nyt Nordisk Forlag Arnold Busck, København.
- Sands, P., (2003), *Principles of International Environmental Law*, Cambridge University Press, Cambridge.
- Tamm, D., (1976), *Fra »Lovkyndighed« til »Retsvidenskab«*. Studier over betydningen af fremmed ret for Anders Sandøe Ørstedes privatretlige forfatterskab, Jurist- og Økonomforbundets Forlag, København.
- Tamm, D., Vogt, H., (2009), *Dansk retshistorie og vestlig forfatningsudvikling*, Studieudgave, Jurist- og Økonomforbundets Forlag, København.
- Torp, C., (1892), *Dansk Tingsret*, GEC GAD, Kjøbenhavn.

- Tuomela, J. (1995), *Hållbar utveckling som miljörettens grundprincip*, In: Basse, E. M. (ed.), *Bæredygtighed – en retsteoretisk begrebsanalyse*, GADJura, København.
- Vinding Kruse, F., (1929-33), *Ejendomsretten* vol. I-V, Nordisk Forlag Arnold Busck, København.
- Zahle, H., (1995), *Naturen ind i grundloven – menneskeret eller naturbeskyttelse*, In: Christiansen, O. (ed.), Christensen, B., and Koester, V., *Retligt set – synsvinkler på arealanvendelseslovgivningen*, Miljø- og Energiministeriet, København.
- Ørsted, A.S., (1815), *Om regeringens ret til at ophæve eller forandre stiftelser, som private mænd have oprettet*, (1801), Eunomia, 1, Andreas Seidelin, Kjøbenhavn.
- Aagesen, A., (1881), *Indledning til den danske formueret*, posthumt A.C. Evaldsen and C.Goos, Gyldendalske Boghandels Forlag, Kjøbenhavn.

The Severn Tidal Barrage Project: a legal paradox? A case study looking at the possible conflict between climate change policy and environmental legislation in the UK.

¹Tilak Ginige, ²Ann Thornton, ¹Fraser Ball, ¹Natalie Sparks.

¹School of Conservation Sciences, Bournemouth University, Talbot Campus, Fern Barrow, Poole, Dorset, BH12 5BB, United Kingdom

²Lindsay Carrington Ecological Services Ltd, The Old Squash Court, Rempstone Hall, Rempstone, Corfe Castle, Wareham, Dorset, BH20 5JQ, United Kingdom.

Abstract

The proposed construction of a tidal barrage to generate electricity in the Severn Estuary between England and Wales could provide an economically attractive and environmentally acceptable way of supplying up to 7% of England and Wales's electricity consumption with low-cost, low-carbon electricity by 2020. This helping the UK government to meet its obligation under its current Climate Change Policy and specifically those contained within the Renewable Energy Strategy (2009) of achieving 15% of energy supply from renewables by 2015.

This development will however have a huge impact upon 63,000 overwintering birds; destroy protected areas of wetland; and alter the estuarine ecosystem beyond repair.

If projects like this were permitted, the UK would have to take compensatory measures to ensure the overall coherence of Natura 2000 was protected (Art. 6(3) & 6(4) of the EC Habitats Directive 92/43/EEC). Will these compensatory measures be effective considering that we do not fully appreciate the role and function of ecosystem services provided by areas such as those around the River Severn? Perhaps more importantly are we prepared to gamble that their loss will be less damaging to us than the impact from future climate change?

This then obliges us to question the weight given to conservation of species and habitats in an era of economic transformation and climate change obligation. We need to ask ourselves if we truly are at a point in our evolution where we will be making decisions of whether or not to sacrifice one good for another greater good (Alder, J. and Wilkinson, D., (1999).

Given the overwhelming legal protection afforded to this area one of the issues that needs to be explored is how such a proposal was even considered? Have we approached the moment when Article 6(4) of the Habitats Directive ‘**Imperative Reasons of Overriding Public Importance**’ (IROP) will include combating climate change? If such a project is allowed to go ahead in this area, does the prevention of climate change represent the ‘trump card’? Does this present an open door to developers wishing to capitalise upon opportunities presented by the government’s binding targets? Provided that it can be shown that the project or development contributes towards these targets, will all other environmental considerations be ignored?

Thus placing us in a paradoxical situation where the environment will be irrevocably changed in order to prevent the environment being irrevocably changed.

Keywords:

Climate Change Policy, Renewable Energy Strategy, Habitats Directive (Dir.92/43/EEC), Paradox, Balancing interests.

Introduction

It has been estimated that the offshore renewable energy industry could generate the same amount of electricity a year by 2050 as one billion barrels of oil (Clarke, 2010). If the UK were to harness just 29 per cent of its wind, wave and tidal resources it would be able to match the electricity generated by North Sea oil and gas production. Such a move would be a huge boost in cutting emissions, saving a cumulative total of 1.1 billion tonnes of carbon dioxide over the next 40 years (The Offshore Valuation Group (OVG) 2010). The OVG estimated that it would cost £443bn to harness 29 per cent of the offshore resources, but at the same time that the scheme would generate £62bn in annual exports. This figure could rise to £164bn if the government invested around £993bn – or 76 per cent of available resources.

In the UK, using tidal power or tidal range in the form of barrages or lagoons could make the largest contribution for future renewable energy. (OVG 2010) A tidal barrage works in a similar way to hydroelectric power schemes, in that it is a way of converting the kinetic energy of water into electric power (Frontier Economics, 2008). This technology would extract the energy associated with the difference between high and low tides, but out of all the renewable energy technologies it is by far the most expensive (Clarke, 2010) (OVG) 2010).

Severn Estuary

The Severn Estuary located at the south-west coast of the UK between Wales and England at the mouth of the Severn, Wye and Avon rivers has the second highest tidal range in the world after the Bay of Fundy, Canada with a mean tidal range of 4.68m at Milford Haven and 8.52m at Avonmouth (Sustainable Development Commission October 2007). The tidal range has resulted in around 200km² of inter-tidal zone in the Severn Estuary that supports a wide variety of fauna. A high tidal range, as exhibited in the Severn Estuary, is caused by the combination of estuary shape and the velocity at which the tide wave propagates causing tidal resonance. As a result, the high spring tides at Avonmouth can exceed 14m (Xia, J.Q., 2010), making it an ideal site for constructing a tidal power project.

Tidal Barrage Scheme

The idea of a barrage across the Severn estuary is not new and was first mooted during the latter half of the 19th century for the purposes of harbour creation, flood defence and as a rail transport link. Its use for energy generation was brought to the forefront of government thinking during the fuel crisis of the 1970s (Pethick et al. 2009). The Severn Barrage Committee (SBC) was set up as a result. Additional fuel issues during the 1980s, including the Chernobyl nuclear disaster and disturbances following pit closures, led the SBC to endorse a barrage scheme. As a result in 1989 a report was produced for the then Conservative government. It concluded that the Severn barrage could provide significant economic, social and environmental benefits (HMSO (1989) Energy Paper 57). However, following the privatisation of the electricity industry, the Severn barrage project was effectively 'shelved'.

In 1999 Energy Minister Ian Battle controversially refused to approve a reappraisal of the proposed barrage despite pressure from the Severn Tidal Power Group (STPG) - (a consortium of McAlpine, Balfour Beatty, Alstom, Rolls Royce, Taylor Woodrow and Carillion). However, a year later a recommendation by the Royal Commission on Environmental Pollution provided the impetus for a re-assessment of tidal energy production as a way of reducing CO₂ emissions (22nd Report: RECP, 2000).

The Kyoto protocol (United Nations, 1997) and sustainable development implications led the new Labour government to instigate a discussion on the feasibility of a Severn tidal barrage. In 2002 the STPG produced 'The Severn Barrage – Definition Study for a New Appraisal of the Project' (Taylor, 2002). This study concluded that a reappraisal of the Severn barrage was justified on several grounds: reductions in the capital building costs of the scheme; Kyoto targets for greenhouse gas emission reductions; flooding risk in the Severn basin due to

climate change which would be mitigated, in part, by the barrage.

However, by 2004, the barrage proposal was again dismissed by Industry Minister Lord Sainsbury due to the likely ‘significant effects’ on the ecosystem of the area. This was despite calls from other peers suggesting that the barrage could provide additional transport links to the region in the form of a rail crossing. Lord Sainsbury also added. "We would be required under our international obligations to recreate inter-tidal habitats in order to compensate for that [ecosystem] loss." (ENDS Report No. 348, 2004) The STPG consortium proposed that predictions of habitat loss from climate change should form the baseline of any environmental assessment - although these may be extremely difficult to make in practice.

In 2007, the Sustainable Development Commission (SDC) produced ‘Turning the Tide’, a study into the potential of tidal energy production in the UK (SDC, 2007). It incorporated recommendations for a sustainable Severn Estuary Barrage. The Government’s five sustainability principles (Cullingworth, B. and Nadin, V. (2006)) would form the basis upon which a decision could be made as to the feasibility of such a large-scale development. From a sustainable development perspective, this document contained a number of pertinent points regarding the environmental issues and legislation covering the Severn basin area. The Commission highlighted the ‘significant adverse effect on the integrity of the protected sites’ although it acknowledged that these effects would be different dependant upon the location and scale of the chosen development.

The report also notes views expressed concerning the suitability of both the Habitats and Birds Directive in coping with the implications of climate change ((SDC), 2007 pp. 130). Furthermore, the SDC heard views suggesting that ‘one-off derogation’ would be appropriate for matters of climate change mitigation ((SDC), 2007 pp. 131). The Commission certainly remained convinced that such derogation would send entirely the wrong signal to other Member States. In addition, it could create difficulties in the scale of development that would be deemed appropriate for such derogation. The Commission also acknowledged that increased housing and other development around the Barrage could reduce the effectiveness of any climate change benefits.

The SDC concluded that “any proposal for a Severn barrage must fully comply with the Directives and adhere rigorously to the process they set out”. It stated emphatically that they “would be firmly against any moves to revise or derogate from the Directives to facilitate proposals for a Severn barrage”. Despite these reservations and contrary to the findings of the 1980s feasibility study, the now defunct SDC (ENDS Report 426, July 2010, pp. 5-6)

advocated a 10 mile barrage between Cardiff Bay and Weston-super-Mare (ENDS report 393, 2007).

Following the publication of 'Turning the Tide', the Department for Energy and Climate Change (DECC) called for proposals for tidal energy schemes for the Severn Estuary in 2008. As a result of the consultation that formed part of Phase 1 of the Feasibility Study, a long-list of 10 projects was formed (DECC, 2009), these included barrages, tidal lagoon and tidal fence schemes. In January 2009, a short-list of five was developed based upon the findings of the feasibility study together with best available technology. The five short-listed schemes (fig 1) currently under consideration are:

- **Shoots Barrage** – near the Severn road crossings, estimated to cost £3.2bn to construct and generate 2.7TWh/year, which would account for just under 1 percent of UK electricity
- **Beachley Barrage** – slightly smaller and further upstream than Shoots Barrage, and upstream of the Wye, estimated to cost £2.3bn and generate 1.6TWh/year
- **Welsh Grounds Lagoon** – an impoundment on the Welsh shore of the Estuary between Newport and the Severn road crossings, estimated to cost £4bn and generate 2.3TWh/year
- **Bridgwater Bay Lagoon** – an impoundment on the English shore of the Estuary between Hinkley Point and Weston Super Mare, estimated to cost £3.8bn and generate 2.6TWh/year
- **Cardiff-Weston Barrage** (commonly known as 'The Severn Barrage') – located between Brean Down and Lavernock Point, estimated to cost £20.9bn and generate 16.8Twh/year , which would account for 5 percent of UK electricity.

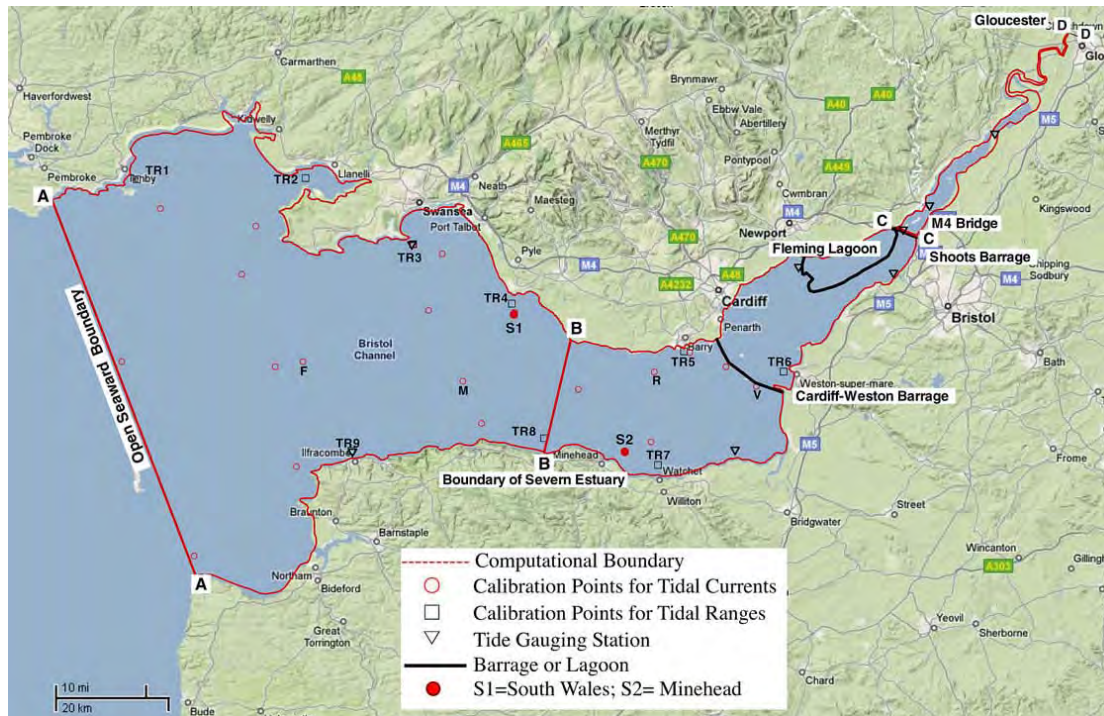


Fig 1 Map of the Bristol Channel and Severn Estuary showing the location of the Cardiff-Weston (or Severn) and Shoots barrages and the Fleming lagoon (Xia, J.Q., 2010).

An inquiry conducted in April 2010, by the All Party Parliamentary Group on Angling concluded that the largest scheme, the ‘Severn Barrage (Cardiff – Weston)’ would damage the estuary and recommended a shorter barrage upstream (Salter and Walker, 2010). Not surprisingly the Department for Energy and Climate Change denied suggestions that this scheme had been abandoned and stated that it will continue to appraise the project along with the other four (DECC, 2009).

The Severn Barrage is considered to be the best, although the cost would be the highest of all the schemes, it is thought that the barrage would generate more power than the other proposals and in turn would help contribute to the renewable energy targets.

As the largest and potentially most damaging of all the schemes, the ‘Severn Barrage’ has courted most controversy. In February 2010, the RSPB highlighted the potential for increased flooding risk from the barrage. It is widely acknowledged that the Severn Barrage will result in significant impact upon the habitat and wildlife of the basin. This, according to the ENDS report (409, Feb 2009) could extend inland and impact upon salt marsh habitat and the ecosystems of the Somerset Levels.

The Severn Tidal Barrage

The construction of the barrage would contain 216 turbines, each predicted to generate 8,640MW in total. The array of sluices would let the tide in then close to force the water back out through the turbines after the tide has gone out. The purpose for building a head on the water is to create pressure which increases the efficiency of the turbines. Shipping would be accommodated through a series of locks which would handle the larger container vessels. The barrage would take around eight years, requiring around 35,000 employees at peak building time. The lifespan of the barrage is predicted to be 120 years (About three times that of a nuclear reactor), but this could be increased to 200 years with a comprehensive maintenance program (Anderson, S. H., 2007).

In the Severn estuary high sediment load is maintained in the water column is maintained by the current passing through the deep channels, which are narrow with shallow areas, bordered by ground which floods in high tides. The mud suspended within the estuary mostly comes from erosion from the rivers and streams that feed into the estuary, such as the River Wye. (Parsons Brinckerhoff Ltd (2008)). The area also contains large amounts of inter-tidal mudflats, sand flats, salt marsh, rocky inter-tidal areas and sand dunes. These sub tidal habitats are well known for their importance to all the wildlife in the area. However little is known about the wildlife on the estuary floor, due to the dynamic movement of the tides within the estuary. This in turn may alter distributions and concentrations of contaminants in estuary water and sediments. (Pethick, et al., 2009). The changes in suspended sediments may also affect primary productivity and may contribute to eutrophication (Mettam, C. 1978). The government itself acknowledges that the vibrations and noise from the barrage would damage and confuse migratory and estuarine fish species such as the protected Allis and Twaite shad (DECC, 2009).

Climate Change & Renewable Energy Implications

Under the Kyoto Protocol, a number of countries agreed to reduce their greenhouse gas (GHG) emissions. In order to fulfil its Kyoto obligations, in June 2009, the European Commission published the Climate Change and Energy Package of legislation (also known as the "20-20-20 package" (COM (2010) 265 Final (May 2010)) to implement the new EU energy policy adopted in March 2007. The proposals contain measures to reduce GHG emissions in the EU by 20% by 2020, and ensure that 20% of the EU's overall energy needs are met by renewable energy by 2020.

As part of the 20-20-20 package, the Commission published Directive 2009/28/EC OJ L140/16 which would:

- A common EU framework for the promotion of renewable energy.
- 20% of the EU's final energy consumption to come from renewable sources by 2020.
- 10% of each member state's transport energy consumption to come from renewable sources by 2020.
- That the UK achieves a 15% share of its energy from renewable sources by 2020, as its contribution to the EU's overall 2020 target.

This directive took into account the fact that UK's renewable energy share is very low, at 1.3%, in the baseline year of 2005, and only about 2.25% in 2008 (DECC 2009. *The UK Renewable Energy Strategy*). In addition to its international targets, the UK Government has agreed to commit itself to national targets which were made legally binding under the Climate Change Act 2008. Thus making the government legally bound to ensure that the UK reduces its overall GHG emissions by 80% (below 1990 levels) by 2050 (s1(1) Climate Change Act 2008).

To sum up the EU renewable energy targets have committed the UK Government to producing 20% of its total energy from renewable resources by 2020, which corresponds to about 35% of the UK's electricity demand (DTI, 2003; Kirby and Shaw, 2005). At present only about 5% of the UK's electricity comes from renewable energy resources, so the UK Government is currently exploring the scope for increasing its generation of renewable energy from tidal sources, with tides having the advantage over wind and waves in that they are predictable (Falconer et al., 2009)

Economics and Ecosystem Services

The UK Government's 20% targets are higher than EU targets of achieving 15% of total energy consumption from renewable sources by 2020. These targets equate to 93,000GWh and 187,000GWh respectively of electricity generation from renewable sources (Frontier Economics 2008).

A variety of renewable energy sources could be utilised to contribute towards these targets, including the use of on-shore and off-shore wind, tidal, nuclear and hydro projects. The Government needs to consider the cost effectiveness of each method of generation as well as the capacity levels available from each source.

A feasibility study produced by DECC 2008 puts forward the case for the Severn Tidal Power scheme, within which 5 options were short-listed.

A financial analysis of these 5 schemes highlights the cost effectiveness of each proposal after looking at the costs of construction, costs of habitat compensation and maintenance costs. These costs have then been set against the savings produced through avoiding CO₂ emissions and the costs saved by not generating electricity using Combined Cycle Gas Turbines.

Cardiff	Weston Barrage	Shoots Barrage	Beachley Barrage	Fleming Lagoon	Bridgewater Bay Lagoon
Cost £bn	43.4	6.3	4.8	10.4	7.6
Benefit £bn	16.3	2.7	1.6	0.7	3.0
NPV (best estimate) £bn	-27.1 -3.5		-3.2	-9.7	-4.5
Value of CO ₂ emissions avoided (undiscounted)	£1.5bn £235m		£132m	£196m	£224m

(Source: DECC Severn Tidal Power Feasibility Study(2008)

On the basis of these figures, it is clear to see that none of the options provide a positive Net Present Value (NPV). It is important to note that a positive NPV means that the investment is worthwhile, while the highest NPV among different options means higher net value of the specific investment decision (Mott MacDonald Limited, April 2007).

Therefore the savings produced in terms of CO₂ emissions do not compensate for the large construction, compensatory and maintenance costs incurred. Shoots Barrage would provide the greatest saving of CO₂ emissions in relation to cost therefore appears to be the most attractive option. However the Shoots Barrage would only produce a 0.8% of Annual electricity generation (based on 2020 figures) compared to the 4.8% generation projected for the Cardiff Weston Barrage.

The Cardiff Weston Barrage is estimated to be able to produce 17,000GWh of electricity per annum representing a major contribution to the Government targets. However the costs of generating electricity from this source are deemed to be expensive. Projections have

estimated the cost of producing electricity from a large barrage to be in the region of £60-154/MWh compared against £41-62/MWh from offshore wind (Frontier Economics 2008). Indeed the cost of electricity generation from a large barrage is more expensive than other options with the exception of solar energy (260/MWh) (Frontier Economic 2008) and even nuclear energy provides a cheaper option at £38/MWh (DECC 2008).

Although the NPV calculations have factored in the costs of compensation for destruction of habitat at £65,000/hectare, there are many other economic considerations to account for. Additional environmental costs may exist relating to the loss of species and damage to habitat. The economic cost to local industry (fishing and port activity) needs to be accounted for. Lost employment, although balanced by the creation of jobs from such energy schemes, must also be factored in.

Any NPV should also reflect an additional, emerging, consideration namely the value of the ‘services’ provided by the Severn Estuary coastal and marine ecosystems. These provide ecological functions that directly or indirectly translate to a variety of tangible benefits to human society. For example they support the production of food, climate regulation, flood protection, pollution sinks and recreational and aesthetic benefits (DEFRA, 2007 *An Introductory Guide to Valuing Ecosystem Services*; Remoundou, 2009) .

Ecosystem services can be defined in a variety of ways, including “the benefits human populations derive, directly or indirectly, from ecosystem functions” (Costanza, R., et al, 1997); “the benefits people obtain from ecosystems” (MEA 2003); “the direct and indirect contributions of ecosystems to human well-being” (TEEB, 2008), and the “services provided by the natural environment that benefit people” (Defra, 2007). The common link between the various definitions of ecosystem services is the emphasis placed upon the beneficial role played by ecosystems in enhancing or maintaining aspects of human well being and thereby human society.

By appreciating the value of these ‘services’, an assessment can be made as to the economic implications of changes to ecosystem balance. Such changes can arise from dynamic environmental conditions, modification in management practices, and development activity. A valuation of these ecosystem services could provide a convenient method to monitor the impacts of ecosystem change which can also be used identify triggers for intervention in ecosystem management (Watson, R., 2010).

There is considerable debate over how ecosystem services should be defined and classified

which reflects differing interpretations of how social benefits are linked with ecosystem functionality (Beaumont and others 2008; Constanza 2008). Several approaches to the measurement and classification of ecosystem services have been proposed (Efteci, 2006; Frid, 2008). Each of these classification systems uses similar, but distinctive categories, which results in a lack of consistency and comparability between assessments. It has been suggested by Fisher and others (2009) that a single ecosystem service classification system will not be applicable in all circumstances and therefore a classification system should be tailored to meet the specific needs of a given assessment. In contrast, there have been recent calls to develop a single internationally standardized list of ecosystem services that can serve different purposes (Haynes-Young, R., and Potschin, M., 2008), but as yet, no such universal approach exists. The most well-known and widely applied classification of ecosystems services was developed by the Millennium Ecosystem Assessment (MEA) (2003) which although useful for educational purposes, is not suited for economic valuation. The MEA firmly established the concept of ecosystem services as an approach for linking ecosystem function to human welfare within the marine environment (UNEP, 2006). It categorises ecosystem services as provisioning, regulating, cultural, or supporting (MEA, 2003). The MEA ecosystem service classification has been complimented as being an intuitive and highly useful as an educational and policy tool (TEEB, 2008). However, it has been widely criticised as not 'fit for purpose' (Boyd, J., Banzhaf, S. (2007) as it exhibits logical inconsistencies within and between categories, it mixes processes (means) and benefits (ends), it is therefore particularly prone to double counting (Haines-Young & Potschin, 2007). It is therefore unable to produce economically robust valuation of ecosystem services (Fisher and others 2009). However, the classification system developed through 'The Economics of Ecosystems & Biodiversity' (TEEB 2009) is based on a distinction between ecological processes and the benefits experienced by humans. This model, it is suggested, is far more consistent with the framework developed by Fisher and others (2009), which similarly focuses on valuing the benefits to human wellbeing, and avoids the risk of double counting by separating such benefits from underlying ecosystem processes. The classification has three components:

- Core ecosystem processes: these describe the basic ecosystem processes supporting ecosystem functions.
- Beneficial ecosystem processes: these are the specific ecosystem processes that directly underpin benefits to people.
- Beneficial ecosystem services: these are the products of ecosystem processes that directly impact human wellbeing.

One such example is that of the saltmarsh habitat that, if lost from around the UK, could result in an additional £17 – 32bn per annum on flood defence (Beaumont et al. 2008).

Therefore based upon the SDC's estimate that the Severn barrage would result in an 'unquantified but substantial loss of existing 539ha saltmarsh resource' it could be argued that this would represent a £205 – £383million increase in flood defence costs.

It is also worth noting that Public Service Agreement 28 (Watson 2010) on the Natural Environment contains within its vision a commitment '....where the value of the services provided by the natural environment are supposed to be reflected in decision-making' (Watson, 2010). Public Service Agreements (PSAs) are essentially detailed aims and objectives or key priority outcomes that the Government wants to achieve in the current spending period (2008 – 2011). As such there is a danger that they will be disregarded in favour of climate change mitigation (Local Government Improvement and Development, 2009). Furthermore evaluation of 'ecosystem services' in purely monetary terms provides, at best, a wide range of values. At worst the services are impossible to value, for instance cultural or religious benefits, thereby reducing the service to 'value-less' and thus ignored (Ruhl, J.B., Kraft S. E., and. Lant, C. L., 2007).

Proponents of the Severn barrage cite La Rance as a model and highlight the slight impact this latter development made upon the environment. However, the economic implications of the likely more severe environmental impact caused by the Severn barrage is less investigated. By using a more appropriate scheme in Holland for comparison, Morris (2008) highlights the very real possibility of additional flood defence costs incurred as a result of the increased sedimentation of the Severn basin. This is supported by models showing erosion caused by increased wave/wind energy on the standing body of water held by the barrage. His study indicates significantly greater economic costs associated with the rise in flood risk caused by the very barrage that is supposed to reduce flooding. This risk is, of course, further exacerbated by the expected sea level rise as a result of climate change.

The Sustainable Development Commission, however, concluded that 'environmental concerns and the unknown costs of habitat creation should not hold back the development of a barrage across the river [Severn]. This is despite the limited carbon savings it would achieve for its £15bn cost. Indeed, as Friends of the Earth (FoE 2007) point out, £15bn spent upon other renewable energy technology, such as offshore wind farms, could generate 6TWh of electricity more than the barrage and save almost three times as much CO₂ (ENDS report No. 393, 2007).

The economic argument for the introduction of one or more of these schemes is not proven. The wide-ranging estimates used to calculate these forecasts create large margins of error.

However, even the best-case scenarios produce forecasts showing large negative NPVs. This does not make the Severn Tidal Barrage schemes appear cost effective or desirable. Even if the Government's desire to achieve renewable energy targets override the initial economic arguments, when these forecasts are combined with information showing cheaper alternative methods for renewable energy generation, there seems to be little compelling evidence for progressing the Severn schemes.

Environmental implications

As the largest and potentially most damaging of all the schemes, the 'Severn Barrage' has courted most controversy. In February 2010, the RSPB highlighted the potential for increased flooding risk from the barrage. It is also widely acknowledged that the Severn Barrage will result in significant impact upon the habitat and wildlife of the basin. This, according to the ENDS report (409, Feb 2009), could extend inland and impact upon saltmarsh habitat and the ecosystems of the Somerset Levels.

If any of the above mentioned schemes were to go ahead for imperative reasons of overriding public interest (Art.6(4) Habitats Directive 92/43/EEC OJ L206), then compensatory habitats would need to be provided (Art.6(4) Habitats Directive). This could result in managed realignment schemes on an 'unprecedented scale' (SDC, 2007). In 2007, the Severn Estuary and the area surrounding it was put forward by DEFRA as a candidate Special Area of Conservation (Habitats Directive) (DEFRA 2009). The area is a Site of Community Importance (SCI) under the Habitats Directive for intertidal and subtidal habitats and migratory fish species and a Special Protection Area (SPA) designation under the Birds Directive. These designations place a requirement upon the UK government to take compensatory measures to ensure the overall coherence of Natura 2000 is protected (Art 6(3) & 6(4) Habitats Directive 92/43/EEC).

According to the RSPB (2010) the Severn Barrage will result in a loss of intertidal feeding habitat with devastating consequences for the breeding and overwintering bird populations. They support the view held by Morris (2008) that La Rance in France is unsuitable for comparison due to its rocky substrate and agree that environmental data be taken from the Dutch storm surge barrier sited at Oosterschelde. The sediment budget has been changed with less water flow washing sediment out to sea. This has resulted in increased siltation up stream with potential consequences for myriad bird and migratory fish species (van Zanten and Adriaanse, 2008).

Applicable Legislation.

Environmental protection has not always been the goal per se of the European Union and an explicit legislative mandate with respect to environmental policy only appeared after the Single European Act in 1986 (Ginige, T (2002) Mining Waste Part I & II The Aznalcóllar Tailings Pond failure, European Environmental Law review – 76-88 & 102 – 113). It has been suggested by some academic that the Treaty of Lisbon, which was agreed in 2009 reinforces the significance of the environment in the European agenda and the continued influence of EU policy-making on environmental law. The new Article 3 (3) of the Treaty of European Union emphasises the protection of the environment as a key part of sustainable development set in the context of balanced economic growth and price stability intended on ensuring full employment and social progress. (McEldowney, S & J (2009). This is a very naive interpretation and it is suggested that an alternative reading of Art 3 (3) of the Treaty of European Union is that it is indicative of the current trend in the EU of reducing the high level environmental protection by emphasising the importance by given to economical and societal issues.

There is no specific legislation that expressly regulates the activities of renewable energy projects like those proposed in the Severn Estuary either at European Union or at UK level. There are, however, international, European and UK laws governing environmental and nature conservation. Applicable legislation, including the OSPAR, Ramsar, EC Habitats and Birds Directives, EC Water Framework Directive, EC Environmental Liability Directive, Aarhus Convention and the relevant Public Participation Directive 2003/35/EC and the Access to Information Directive 2003/4/EC, Marine and Coastal Access Act 2009, is discussed below.

OSPAR Convention

UK is a signatory to the 1992 Convention for the Protection of the Marine Environment of the North-East Atlantic (OSPAR Convention), which was signed in Paris in 1992. It has been ratified by the contracting parties – Belgium, Denmark, the Commission of the European Communities the European Community United Kingdom of Great Britain and Northern Ireland, (...). and came into force on 25 March 1998. It is suggested that based on Art.1 of the Convention, the Severn Estuary and the seas adjacent to it, falls within its remit. (DECC, 2008 Severn Tidal Power – Scoping Topic Paper, Marine and Estuarine Water Quality Dec 2008 by Parsons Brinckerhoff Ltd in association with Black & Veatch Ltd) As such UK has a duty to: take all possible steps to prevent and eliminate pollution and (.....) protect the marine

area from adverse effects of human activity (Art.2); all possible steps to prevent and eliminate pollution from land-based sources (....)(Art.3).

Marine & Coastal Access Act 2009 (MCAA)

It is a fascinating bit of draftmanship combining key considerations of protection of the marine environment, human health, prevention of interference with legitimate uses of the sea, (s69 MCAA 2009) and a strong emphasis on renewable energy(s12) and marine economic zones (s116 MCAA 2009). It will require a delicate balance being made between environmental protection and the need to develop a network of renewables to decarbonise the electricity sector in the fight against climate change, if it is to benefit the natural environment and the renewables industry.

Ramsar Convention

The Ramsar Convention came into force in 1975. The treaty imposes a general duty on contracting parties to promote the conservation of wetlands and waterfowl. “Wetland” is defined in a very wide sense and includes areas of marsh, (...) with water that is static or flowing, fresh, brackish or salt, including areas of marine water the depth of which does not exceed six meters. UK ratified the Convention in 1976. In doing so agreed to: designate and monitor wetlands in its territory including those on the list of most important wetlands -this includes the Severn Estuary and its surrounding areas (CCW & Natural England (2009) Severn Estuary SAC, SPA and Ramsar Site: Regulation 33 Advice from CCW and Natural England, June 2009); draw up a national wetland policy as the key means for delivering “wise use”; create a database and the promotion of the exchange of information between citizens and administrative authorities and between the parties to the convention.

It is argued that the convention by interpreting wetland in such a wide manner and establishing protection duties in very general terms runs the risk of being unenforceable (Bell, S & McGillivray, D (1997). In the UK all terrestrial areas included within listed Ramsar sites in England are currently Sites of Special Scientific Interest (SSSI's) and are legally protected under the Wildlife and Countryside Act 1981, as amended by the Countryside and Rights of Way (CROW) Act 2000 and the Natural Environment and Rural Communities (NERC) Act 2006 (DEFRA 2006, Ramsar sites in England - a Policy Statement). One would rightly assume that with good nature conservation and protection laws in place and sufficient funding (395 million pounds have been spent since 2000), the natural environment would be protected to a very favourable level. It seems that the reality could not be further from the truth. According to the recent National Audit Office Report (National Audit Office, 2008) about a

third of the sites did not have conservation objectives and the quality of the record keeping was variable. The inherent defects of the Ramsar designation coupled with Natural England's track record to date does not bode well for the protect fauna and flora in the Severn Estuary.

Natura 2000

The centrepiece of the European Union's ambitious Biodiversity Strategy is the Natura 2000 network of European protected sites. The legislation that means to deliver this ecological network is the Habitat Directive 92/43/EEC OJL206/7 on the conservation of natural habitats and of wild fauna and flora). Once a SAC is placed on the EC Commission list, the Member State must:

- establish priorities for the maintenance or restoration –Art.4(4).
- management plans which correspond to the ecological requirements of the site –Art.6(1).
- avoid the deterioration of the natural habitats and disturbance of the species –Art.6(2).
- conduct an appropriate assessment of the implication for the site of any project not directly connected with or necessary to the sites management, but which is likely to have significant effect on it –Art.6(3).
- undertake surveillance of the habitats –Art.9&11.

The environmental assessment in Art.6(3) is a cause for concern as it is site specific and its focus is on conservation, rather than the significant effect of the project. (Ginige TA (2002). Furthermore under the EC Habitats Directive of 1992 it is suggested that there will be a negative assessment of the project under Art. 6(3), owing to the fact that the scheme will probably cause the disappearance of between about one-third (at neap tide) and two-thirds (at spring tide) of the existing intertidal areas, even before consideration of species specific impacts (FoE,2007).

Another cause for concern is the possibility that per Art.6(4) of the Habitats Directive that a member state may permit a project with damaging implications for a site to be carried out for imperative reasons of overriding public interest so long as compensatory measures are taken to ensure that the overall coherence of Natura 2000 is protected(e.g. habitat restoration) (Scott, J (1989) pg.112).

The European Court of Justice has confirmed in Castro Verde case (Commission v. Portugal (Case C-239/04, 26.10.06)), Art. 6(4), as a derogation from Art.6(3) and the general protection conferred by the Directive, must be interpreted strictly. Article 6(4) reads:” If, in spite of a negative assessment of the implications for the site and in the absence of alternative solutions, a plan or project must nevertheless be carried out for imperative reasons of

overriding public interest, including those of a social or economic nature, the Member State shall take all compensatory measures necessary to ensure that the overall coherence of Natura 2000 is protected. It shall inform the Commission of the compensatory measures adopted.

Article 6(4) of the Habitats Directive makes it clear that the obligation to consider alternative solutions is upon the Member State. Further more Regulation 62(1) of the Conservation of Habitats and Species Regulations 2010, SI 2010/490, reads:

“If they are satisfied that, there being no alternative solutions, the plan or project must be carried out for imperative reasons of overriding public interest (which subject to paragraph 2 (concerned with priority habitats and species) may be of a social or economic nature), the competent authority may agree...to the plan or project notwithstanding a negative assessment of the implications of the site”.

As highlighted by the *Castro Verde* case, “Solutions”, and the “reasons”/“public interest” should form the main part of the IROPI test and in the opinion of Advocate General Kokott (*Commission v. Portugal* (Case C-239/04, 26.10.06 at 53))

“The absence of alternatives cannot be ascertained when only a few alternatives have been examined, but only after all the alternatives have been ruled out. The requirements applicable to the exclusion of alternatives increase the more suitable those alternatives are for achieving the aims of the project without giving rise – beyond reasonable doubt – to manifest and disproportionate adverse effects. The choice requires a balance to be struck between the adverse effect on the integrity of the SPA and the relevant reasons of overriding public interest. The necessity of striking a balance results in particular from the concept of “override”, but also from the word “imperative”. Reasons of public interest can imperatively override the protection of a site only when greater importance attaches to them. This too has its equivalent in the test of proportionality, since under that principle the disadvantages caused must not be disproportionate to the aims pursued.”

Alternative solutions specific to the Severn Tidal Barrage are all manner of technology capable of producing reliable, low carbon electricity over the long term for the UK. This would include examination of tidal resources throughout the UK, including both tidal range and tidal stream resources (Royal Academy of Engineering, 2008). It will also require the examination of the suitability of other renewable low carbon energy resources. There is an obligation upon Member States under the Directive to examine alternative sites for the project (Krämer L., 2009). The competent authorities have to examine the possibility of resorting to

alternative solutions which better respect the integrity of the site in question ... They could involve alternative locations, different scales or designs of development, or alternative processes. The “zero-option” should be considered too”. (European Commission in the form of Managing Natura 2000 Sites, published in 2000 at Section 5.3.1). Further more they must consider if alternative solutions exist or not, and that this consideration “will also include other alternative solutions that may be suggested by other stakeholders and should not limit consideration of alternative solutions to those suggested by a project's proponents and that alternative solutions could be located even in different regions or countries” (EC Commission Assessments of plans and projects significantly affecting Natura 2000 sites November 2001).

Assuming the UK Government can satisfy the “no alternative solutions” and IROPI tests, the final hurdle under Art.6(4) is that it would have to “*take all compensatory measures necessary to ensure the overall coherence of Natura 2000 is protected*”. Thus when a competent authority is considering whether habitat or species losses can be made up by compensation, it must look at two different aspects per Art. 3(1) of the EC Habitats Directive, i.e. “*on the one hand the targeted species and habitats in terms of quantity and quality, and on the other hand the role of the site in ensuring the adequate geographical distribution in relation to the range*” and be mindful that these criteria are met. (European Commission. 2007 .Guidance document on Article 6(4) of the 'Habitats Directive' 92/43/EEC January 2007)

Furthermore all possible mitigation must be addressed before thinking about compensation. One must not make error of dressing up compensation as mitigation in an attempt to show that the integrity of the site is not affected, as attempted by Associated British Ports who sought unsuccessfully with regard to the Dibden Bay Container Terminal. Here the appropriate assessment was fundamentally flawed in this respect. (Dibden Bay Container Terminal Inspector’s Report, at para.36.166)

Finally a Member State cannot propose a project under Art.6 (4) which cannot provide a package of compensation which falls short of being equivalent to that which is lost (European Commission 2007 at 1.51). It is questionable as to how this might be achieved for the Severn Estuary when considering amongst other things, the fact that duration of DECC’s 2008 review was too short in comparison with what is considered an appropriated timescale for such evaluation (Sally A. Keith, et al 2009).

Art.16 of the Habitat Directive states that if no satisfactory alternative exists and the derogation is not detrimental to the maintenance of the populations of the species concerned at a favourable conservation status in their natural range, Member States may derogate from the provisions of Articles 12, 13, 14 and 15 (a) and (b) the article that specify the methodology required for the species protection. Thus, instead of having a Directive with a high level of protection for the environment, what we have ended with is one that represents a dramatic reassertion of member states sovereignty over their natural resources (Scott, J., 1989). Furthermore there were numerous flaws in the UK transposed, EC Habitat Directive 1992 legislation i.e. Conservation (Natural Habitats, etc) Regulations 1994, (SI 1994/2716). These flaws included the failure to apply the law fully offshore, the presence of excessively wide defences permitting harmful acts if they were the incidental results of lawful operations, inadequate measures for surveillance and monitoring and the absence of procedures to assess the impact of development plans on designated sites as stated by the European Court of Justice (ECJ) in the recent case of *Commission v UK* 2005 ((C-6/04) [2005] ECR I-9017, [2005] All ER (D) 231 (Oct)). This in turn has led to the revised and consolidated Conservation of Habitats and Species Regulations 2010, SI 2010/490, which took effect as of 1 April 2010, which should now provide a single source for the law in England and Wales, conserving the plants, animals and habitats awarded protection at European level notwithstanding Art 6(4) of the Habitats Directive 92/43/EEC.

EC Environmental Liability Directive

The Environmental Liability Directive 2004/35/CE (ELD) is an EU wide legislation with the objective of “establishing a common framework for the prevention and remedying of environmental damage at a reasonable cost to society...” The Environmental Liability Directive of April 2004 came into force in the UK on 1 March 2009 via the Environmental Damage (Prevention and Remediation) Regulations 2009. It will apply to “environmental damage” as defined in Art.2 of the ELD which applies to protected species and natural habitats under the Birds and Habitats Directives. Therefore, if the government reduces intertidal range after a properly carried out exercise under Arts 6(3) and (4) of the Habitats Directive 92/43/EEC, that reduction would not amount to “environmental damage” within the ELD. However, it should be noted that this protection extends only to “*previously identified adverse effects*.” If for argument sake all prior investigations do not reveal any adverse effect, which then emerges during the building of the barrage, then the government may in principle be liable if the damage was caused either by an Annex III activity (it is suggested that water impoundments might well be included) or, if caused by a non – Annex III activity, where the operator was at fault or negligent- Art.3(1)ELD.

Water Framework Directive

The legislation regarding water quality standards are well developed and can be found at national, European International levels, all covering different aspects(Bell,S 2009). The Water Framework Directive 2000/60/EC (WFD) is comprehensive and provides a framework that classification of water bodies: the setting of objectives and the monitoring of achievement as well as identifying what means are required to maintain energy improvement in water quality requiring that all inland and coastal waters achieve “good status” by 2015 (McEldowney, S., & J., 2009) and it includes “transitional water bodies” or estuaries. The physical and hydrological characteristics of the Severn Estuary support a wide variety of the species; giving consideration of depth variations, the structure and condition of intertidal zones, flow, currents and wave exposure all of which including the water quality status will be affected by a tidal barrage. The Environment Agency(EnA) has made it clear that the Cardiff-Weston barrage would make it impossible for the UK to be in compliance with the WFD by 2015. (Environment Agency, 2010)

The possibility exists that a Severn tidal power scheme may change the components that define the status of the transitional and coastal water bodies in the river basin, and may alter some freshwaters too. Perhaps the question we need to ask, is whether the river system, its sediments and components may be altered drastically further up river by a barrage system? To answer that question we may need to study the effects that the Petitcodiac causeway in south-eastern New Brunswick, Canada had on the Peticodiac river instead of La Rance in France(Sentinells Petitcodiac Riverkeeper. 2010).

Aarhus Convention

The dialogue between the government and the public i.e. what in the UK is known as public participation which “has been pointed out as being limited by consultant institutions and their commissioning bodies pre-framing issues, publics and consultations themselves. This has led to consultations labelled moribund, tokenistic and even deceptive”(Gerlach, J. 2008). This lack of dialogue between the government and individual citizens with regard to the Severn Tidal Barrage is another of the many indication of the glaring problems that exist with regard to the application of the Aarhus Convention in the UK (UKELA 2008). This is not surprising given that, historically, excessive secrecy has been the hallmark of UK’s approach to sharing of environmental knowledge with its citizens (McEldowney, S., & McEldowney, J., 2009). The rights afforded by the Aarhus Convention i.e. access to information; public participation in decision-making; and access to justice were further strengthened by the

following EC directives -EIA Directive 85/337/EEC as amended by 97/11/EC and Public Participation Directive 2003/35/EC; IPPC Directive 96/61/EC and as amended by the Access to Information Directive 2003/4/EC.

These rights should be regarded as minimum requirements that operate as a minimum standard rather than an optimum level to be attained; a floor, not a ceiling. Articles 3(5) and (6) state that the Aarhus Convention provisions:

shall not affect the right of a Party to maintain or introduce measures providing for broader access to information, more extensive public participation and wider access to justice than required by [the Convention] ... and shall not require any derogation from existing rights.

The government considers that since the introduction of new Environmental Information Regulations 2004 (SI 2004/3391) (the Environmental Information Regs. 2004), that the UK is Aarhus compliant. It is suggested that we apply the letter and spirit of the law.

The proposed Severn Tidal Barrage ought to provide the government with the prime opportunity to engage the public meaningfully and to instigate this debate according to a reconfigured geography of more deliberative, inclusive consultation. Without precedent and with the socio-ecological stakes so high, the government cannot afford not to consult the public on the issue”(Gerlach, J. 2008).

Conclusion

The centrepiece of the European Union’s ambitious Biodiversity Strategy is the Natura 2000 network of protected sites, under the EC Habitats Directive 92/43/EEC. The environmental assessment in Art. 6(3) is a cause for concern as it is site-specific and its focus is on conservation rather than the significant effect of the project. Another cause for concern is the possibility per Art. 6(4) that a Member State may permit a project with damaging implications to be carried out for reasons of overriding public interest so long as compensatory measures are taken to ensure that the overall coherence of Natura 2000 is protected (e.g. habitat restoration).

The fundamental issue, therefore, remains unresolved. Are we now at the stage where ‘imperative reasons of overriding public interest’ covers all climate change mitigation development?

Although this paper has focussed upon a significant proposed development in the form of the Severn Tidal Barrage, it applies equally to any development designed to combat the, as yet unknown, effects of climate change. Is this not akin to marching into an area of pristine Brazilian rainforest, chopping down trees and putting wind turbines in their place? The protection afforded to the Severn River basin is on a similarly national, European and international level. Is the legacy of this development to be the irrevocable loss of this ecosystem?

Essentially, if this project can go ahead with the acknowledged significant environmental impacts, then it could represent the future benchmark for other developments. The door will be opened for development anywhere in England and Wales that mitigates climate change, to be allowed to proceed no matter what statutory protection is in place.

In addition, the balancing of conflicting interests inherent within this scheme need to be supported by a scientific study of the environmental impact, an economic study of the NPV and a social study into the pros and cons of this barrage. All three elements need to be sufficiently robust to truly enable the balancing of conflicting interest to take place. Restraint from the current 'eyes wide shut' headlong rush into climate change mitigation is needed to fully appreciate the consequences of the development. As Aronson (et al. 2006) suggest

‘Ecologists and economists working together with NGOs, community leaders and governments, may help forge a way forward with a new paradigm: ecology as if people mattered; economics as if nature mattered.’

The implementation of the new EC laws incorporating the Aarhus Convention Principles will require the development of new relational capacities both between social agents, in the form of learning how to collaborate and understand others roles and capacities differently, and between social-ecological systems (Pahl-Wostl, et al. 2008). New institutional arrangements will also be required to facilitate the more sustainable relationships, based on new framings of the issues at stake and agents involved. This will necessitate the development of new identities, as well as institutions and individual capacities that are more socially and environmentally robust with the common goal of sustainable nature conservation. Without this debate, there can be no balancing of conflicting interests and, thus, no sustainable development.

Alder and Wilkinson (1999) take the view that:

‘Decision makers sometimes have to sacrifice one good for another greater good. The broader our area of environmental concern the more difficult becomes the balancing act’ (Alder, J. and Wilkinson, D. 1999).

However, the current situation would suggest that climate change mitigation represents the ‘trump card’. Thus the paradox remains: the damaging impact to the environment caused by development designed to mitigate the damaging impact to the environment. Only by addressing this paradox can development for climate change mitigation be undertaken with ‘eyes wide open’ and consequences accepted and balanced.

References

- Anderson, S. H., 2007. *Barrages And Impoundments: A Preliminary Study of the Potential for Sustainable Net Energy Capture Obtainable by Storage for Tidally Amplified Release* (‘Ecostar’ Scoping Model.) Available from:<
<http://www.marinet.org.uk/refts/7ecostar1.pdf>>[Accessed 20 July 2010].
- Alder, J., and Wilkinson, D., 1999. *Environmental Law & Ethics*. London: Macmillan Press.
- Aronson, J., Milton S.J., Blignaut J.N., Clewell A. F., 2006. Nature Conservation As If People Mattered *Journal For Nature Conservation* 14 (2006) 260—263
- Beaumont, N. J., Austen, M. C., Mangi, S. C. and Townsend, M., 2008. Economic valuation for the conservation of marine biodiversity. *Marine Pollution Bulletin* 56, 386 – 396.
- Bell, S and McGillivray, D., (1997). *Environmental Law*, London, Blackstone Press.
- Bell, S., and McGillivray, D., 2009. *Environmental Law*. Seventh Edition. Oxford :Oxford University Press.
- Boyd, J., Banzhaf, S., 2007. What are ecosystem services? The need for standardized environmental accounting units. *Ecological Economics. Volume 63, Issue 2-3, August 2007, Pages 616-626*
- Clarke, N. (2010, May). *Independent Newspaper*. Retrieved July 20, 2010, from
<http://www.independent.co.uk/news/business/news/britains-offshore-renewable-energy-worth-a-billion-barrels-of-oil-and-145000-new-jobs-1977531.html>

Commission of the European Communities, Analysis of options to move beyond 20% greenhouse gas emission reductions and assessing the risk of carbon leakage , COM (2010) 265 Final (May 2010) (hereinafter Analysis of options to move beyond 20%)

Costanza, R., 2008. Ecosystem Services: Multiple classification systems are needed. *Biological Conservation* 141:350-352

Cullingworth, B. and Nadin, V. (2006) *Town and Country Planning in the UK*, Routledge, 14th Edition, pages 200-205.

DEFRA 2006, Ramsar sites in England - a Policy Statement. London:HMSO

DEFRA, 2007 *An Introductory Guide to Valuing Ecosystem Services*. London: HMSO

DECC, 2008 SEVERN TIDAL POWER – SCOPING TOPIC PAPER, Marine and Estuarine Water Quality Dec 2008 by Parsons Brinckerhoff Ltd in association with Black & Veatch Ltd)

Department of Energy and Climate Change, 2009. *Severn Tidal Power Phase 1 Consultation*. London: DECC.

DECC, 2009. *The UK Renewable Energy Strategy*, Cm 7686.London: HMSO

DTI, 2003 Our Energy Future—Creating a Low Carbon Economy. TSO, London, p. 138.

DTI, 2003 Government Gives Green Light to Renewables Future. Press Release, DTI, Dec 1st, 2003, available from <http://213.38.88.221/gnn/national.nsf/TI/E42654EB406D834F80256DEF00569E89?opendocument>.

Dibden Bay Container Terminal Inspector's Report, at para.36.166 as stated in Owen, R., European Nature Conservation Sites and the Appropriate Assessment of Plans and Projects in *Journal of Planning & Environment Law* 2007, December, Sweet & Maxwell

Environmental Data Services, 2004. *Severn Barrage kicked further into touch*. ENDS report, 348, pp. 33, January 2004.

Environmental Data Services, 2007. *Government advisers call for Severn barrage*. ENDS Report, 393, pp. 12 – 13

Environmental Data Services, 2009. *Severn Barrage study unleashes tide of paper*. ENDS report 409, pp. 40 - 41, February 2009.

Environmental Data Services, 2010. *Environmental bodies fall to government's quango cuts*. ENDS Report 426, pp. 5-6, July 2010.

European Commission in the form of Managing Natura 2000 Sites, published in 2000 at Section 5.3.1.

European Commission. 2007 .Guidance document on Article 6(4) of the 'Habitats Directive' 92/43/EEC January 2007.

Falconer, R.A., Lin, B.L., Ahmadian R., and Xia, J.Q., (2009). The Severn Barrage: hydro-environmental impact assessment studies, *Proceedings of 33rd IAHR Congress*, Vancouver, Canada, pp. 2075–2082.

Eftec, 2006. Valuing our Natural Environment. Report for DEFRA.

Environment Agency, 2010 Tidal barrages and barriers: Position Statement <
<http://www.environment-agency.gov.uk/research/library/position/41199.aspx>> Last Accessed on 30 July 2010.

Friends of the Earth Cymru (FoE) . 2007 *The Severn Barrage* Friends of the Earth Cymru: Cardiff.

Fisher, B., Costanza, R., Turner, R. K., and Morling, P. 2009. *Defining and Classifying Ecosystem Services for Decision Making*. CSERGE Working Paper EDM 07-04.

Frid CLJ, Paramor OAL, Brockington S, Bremner J (2008) Incorporating ecological functioning into the designation and management of marine protected areas. *Hydrobiologia* 606: 69-79

Frontier Economics. (2008). *Analysis of a Severn Barrage*. London: Frontier Economics Ltd.

Gerlach, J. 2008. Turning the Tide of Opinion – The Severn Barrage and Spaces of Public Consultation ,Oxford University Centre for the Environment, Oxford

Ginige, T.A., 2002 Mining Waste Part I & II -The Aznalcóllar Tailings Pond failure, *European Environmental Law Review* . Amsterdam: Elsevier pp76-88 & 102 – 113.)

Haynes -Young, R., and Potschin, M., 2008 *England Terrestrial Ecosystem Services And The Rationale For An Ecosystem-Based Approach*. Overview Report London:DEFRA

Haines-Young, R. and Potschin, M. (2007): *The Ecosystem Concept and the Identification of Ecosystem Goods and Services in the English Policy Context*. Review Paper to Defra, 21pp.

HMSO. (1989) *The Severn Barrage Project: General Report: Energy Paper 57*. London: HMSO

Krämer, L., 2009 .The European Commission's Opinions under Article 6(4) of the Habitats Directive J Environmental Law .2009 21: 59-85

Kirby, R., and Shaw, T.L., 2005. Severn Barrage, UK-environmental reappraisal, *Proceedings of the Institution of Civil Engineers. Engineering Sustainability* 158 (ES1) (2005), pp. 31–39

Kyoto Protocol." *United Nations Framework Convention on Climate Change*. Web. 30 July 2010. <<http://www.unfccc.int>>.

Joint Nature Conservation Council, 2010. *UK SAC/SCI sites summary. Special Areas of Conservation/Sites of Community Importance in the UK as at 14 December 2009*. Peterborough: JNCC. Available from <http://www.jncc.gov.uk/page-1456> [Accessed 23 July 2010].

McEldowney, S., & McEldowney, J., 2009 *Environmental Law*, Pearson Education, Harlow.

Mettam, C. 1978. Environmental effects of tidal power generating schemes. *Aquatic Ecology* , 307 - 321. Springer Netherlands

Millennium Ecosystem Assessment, 2003. *Ecosystems and Human Well-being: A Framework*

for Assessment. Island Press, Washington, DC. World Resources Institute.

Millennium Ecosystem Assessment, 2005. *Ecosystems and Human Well-being: Biodiversity Synthesis*. Island Press, Washington, DC. World Resources Institute.

Mott MacDonald Limited. (April 2007). *Appraisal of Costs & Benefits of Smart Meter Roll Out Options*. London: BERR.

Morris, R. K. A., 2008. *A Severn Barrage - planning for environmental consequences*. Presentation for the Renewable Energy Association Wave and Tidal Technologies Symposium 23 September 2008, Cardiff

Morris, R.K.A. 2009. The impact of tidal energy barrages on estuarine geomorphology. In: Aranoff, J.B. *Handbook of Nature Conservation: Global, Environmental and Economic Issues*. Nova Science Publishers.

National Audit Office, 2008. *Natural England's Role in Improving Sites of Special Scientific Interest* (19 November, 2008) HC 1051 Session 2007-08.

Pahl-Wostl, C., Tabara, D., Bouwen, R., Craps, M., Dewulf, A., Mostert, E., Ridder, D., Taillieu, T., 2008. The Importance of Social Learning and Culture for Sustainable Water Management. *Ecological Economics* Volume 64, Issue 3, 15 January 2008, Pages 484-495

Parsons Brinckerhoff Ltd and ABP Marine Environmental Research Ltd for DECC ,2008. *SEVERN TIDAL POWER - SCOPING TOPIC PAPER. Marine and Estuarine Water Quality*. London: Department of Energy and Climate Change (DECC)

Pethick, J. S., Morris, R. K. A. and Evans, D. H., 2009. Nature conservation implications of a Severn tidal barrage – A preliminary assessment of geomorphological change. *Journal for Nature Conservation* 17, 183 – 198.
(Price, 2007)

Remoundou, K., Koundouri, P., Kontogianni, Paulo, A., Nunes, A.L.D., and Skourtos, M., (2009) Valuation Of Natural Marine Ecosystems: An Economic Perspective. *Environmental Science & Policy*, Vol 12, pg 1040 – 1051. Amsterdam: Elsevier

Royal Academy of Engineering (2008). The Severn Barrage Conference 22 May 2008.
London: The Royal Academy of Engineering.

Royal Commission on Environmental Pollution, 2000. 22nd Report: Energy - the Changing Climate. Norwich: HMSO.

Royal Society for the Protection of Birds, 2010. *Severn Tidal Power RSPB briefing on the implications of the Eastern Scheldt (Oosterschelde) for the Severn Estuary*. Available from: <http://www.rspb.org.uk/ourwork/casework/details.asp?id=tcm:9-228221> [Accessed 18 July 2010].

Ruhl, J.B., Kraft S. E., and. Lant, C. L., 2007, *The Law and Policy of Ecosystem Services*, Island Press, Washington, DC, USA.

Sally A. Keith, Adrian C. Newton, Michael D. Morecroft, Clive E. Bealey and James M. Bullock . **2009** .Taxonomic homogenization of woodland plant communities over 70 years
Proceeding of the Royal Society B (2009) 276, 3539–3544

Salter, M. and Walker, C., 2010. *Severn Barrage Digest: An Environmental Gamble Too Far*. Available from: <http://www.martinsalter.com/index.php/reports/severn-barage-report/> [Accessed 10 June 2010].

Sentinells Petitcodiac Riverkeeper.2010. Petitcodiac River Restoration Moncton, New Brunswick, Canada< http://petitcodiac.org/index.php?page=petitcodiac-river-restoration&hl=en_US> accessed on 20 July 2010

Scott, J., 1989 *EC Environmental Law*. Addison Wesley, Longman: Essex.

Sustainable Development Commission, 2007. *Turning the Tide: Tidal Power in the UK*. London: Sustainable Development Commission.

Taylor, S. J., 2002. *The Severn Barrage – Definition Study for a new appraisal of the Project*. Final Report by Sir Robert McAlpine Limited on behalf of the Severn Tidal Power Group. ETSU Reort No. T/09/00212/00/REP.

TEEB – The Economics of Ecosystems and Biodiversity (2008) The Economics of Ecosystems and Biodiversity: An interim report. European Commission, Brussels. URL:

www.teebweb.org

(last access: July 2010).

TEEB – The Economics of Ecosystems and Biodiversity for National and International Policy Makers –Summary: Responding to the Value of Nature 2009. URL: www.teebweb.org (last access: July 2010).

The Offshore Valuation Group (OVG), 2010. *A Valuation of the UK's Offshore Renewable Energy Resource*. Powys: Public Interest Research Centre.

UKELA 2008 . Response of 25 March 2008 to UK Government on Implementation of the Aarhus Convention. Working party on Environmental Litigation. Dorking.

United Nations, 1997. Kyoto Protocol to the United Nations Framework Convention on Climate Change, Conference of the Parties on Its Third Session, FCCC/CP/1997/L.7/Add.1, 10 December.

United Nations Environment Programme.2006.Marine and Coastal Ecosystems and Human Well-being: A Synthesis Report Based on the Findings of the Millennium Ecosystem Assessment.Nairobi: UNEP.

Watson, R., 2010. *The Ecosystem Approach*. Presentation to the Institute of Ecology and Environmental Management Conference 24 March, 2010.

Xia, J.Q., Falconer, R. A., Lin, B. (2010). Impact Of Different Tidal Renewable Energy Projects On The Hydrodynamic Processes In The Severn Estuary, UK. *Applied Energy*, Volume 87, Issue 7, July 2010, Pages 2374-2391

Zanten, E. van, Adriaanse, L. A. 2008. *Sand demand Oosterschelde: An analysis of the erosion of the tidal flats of the Oosterschelde after construction of the storm surge barrier and its effect on flood safety, nature, shipping and fisheries*. Available from: http://www.rspb.org.uk/Images/RSPBbriefEasterScheldtreportfinal_tcm9-240984.pdf. [Accessed 18 July 2010].

Assessing the roles of regulation, litigation, management and environmental advocacy in protecting groundwater resources

Robert W. Schultz¹, Donald C. Daniels¹, and Thomas F. Vandenburg²

¹AMEC Geomatrix, Inc., 2101 Webster Street, 12th Floor,
Oakland, California, 94612, USA

²Dongell Lawrence Finney LLP, 707 Wilshire Boulevard, 45th Floor, Los Angeles,
California, 90017, USA.

Email: TVandenburg@dlflawyers.com; Robert.Schultz@amec.com;
Don.Daniels@amec.com

Abstract:

As a society, we employ a variety of methods to protect groundwater resources. We regulate industrial activities to ensure proper methods are used to avoid releases of hazardous materials to the environment. If a release of contaminants occurs or has already occurred, we closely regulate remediation of the contaminants, and we may institute litigation either to recover costs spent or to ensure adequate cleanup.

We manage our groundwater resources for their protection and long term sustainability by monitoring groundwater conditions, mitigating contamination of water supply aquifers, instituting conservation programs and utilizing storage solutions. Finally, environmental advocacy, including: raising public awareness about groundwater contamination; educating business leaders and elected officials on sustainability issues; and securing public funds to protect groundwater resources is an important method of groundwater protection. Protecting groundwater resources is challenging. Significant financial resources, both public and private, are required, as well as a firm commitment to this goal by all segments of society. This paper will analyze the interplay between these different elements used to protect groundwater resources, and will discuss the efficacy and proper roles of environmental advocacy, litigation, regulation, management and education in this effort.

Keywords:

CERCLA, Contamination, groundwater protection, litigation, water supply.

1 Introduction

Safe, reliable drinking water supplies are critical to public health and economic growth (UNDP 2006). In the United States, expansion of surface water supplies is limited in many areas by the environmental impacts of surface water development projects, limited existing water storage capacity, and the high cost, both financial and environmental, of expanding surface water storage. Consequently demand for groundwater is increasing. Groundwater resources, however, can be particularly vulnerable to degradation.

For groundwater to be a reliable source of drinking water, we as a society must take actions to equitably manage the risks and costs of contamination. Our overarching needs include protection of drinking water aquifers from future contamination; implementation of treatment and management strategies that allow continued operation in areas with impacted groundwater; and identification of new groundwater resources, including restoration of impacted aquifers or areas.

In this paper, we present five methods used to protect groundwater resources, with examples taken predominantly from California, where the authors are practitioners in the groundwater protection and cleanup field. Groundwater in California and the United States (US) is typically managed by municipalities or local government agencies and frequently extracted and delivered to consumers by sister agencies within the local governments or through a public-private partnership between the local agency and a for-profit water company. US federal law makes no distinction between historical and recent contamination in cost allocation or cleanup requirements.

To protect the groundwater resources, local agencies in California play a major role in advocating the public benefits of water projects. Advocacy has enabled many agencies to secure public funding for groundwater resource management and for cleanup of contamination. Litigation is another common method, and is frequently used to draw attention to problems and to obtain funding from those who can pay. Regulation is used to deter future contamination and to oversee the contaminant cleanup process. Water agencies and purveyors may also educate the public they serve to increase awareness of the importance and vulnerability of groundwater resources. And finally, groundwater management is used as a means of sustainably using the resource. After presenting each method, we consider the common shortcomings and successes of each protection element.

2 Advocacy: State Funded Cleanups and State Assurance Funds

Successful local water agency, or other stakeholder, advocacy for public funding can obtain funding that is directly put toward solving water quality problems. Recent public funding of water resource projects in California has typically been through bond propositions approved by voters in statewide general elections. State assurance funds, also called “cleanup funds,” are another means for providing direct funding of groundwater protection projects. Financing of assurance funds can include private sector funding, for example matching funds may come from private party stakeholders as opposed to the involved public agencies that frequently contribute to direct publicly funded projects.

2.1 California Examples of Groundwater Protection Bond Propositions

California maintains a vast network of intra- and inter-state canals and dams to transport and store surface water supplies from the Sacramento-San Joaquin Delta, the Colorado River and other water bodies, yet the State also relies heavily on groundwater. Between approximately twenty and forty percent of the State’s water supply comes from wells (LAO, 2010). Recent bond measures have included funding for groundwater projects and, more specifically, groundwater protection. Propositions 50 and 84 are key examples.

2.1.1 California Proposition 50

California voters approved the Water Security, Clean Drinking Water, Coastal and Beach Protection Act (Proposition 50) in 2002. The stated intent of this proposition was to provide a safe, clean, affordable, and sufficient water supply to meet the needs of California residents, farms, and businesses. Through the development of annual priority lists, Proposition 50 has funded local and regional projects in each of the three past funding years. In 2009, Proposition 50 provided \$75 million for groundwater protection projects.

2.1.2 California Proposition 84

In passing The Safe Drinking Water, Water Quality and Supply, Flood Control, River and Coastal Protection Act of 2006, the State's voters approved \$1 billion for a combination of water delivery infrastructure and groundwater protection. Of this total amount, \$60 million was set aside for groundwater contamination prevention and reduction projects. Example projects on the first round project priority list included several wellhead treatment projects and a groundwater remediation project in the Los Angeles Basin.

2.1.3 The Safe, Clean, and Reliable Drinking Water Supply Act of 2010 (SBX7)

This landmark ballot proposition was passed by the California legislature in early 2010; and was scheduled to go to the voters in November 2010; however, on June 30, 2010, the ballot proposition was pushed back to 2012. Voter approval of this bond is uncertain at this time; however, two-thirds of the legislature, representing both democrats and republicans, voted for the bond to get it on the ballot. This massive water bill would put over \$1.1 billion toward addressing multiple water supply issues. It includes approximately \$1 billion for groundwater quality protection and cleanup.

2.2 Examples of State Assurance Funds

Assurance funds are another example of direct funding of groundwater protection projects. Most assurance funds are financed by assessing fees on the industry that resulted in the contamination. In these programs, parties responsible for the contamination (Potentially Responsible Parties, or PRPs) manage or conduct the work and are reimbursed for certain costs.

2.2.1 Petroleum Underground Storage Tank (UST) Cleanup Funds

The most prevalent type of assurance fund for groundwater cleanup in the United States are commonly called "UST Funds," and provide PRPs cost recovery for petroleum underground storage tank (UST) cleanups. In 1996, there were 34 state UST funds approved by the United States Environmental Protection Agency (USEPA). The number had increased to 37 in 2005, and at the time of writing, the USEPA website identified 43 such funds (USEPA, 2010).

Total funding available through assurance funds can be significant. The California UST Fund provided \$190 million for cleanups in FY2008-2009, and between \$200 and \$300 million in each of the seven preceding years (Sjoberg Evaschenk, 2010). Nationwide, state assurance funds have spent over \$5 billion on UST cleanups over the past 10 years (USEPA, 2010).

2.2.2 Dry Cleaner Cleanup Funds

Tetrachloroethene (PCE) has been used in dry cleaning operations worldwide, and the majority of PCE used in the United States for the past 30 years has been by the dry cleaning industry (CVRWQCB 1992). PCE is the second most frequently detected volatile organic compound (VOC) in groundwater (USGS, 2006). To respond, thirteen states have mandated programs and provided funding for dry cleaner site remediation. The Florida dry cleaner fund was open for application from 1996 until 1998, and over 1,400 sites were entered into the program. To date 129 of those sites have been cleaned up and received rehabilitation completion orders (SCRD, 2010).

2.3 State Funded Cleanups: Common Successes

State-funded cleanups demonstrate a number of advantages over litigation and/or regulatory enforcement. Unlike other groundwater protection strategies, public funds can be put quickly and directly toward cleanups. In many instances, taxpayer investment in groundwater cleanup can be recovered after-the-fact. This approach reduces the time period to cleanup or containment of groundwater contamination, and by determining the cost of a groundwater cleanup up front, can reduce the costs and risks of cost-recovery litigation. Moreover, state cleanup funds provide for a type of public-private partnership on cleanups, with the associated opportunities for innovation and cost-com petition. PRPs are incentivized by state cleanup funds to cooperate with regulatory oversight agencies and water agencies to protect groundwater.

2.4 State Funded Cleanups: Common Shortcomings

Despite the efficiency and other advantages to groundwater protection, voters in the current political climate may not be receptive to new taxes. Further, not all “polluting industries” can be realistically funded through fee-based systems; drycleaners are a prime example. PCE groundwater contamination from drycleaners is very expensive to cleanup, but the revenue associated with a typical independent drycleaning operation is very modest. Conversely, petroleum funds have been successful in part because the fees set aside to fund cleanups generally amount to less than one percent of the retail cost.

2.5 Opportunities to Maximize Results

Within the current publicly funded cleanups in the United States, there are opportunities for stakeholders to increase the resources put toward groundwater protection and to maximize the overall protection afforded by a groundwater project. Regulatory agencies can be encouraged to oversee cleanup of priority sites. Fee-based state assurance funds could be expanded to cover other high-revenue “polluting industries.” Nitrate is arguably the most pervasive and threatening groundwater contaminant in California today, and hundreds of thousands of metric tons of nitrogen are sold in fertilizers each year. In California, the combination of public funding for cleanup with other water infrastructure projects has proved to be a politically feasible means of funding water resource protection projects.

3 Groundwater Protection Litigation

The United States Congress passed the Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA), in part, to allow parties which pay for

remediation of hazardous materials to recover those clean up costs from “responsible parties”. Under CERCLA, a party which incurs cleanup costs can recover those costs from four categories of parties: Owners of “facilities” from which hazardous materials are released; operators of facilities from which hazardous materials are released; transporters of hazardous materials to facilities from which they are released; and “arrangers” of the disposal of hazardous materials from a facility¹.

Entities which have rights to use groundwater have often sued under CERCLA to recover “response costs” associated with cleanup of contaminated groundwater, or have sued for a Court Order requiring a PRP to cleanup groundwater contamination.

The phenomenon of methyl-t-butyl ether (MTBE) groundwater contamination led to an evolution of legal theories concerning such claims. Instead of suing individual gas station operators who allegedly allowed gasoline containing MTBE to leak from underground storage tanks, suits were brought against the oil companies whose product – gasoline – contained the contaminant. This litigation culminated in 2008 with one of the largest environmental settlements in United States history worth more than \$422 million (AP, 2008).

The legal theory for the MTBE suits was “strict product liability” which is commonly used for defectively manufactured or designed products which either cause personal injury and/or property damage. In part due to the success of the MTBE litigation, these theories are now being used in cases involving agricultural products (soil fumigants) and PCE used in the dry cleaning industry.

3.1 City of Lodi Action

The City of Lodi Action is a cautionary tale regarding a local city which passed its own statute which was patterned after CERCLA but went even further in allowing the City to recover its costs for remediation.

PCE groundwater contamination was discovered in the City of Lodi in the early 1990s. In 1997 the City of Lodi executed an agreement with the State of California regarding investigation and abatement of that groundwater contamination. Three months later, City of Lodi enacted “MERLO” a municipal ordinance which granted the City wide powers concerning the recovery of its cleanup costs for PCE groundwater contamination.

Two insurance companies filed suit to prevent City of Lodi from enforcing the terms of MERLO against companies they insured. Lengthy and expensive litigation followed. The Federal Court issued a series of rulings which decimated City of Lodi’s position, finding that MERLO was preempted by Federal Law and therefore unenforceable and the City of Lodi itself was a “PRP” due to PCE releases from its sewer system. The City not only found itself facing significant legal bills but also owed a Wall Street Firm \$16 million for litigation costs².

¹ See generally 42 U.S.C. 9601 et seq.

² See generally *Fireman's Fund Insurance Company, et al. v. City of Lodi*, 296 F.Supp.2d 1197 (E.D.Ca. 2003)

The case remains the “poster child” of cautionary tales concerning groundwater contamination litigation

3.2 MTBE Litigation

In the United States, MTBE was added to gasoline, in varying amounts, from approximately 1979 through 2007. When gasoline containing MTBE was stored in underground storage tanks at gas stations, leaks occurred which led to groundwater contamination. MTBE is considered a drinking water contaminant and the US Environmental Protection Agency (EPA) states that MTBE is a potential human carcinogen.

Significant MTBE groundwater contamination litigation was brought throughout the United States during the years of 1995 to 2008. The cases were brought against the oil companies which sold the gasoline, not the “mom and pop” gasoline station operators. These cases were consolidated in the Multi-District Litigation Federal Court in New York which was established to handle mass tort litigation.

In 2008 a partial settlement was reached in which a number of defendants agreed to pay at least \$422 million dollars to 153 public water providers to settle their liability for MTBE groundwater contamination. Subsequently, a party notably absent from this settlement, ExxonMobil, was hit with a jury verdict of \$105 million in an MTBE groundwater contamination case brought by the City of New York. Soon thereafter, ExxonMobil settled 50 MTBE lawsuits for \$17.5 million (Rita Cicero, 2009).

3.3 TCP Litigation

Starting in approximately 2004 there have been 13 cases filed in California concerning groundwater contamination of a chemical called trichloropropane (TCP). TCP allegedly was an impurity in soil fumigants manufactured by Dow Chemical and Shell Chemical Company. The cases filed have alleged that TCP contamination from these products has impacted groundwater resources affecting 13 different water purveyors in California.

The first of these cases went to trial in January 2010, solely against Shell Chemical Company. In May 2010 the trial resulted in a defense verdict³. The Court is now addressing motions concerning the scheduling of trials for the remaining matters.

3.4 Litigation against the Dry Cleaning Industry

Another approach has been taken by California Water Service (“Cal Water”), a water purveyor with 14 different subdivisions throughout California. Cal Water asserts that it owns over 200 production wells which have been contaminated by PCE, which Cal Water asserts has resulted from dry cleaning operations near their wells. Cal Water has sued 58 defendants for this contamination⁴. Cal Water has sued the manufacturers and distributors of solvents used in dry cleaning operations and manufacturers and

³ City of Redlands, et al. v. Shell Oil Co. et al., JCCP 4435, San Bernardino Superior Court Case No. SCVSS120627

⁴ California Water Service Co. et al. v. Dow Chemical Corp. et al., San Mateo Superior Court Case No. CIV 473093

distributors of dry cleaning equipment, alleging that these parties are responsible for this contamination and should be held liable for resulting damages and/or cleanup costs.

3.5 Efficacy of Groundwater Litigation

It is difficult to quantify the efficacy of groundwater litigation. It is undoubtedly an expensive enterprise. Critics of the tort system in the United States have estimated that 50-57% of the total monies awarded in tort litigation are absorbed by transaction costs, legal fees and court costs (Cody, 1997). Might groundwater litigation result in a similar percentage of recovery?

Every success story can be offset by a story of setbacks. The City of Lodi is a cautionary tale. The nearby City of Modesto also pursued novel legal theories regarding PCE groundwater contamination and obtained a much better result than its neighbor. Press reports assert the City's recovery to be over \$41 million (Ashton, 2009). The City of Modesto likely considers its groundwater litigation to be a success, although that litigation lasted over 10 years. MTBE litigation is certainly considered a success by those parties that filed suit. The jury is still out (literally) on TCP and PCE / dry cleaning litigation.

The preventative effect of groundwater litigation is open to debate as well. Many of the cases concern actions and/or decisions made decades ago. Jury verdicts in the 21st century can have little effect upon decision makers of the 20th century.

Groundwater contamination cases appear to be growing in scope. Entire industries are now being sued as water purveyors appear to be seeking "economies of scale" for such lawsuits rather than litigation on a site by site basis. These suits are pushing the envelope of established legal theories and are causing a fusion of two areas of law: environmental law and products liability law. How that fusion will mutate and/or evolve remains to be seen.

4 Regulation: Resource Protection and Cleanup Oversight

PRP-funded cleanup under an administrative order is probably the most common framework for groundwater cleanup in the United States. Regulation of potentially polluting industries helps to limit or prevent new releases. Regulated site data is publicly available in many cases. The availability of site data can serve to increase public scrutiny thereby encouraging compliance, and it can be helpful to water agencies and other groundwater users in assessing threats and improving planning. In many cases, regulatory oversight also serves to confirm the completeness of a cleanup.

Unfortunately, regulatory oversight of cleanups rarely decreases, and frequently increases, the time required to perform site investigation and implement remediation. In most cases, an incentive is necessary for a PRP to quickly address a groundwater quality problem. Assistance from those involved in the use of a groundwater basin is frequently necessary to put under the context of a groundwater problem. At too many sites, effort is expended sampling or cleaning up groundwater contamination that does not threaten regional resources. Conversely, samples are frequently not collected in

locations where a regional problem could be detected because of the focus on site-scale cleanup.

The results of our review suggest that the groundwater protection benefits of regulation can be enhanced when stakeholders assume an active role. Identification of areas of increased vulnerability and evaluation of the relative magnitudes of the groundwater contamination sites in a basin by a water agency, for example, can help a cleanup oversight agency to prioritize its workload. By publicizing or otherwise encouraging public awareness of groundwater problems, water agencies and other stakeholders can create a public relations incentive (or disincentive) that encourages groundwater protection.

5 Management of Groundwater Resources

The long term viability of groundwater as a resource requires sustainable management of both the quantity and quality of the resource. The United States Geological Survey defines sustainable groundwater management as “the development and use of groundwater in a manner that can be maintained for an indefinite time without causing unacceptable environmental, economic, or social consequences” (Alley et al., 1999). Management involves monitoring and tracking of the resource, wellhead protection, contaminant mitigation, implementation of conservation measures, and utilization of groundwater storage.

5.1 Groundwater monitoring

A fundamental element of groundwater management is monitoring of groundwater resources including both quantity as indicated by groundwater elevation and quality as determined by analysis of groundwater samples. In California, monitoring is performed under various programs. In 2000 California created the Groundwater Monitoring and Assessment Program (GAMA), which was expanded in 2001 by the Groundwater Quality Monitoring Act. Monitoring involves analysis of a broad suite of organic and inorganic chemicals in groundwater basins across the state. The resultant data are maintained in a GAMA database, which is available online.

More recently, as a component of a five bill water and groundwater legislation package enacted in 2009, Chapter 1 of SBX7 is dedicated to groundwater and includes statewide monitoring and reporting requirements to be implemented by local agencies. Originally slated to be on the November 2010 ballot, the SBX7 package has recently been pushed back to 2012 due to the current fiscal climate in California.

5.2 Wellhead protection

The cost of mitigating groundwater contamination is significant frequently resulting in abandonment of an otherwise useable aquifer. A fundamental component of groundwater management is the protection of aquifers through wellhead protection to avoid degradation of the resource. Wellhead protection involves: the identification of sensitive recharge areas within a groundwater basin where surface releases could leach to underlying groundwater resources or flow to surface waters that interact with groundwater; the identification of threats such as industrial facilities and agricultural

areas; and the mitigation of those threats through administrative and engineering controls such as stormwater management and restrictions on pesticide use.

5.3 Contaminant mitigation

Due to industrial and agricultural activities, many aquifers have been impacted by various chemicals at concentrations exceeding drinking water standards. As a result, existing wells have often been shutdown and unexploited aquifers have been left undeveloped. To address this issue and bring these resources into production, water purveyors have incorporated a variety of approaches to meeting water quality standards.

5.3.1 Wellhead treatment

One approach is to effectively treat impacted water on demand by engineering an inline treatment system at the wellhead. Water exiting the well is appropriately treated depending on the nature of the groundwater contamination. In the case of volatile organic compounds, the water may pass through an air stripper or through activated carbon to remove the contaminants. For other compounds, treatment may involve flow through an ion exchange column to replace harmful chemicals with innocuous ones. In many instances, wellhead treatment is a cost effective method to mitigate groundwater impacts when compared to in situ remedial alternatives focused on remediating impacts within the aquifer. In California, any wellhead treatment prior to direct potable use requires a permit from the California Department of Public Health (CDPH).

5.3.2 Groundwater blending

Similarly, in impacted areas where wells are screened in multiple aquifers or multiple wells are present within a water supply system, water purveyors will blend water from multiple zones within a well, from multiple wells, or from a combination of groundwater and other water sources such that the resultant water is acceptable for distribution.

Water purveyors extracting contaminated groundwater may also combine wellhead treatment with blending as is the case in the North Hollywood Operable Unit of the San Fernando Superfund Area, where extracted groundwater is treated by air stripping and granular activated carbon prior to blending with water imported from outside of the groundwater basin (MWD, 2007).

5.4 Water Conservation

A fundamental component of groundwater management is the overarching issue of water conservation regardless of the source. Any reduction in demand will decrease the need to exploit additional groundwater resources. Although per capita water consumption in California has decreased over time primarily due to improved efficiency in agricultural water use, the overall demand for water continues to increase due to increases in the state's overall population (UCAIC, 2009). Given that large scale public works projects have fallen out of favor due to cost and environmental impacts, additional conservation measures are necessary to meet this growing demand.

During drought years, local agencies often require a reduction of urban water use; however, generally, per capita water use in the urban environment has increased over

time. The most significant reductions in water demand can be made in agriculture due in part to the fact that agriculture accounts for such a large percentage of overall water use. Such reductions could include further adoption of micro-irrigation techniques and a reduction or phase out of water intensive agriculture in California. In addition to other elements, water conservation is a significant component of SBX7 the recently enacted package of water bills currently scheduled for the 2012 ballot.

5.5 Groundwater storage

There is a growing appreciation for the value of water and the need to develop a sustainable model for use of this resource. In addition to concerns regarding quality of groundwater, there are concerns regarding the growing demand for water, especially in arid areas such as the western United States. To this end, groundwater basins are being viewed more and more as an alternative means of storing water. This is especially true in the western United States where essentially all rainfall occurs in the relatively cool wet season, while water demand peaks during the hot dry season.

5.5.1 Spreading grounds

Recharge of aquifers has been a key component of water management in southern California since the 1940s. Los Angeles County Department of Public Works operates the Hansen Spreading Grounds in Sun Valley, which consist of a 105 acre recharge area below Hansen Dam that is designed to allow surface water to infiltrate at up to 49 cubic feet per second to the underlying aquifer (MWD, 2007). The spreading grounds have a recharge capacity of approximately 35,000 acre feet of water. The five such spreading grounds within the Upper Los Angeles River Area, a total of recharge capacity of approximately 105,000 acre feet per year. In some instances, recharged water comes from other sources outside of the basin.

5.5.2 Recycling/reuse

In addition to recharge of surface waters via spreading grounds, water is also reclaimed through infiltration or direct injection of treated wastewater. The degree of treatment required is generally dependent on the source of the water, the proposed reuse of the water, and the method of recharge.

Due to concerns regarding potential degradation of groundwater resources, states regulate the recharge of groundwater aquifers with reclaimed water. These state regulations and guidelines vary considerably as a function of intent; while some states such as Arizona, California, Colorado, Florida, North Carolina, and Washington focus on the reuse of water, other states are primarily focused on recharge as a means of disposal of wastewater following acceptable treatment. (Stone Environmental, 2005).

In California, recycling and reuse of water by recharge of groundwater aquifers is jointly regulated by the CDPH and the State Water Resources Control Board through a 1996 memorandum of understanding (CDPH and SWRCB, 1996). The CDPH developed draft regulations in 2008 regarding recharge of recycled water, which establish criteria for recharge. The regulations require that minimum treatment standards be applied, no degradation of groundwater occur, recharge be conducted in a manner that allows sufficient residence time in the aquifer prior to reextraction, and

ongoing monitoring be conducted between the recharge location and the point of extraction (Walker, 2009).

5.5.3 Saltwater intrusion

In addition to increasing water storage, recharge can be used to protect groundwater basins from intrusion of non-potable water into an otherwise potable resource through focused hydraulic control at the point of intrusion. In many areas along the western side of Los Angeles and Orange Counties, low permeability fault systems protect over pumped aquifers from salt water intrusion; however, where the faults have been compromised due to erosion, salt water can enter the groundwater basin threatening the viability of the aquifer. To mitigate this threat, water is injected in the vicinity of these compromised areas to mound water and reverse the groundwater gradient effectively pushing the salt back seaward.

6 Groundwater Protection through Education

There are two fundamental purposes to educating stakeholders on issues related to groundwater protection. Firstly, education provides stakeholders including the business community, agricultural interests, political leaders and the public with critical information that will help them understand the importance of protecting groundwater resources. Especially in California, there is general interest in issues related to sustainability and climate change; however, the association of groundwater protection with these broader issues is limited. Stakeholders are more likely to be willing to invest in methods to protect groundwater if they are aware of the importance of this resource in a broader environmental context. For instance, as a sustainability measure, water conservation has the benefit of reducing groundwater extraction and increasing residence times for contaminants moving through impacted aquifers, which reduces the spread of contaminants and may allow for increased biodegradation of contaminant plumes. Regarding climate change, many stakeholders do not associate water conservation with climate change; however, due to the significant energy consumption required to obtain, treat, and distribute water, water conservation can have a significant impact on the reduction of greenhouse gas emissions.

Secondly, from a pragmatic perspective education is a critical tool to ensure awareness of the responsibilities that various stakeholders have regarding protection of groundwater resources. For instance, outreach to the agricultural community is critical to implementing water conservation. Specific activities as well as lifestyle choices have an impact on groundwater resources. The more aware the public is on this issue the more likely these impacts can be limited.

7 References

- Alley, W. M., Reilly, T. E. & Franke, O. L., 1999. Sustainability of Groundwater Resources. US Geol. Survey Circular 1186, Denver, Colorado, U.S.A..
<http://water.usgs.gov/pubs/circ/circ1186>
- Ashton, Adam, 2009. Modesto Awarded \$18 Million in Groundwater Lawsuit. *Modesto Bee*, May 20.

- Associated Press, 2008. Most Oil Companies in MTBE Lawsuits Settle, May 9.
- California Department of Public Health and State Water Resources Control Board, 1996. Memorandum of Understanding between the Department of Health Services and the State Water Resources Control Board on Use of Reclaimed Water, February.
- California Legislative Analyst's Office (LAO), 2010. Liquid Assets: Improving Management of the State's Groundwater Resources, March 24.
- Central Valley Regional Water Quality Control Board (CVRWQCB), 1992. Dry Cleaners – A Major Source of PCE in Groundwater, March 27.
- Cicero, Rita, 2009. ExxonMobil to Settle 50 Water Pollution Suits for \$17.5 Million, *Andrews Publications*, November 13.
- Cody, Troy L., 1997. Disadvantaging the Disadvantaged: The Discriminatory Effects of Punitive Damages Caps, 25 *Hofstra L. Rev.*, 1005, 1019-1021.
- Metropolitan Water District of Southern California, 2007. Groundwater Assessment Study, Chapter IV – Groundwater Basin Reports Upper Los Angeles River Area Basins
- Sjoberg Evashenk Consulting, 2010. Underground Storage Tank Cleanup Fund Performance Audit, February 22.
- State Coalition for Remediation of Drycleaners, 2010. www.drycleaningcoalition.org, viewed March 30, 2010.
- Stone Environmental, 2005. Spray Irrigation of Athletic Fields Using Reclaimed Water: A Review for Fauquier County, Fauquier County Parks and Recreation Department, Stone Project # 041551-W, April 7.
- United Nations Development Programme (UNDP), 2006. Human Development Report 2006, Beyond scarcity: Power, poverty and the global water crisis.
- United States Geological Survey, 2006. Volatile Organic Compounds in the Nation's Ground Water and Drinking Water Supply Wells, Circular 1292.
- United States Environmental Protection Agency (USEPA), 2010. Status of State Fund Programs. <http://www.epa.gov/OUST/states/fndstatus.htm>, viewed March 30, 2010.
- University of California Agricultural Issues Center, 2009. Water Supply and Demand, AIC White Papers on California Agricultural Issues, November.
- Walker, Leah Godsey, PE, Chief, Drinking Water Technical Programs Branch, California Department of Public Health, 2009. Impacts of Water Reuse/Recycling on Groundwater, Biennial Groundwater Conference, October 6.

Carbon capture and storage and the legal challenges of the land use policy: comparing the legislation in the EU and Australia

Inga-Liisa Paavola¹

¹Faculty of Law, Energy and Environmental Law
University of Helsinki
Finland

Email: inga-liisa.paavola@helsinki.fi

Abstract:

Carbon dioxide capture and storage (CCS) is a relatively new technology in the context of climate change mitigation strategies. Its legal and regulatory implications are not yet wholly developed. This paper takes a brief look at the EU (European Union) and Australian law principles especially relevant to CCS and land use and land planning regulations and compares them and the solutions in the field of legislation. It identifies key environmental and safety risks associated with the technology, and highlights significant legal frameworks that pose challenges to the implementation of CCS within the EU under EU and international law. It then notes continuing regulatory gaps that will need to be addressed for large-scale CCS to take place. The paper concludes that the clear inclusion or exclusion of CCS activities from the range of relevant legal frameworks will increase transparency, provide regulatory certainty and ultimately facilitate CCS in appropriate contexts.

Keywords:

Australia, CCS, European Union, Land Use Regulations

1 Introduction

Efforts undertaken to capture, transport and store carbon dioxide, to avoid transboundary harm, may themselves lead to damage to the environment in the areas beyond the territorial limits of national or even the EU jurisdiction. The degree and nature of this risk will largely depend upon the systems chosen to transport and storage CO₂; the means used to transport CO₂, the routes used for this transport, and the location, reach and integrity of the geological storage sites selected for long-term storage.¹

There are a few important considerations concerning CCS and the land use policy. Questions arising are for example, what happens when the predicted CO₂ plume crosses state, how to determine boundaries and differ regulatory jurisdictions, how to create an equal determination of the suitable area of storage and how to deal with differences in different kinds of state-level or international legislation or regulatory interpretation.

2 Current regulation of the CCS and the land use policy

CCS and the land use policies relates essentially to international law principles. It is a general principle of international environmental law that States have the sovereign right to exploit their own resources, but also the responsibility to ensure that activities within their jurisdiction or control do not cause damage to the environment of other States or damage to areas beyond the limits of their national jurisdiction.² The same kind of principle is included in the Directive 2009/31/EC on the geological storage of carbon dioxide (CCS-directive) of the EU, which states that the Member States should retain the right to determine the areas within their territory from which storage sites may be selected.

No single, clear international regulatory framework applies to carbon capture and storage activities undertaken for the purpose land use policies. Only a few frameworks even directly address CCS activities, and include them or exclude them from their scope, like The London Convention³, the London Protocol⁴, the OSPAR⁵ Convention CCS-directive⁶. On the other hand, the EU has its own legal framework concerning CCS-activities, the CCS-directive. However, the CCS-directive is not broadly implemented in the Member States yet.

¹ Mace, M. J., Hendriks, C., and Coenraads, R 2007: Regulatory challenges to the implementation of carbon capture and geological storage within the European Union under EU and international law. *International journal of greenhouse gas control* 1 at <http://www.sciencedirect.com>.

² Sands, P. 2003: *Principles of international environmental law*. 2nd ed., pp. 143; 869–939.

³ 1972 Convention on the Prevention of Marine Pollution by Dumping of Wastes and Other Matter.

⁴ The 1996 London Protocol to the London Convention, entered into force in March 2006.

⁵ The Convention for the Protection of the Marine Environment of the North-East Atlantic (OSPAR) 10.12.1982.

⁶ Directive 2009/31/EC on the geological storage of carbon dioxide.

The 1996 London Protocol to the London Convention entered into force in March 2006. The Protocol supersedes the London Convention for Parties to the Protocol who are also Convention Parties. The London Protocol is notable because, when it was agreed, it specifically prohibited the dumping of wastes, including industrial wastes. Since that time, the Protocol has been amended to expressly allow CCS in certain circumstances.

The 1996 London Protocol differs in significant ways from the earlier London Convention. Under the Convention, dumping is permitted, except for materials that are on a banned list. Under the newer London Protocol, the situation is quite opposite: dumping is prohibited except for materials on an approved list. The Protocol expressly prohibits the deliberate disposal, so dumping, into the sea of all wastes from vessels or manmade platforms. This relates essentially to an option that is thought about, to the storage of CCS into the sea.

However, Annex 1 to the London Protocol mentions seven categories of wastes or other matter that may be considered dumping: dredged material; sewage sludge; fish waste, or material resulting from industrial fish processing operations; vessels and platforms or other man-made structures at sea; inert, inorganic geological material; organic material of natural origin; and bulky items primarily comprising iron, steel, concrete and similarly unharmed materials for which the concern is physical impact. CO₂ did not plainly fall within one of these listed waste categories and the Annex 1 may not be applicable to CCS then. Moreover, Annex I to the London Protocol permits the deliberate disposal under the seabed of 'pollution' from land-based sources, other than for the purpose of offshore activities. Annex II, in contrast, prohibits the 'dumping' of CO₂ (as interpreted 'wastes or other matter') in the seabed and its subsoil from vessels and man-made structures at sea. What is more, Annex III then prohibits the dumping of wastes or other matter from offshore installations, but allows 'discharges or emissions' from these installations, and allows for certain 'placements' but only in the context of offshore activities.

OSPAR's Annexes thus create the potential for different treatment of CCS and land use policy in the field of storage. OSPAR's Annexes regulate CO₂ injection into geological storage sites reached by pipeline from land, by pipeline from vessels, by pipeline from manmade structures at sea that are not related to oil and gas extraction, and by pipeline from offshore installations that are related to oil and gas extraction. They also create the potential for different treatment of CO₂ that arises from offshore activities and that does not arise from offshore activities.

The CCS-directive of the EU regulates CCS quite comprehensively. The Directive has created a legal framework to CCS and regulates also the land use policy in the scope of CCS storage and transportation. Still, the regulations stay at relatively common level and a lot depends on that how the Member states implement the Directive.

On May 16, 2008, the Government of Australia released draft CCS legislation. The draft legislation takes the form of a comprehensive set of amendments to the Commonwealth's Offshore Petroleum Act 2006 and is designed to provide an enabling framework for objective-based regulation for CCS in offshore waters. The proposed legislation accomplishes two main objectives: firstly, it provides a disposition or tenure

scheme for parties to acquire the right to storage in the offshore and secondly it provides a regulatory framework for reviewing and approving CCS operations on a case-by-case basis, with individual site plans and closure plans. In delivering on both of these objectives, the legislation also provides a framework for deciding upon the competing claims of petroleum operations and CCS operations. In the point of view of the land use policy, the draft legislation identifies site selection and approval and site closure as areas for regulation and proposes a framework.⁷

3 Branches of the use of CCS concerning the land use policy

3.1 *The Legal Challenges of the CCS storage and the Land Use Policy*

The primary question in the scope of CCS regulations and the land use legislation must be how to determine if a site is suitable for storage. Site characterization and selection is perhaps the most important step in ensuring the safety and integrity of a storage project. During this phase, much of the site-specific data are collected and permits applications are developed. Even though a site may be economically attractive, data collected during site characterization should be used to assess the technical feasibility of a site and keeping environmental effects in mind. The regulation of storage should set standards for well design. Consistent standards are needed to ensure the highest level of prevention of leakage into the ground, water, and air systems over the long term. Any regulation of well design should include operational practices, materials used, number and age of wells, potential geophysical changes, pathways in the event of leakage, and duration of storage.⁸

3.1.1 *Legislation concerning the storage in the EU*

According to the CCS- Directive, it applies to the geological storage of CO₂ within the territory of the Member States, in their exclusive economic zones and on their continental shelves. The Directive should not apply to projects with a total intended storage below 100 kilotonnes, undertaken for research, development or testing of new products and processes. The storage of CO₂ in storage complexes extending beyond the territorial scope of this Directive and the storage of CO₂ in the water column should not be permitted, as well as it is not permitted in the international treaties either⁹.

⁷Ministerial Council on Mineral and Petroleum Resources 2005: Carbon Dioxide Capture and Geological Storage: Australian Regulatory Guiding Principles at http://www.industry.gov.au/assets/documents/itinternet/Regulatory_Guiding_Principles_for_CCS20051124145652.pdf.

⁸DOE/NETL-2006/1236 2006: International Carbon Capture and Storage Projects Overcoming Legal Barriers at <http://www.netl.doe.gov/energy-analyses/pubs/ccsregulatorypaperfinalreport.pdf>.

⁹ 1972 Convention on the Prevention of Marine Pollution by Dumping of Wastes and Other Matter, The 1996 London Protocol to the London Convention, entered into force in March 2006, The Convention for the Protection of the Marine Environment of the North-East Atlantic (OSPAR) 10.12.1982.

According to the Article 4 of the CCS-Directive, the Member States have right to decide about their land use policy concerning the storage of CCS. The Member States shall retain the right to determine the areas from which storage sites may be selected pursuant to the requirements of this Directive. This also includes the right of the Member States not to allow for any storage in parts or in the whole of their territory. The suitability of a geological formation for use as a storage site shall be determined through a characterisation and assessment of the potential storage complex and surrounding area pursuant to the criteria specified in the CCS-Directive's Annex I. A geological formation shall only be selected as a storage site, if under the proposed conditions of use there is no significant risk of leakage, and if no significant environmental or health risks exist. Article 5 of the CCS-Directive states, that where Member States determine that exploration is required to generate the information necessary for selection of storage sites pursuant to Article 4, they shall ensure that no such exploration takes place without an exploration permit.

Due to CCS-Directive's article 6, the Member States shall ensure that no storage site is operated without a storage permit, that there shall be only one operator for each storage site, and that no conflicting uses are permitted on the site. In this way, the land use of the storage site itself is secured. On the other hand, it may be asked, how the 'storage site' is determined and do all determine it in the same way.

According to the CCS-directive, also the Council Directive 85/337/EEC of 27 June 1985 on the assessment of the effects of certain public and private projects on the environment¹⁰ should be applied to the capture and transport of CO₂ streams for the purposes of geological storage. It should also apply to storage sites pursuant to this Directive.

3.1.2 *Legislation concerning the storage in Australia*

CO₂ is not defined as a pollutant under Western Australian law,¹¹ which is essential in the scope of the storage of CCS. The Barrow Island Act of 2003 provides the legislative mechanism for the disposal of CO₂ by underground injection, including the approvals process for the Gorgon project¹². The Act does not define how CO₂ should be classified, but the injected CO₂ is effectively considered a byproduct of gas processing operations under the Act. The draft legislation offers three-tiered models on the current petroleum regime: firstly, an assessment permit, secondly a holding lease, and lastly an injection license.

Each storage formations' classifications are associated with increased knowledge of the geological formation that is proposed for injection and storage: potential, eligible, or identified. While a tenure holder may inject anything potential and eligible formations

¹⁰ Directive 85/337/EEC of 27 June 1985.

¹¹ The Barrow Island Act of 2003.

¹² <http://www.gorgongasproject.com>.

for appraisal purposes, approval for injection for permanent storage requires that there be a declaration of an identified storage formation.¹³

In Australia there are specific regulations for site selection and approval and site closure. The draft legislation identifies site selection and approval of storage sites and site closure as two major issues that need new regulations for CCS. The approach is outcome oriented in that the overall goal is to achieve “safe and secure storage” in a formation. Site closure builds on the site plan, requiring a program and funding for long term monitoring and verification by the Commonwealth.¹⁴

3.2 *Legal challenges in the transportation of carbon via pipelines*

Another important CCS and land use-related issue that must be well regulated is the transportation of the captured CO₂. Especially the transportation via pipelines, to which this paper concentrates on, has notable impacts on the surroundings and environment. Moreover, transportation issues may get more complicated if the pipelines for the CO₂ transportation are built through many States that may have different kinds of legal frameworks for CCS and land use policies.

Transporting CO₂ from point-of-capture to storage sites is an important linking step in the CCS project cycle. Pipeline transport is considered to be the most cost-effective and reliable method of transporting CO₂ for onshore CCS¹⁵.

Over the North Sea, several existing trunk pipelines appear technically suitable for carbon dioxide transport. These pipes are limited to transporting CO₂ at lower pressures than if new pipelines were built and could only be used assuming that the CO₂ transported is rigorously purified. Any legislation which encourages rapid decommissioning of infrastructure when oil and gas production ceases would need to be reconsidered to encourage reuse for CCS. As platforms reach the end of their service life the potential for reuse decreases. For this reason, this report makes an assumption that new platforms and wells will be required for CCS.¹⁶

To help address potential land use concerns about pipelines, it is important for regulators and the industry to characterize potential risks at the pipeline area. The

¹³ Ministerial Council on Mineral and Petroleum Resources 2005: Carbon Dioxide Capture and Geological Storage: Australian Regulatory Guiding Principles at http://www.industry.gov.au/assets/documents/itrinternet/Regulatory_Guiding_Principles_for_CCS20051124145652.pdf.

¹⁴ Ministerial Council on Mineral and Petroleum Resources 2005: Carbon Dioxide Capture and Geological Storage: Australian Regulatory Guiding Principles at http://www.industry.gov.au/assets/documents/itrinternet/Regulatory_Guiding_Principles_for_CCS20051124145652.pdf.

¹⁵ Svensson, R., M. Odenberger, F. Johnsson, and L. Strömberg, 2004: “Transportation Systems for CO₂—Application to Carbon Capture and Storage.” *Energy Conversion and Management* 45:2343-2353.

¹⁶ Department for Business Enterprise & Regulatory Reform 2007: North Sea CCS Infrastructure Report to the North Sea Basin Task Force.

potential consequences of CO₂ incidents can be modeled on a site-specific basis using several standard methods, taking into account such local conditions as topography, meteorology, and population density.¹⁷

Deploying CCS at a scale required to mitigate global warming will require transporting substantial quantities of CO₂ from capture to storage sites. The nature and extent of the network of CO₂ pipelines that would be necessary to transport such amounts of CO₂ will depend on many factors, including the proximity of storage sites to the capture facilities, the costs to acquire pipeline rights of way and associated permits, the cost to construct the pipelines, and the attendant costs to operate the pipelines and comply with operations and maintenance regulations. The CO₂ pipeline networks developed for the CCS market will most likely evolve over time. Early projects are likely to rely on a mix of options, including use or expansion of the existing CO₂ pipeline infrastructure and the development of dedicated pipelines that are sized and located for individual projects to accommodate the CO₂ specifications of those projects. Under other scenarios, a fully integrated network that utilizes CO₂ from several sources may be practical. In light of the overall costs associated with CO₂ pipelines, including the uncertainty about future material costs and cost recovery, some analysts anticipate that the CO₂ network for CCS will begin with short pipelines from CO₂ sources located close to storage sites, with a larger regional network of interconnected lines developing as the number of projects grows.¹⁸

Another aspect of CCS and transportation regulation is pipeline access and tariff. Ensure that pipelines fulfill common carrier obligations by charging reasonable rates; providing rates and services to all upon request; not unfairly discriminating among shippers establishing reasonable classifications, rules, and practices.

3.2.1 *Legislation concerning the transportation of carbon via pipelines in the EU*

Access to CO₂ transport networks and storage sites, irrespective of the geographical location of potential users within the Union, could become a condition for entry into or competitive operation within the internal electricity and heat market, depending on the relative prices of carbon and CCS. It is therefore appropriate to make arrangements for potential users to obtain such access. According to the CCS-directive, this should be done in a manner to be determined by each Member State, applying the objectives of fair, open and non-discriminatory access and taking into account the transport and storage capacity which is available or can reasonably be made available. Also the proportion of its CO₂ reduction obligations pursuant to international legal instruments and to Community legislation intended must be met through CCS.

¹⁷ IPCC Special Report: Carbon Dioxide Capture and Storage- Summary for Policymakers 2005: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change at http://www.ipcc.ch/pdf/special-reports/srccs/srccs_summaryforpolicymakers.pdf.

¹⁸ World Resources Institute 2008: Guidelines for Carbon Dioxide Capture, Transport, and Storage.

Article 21 of the CCS-Directive regulates access to transport network and storage sites. Member States shall take the necessary measures to ensure that potential users are able to obtain access to transport networks and to storage sites for the purposes of geological storage of the produced and captured CO₂.

The access shall be provided in a transparent and non-discriminatory manner determined by the Member State. The Member State shall apply the objectives of fair and open access, taking into account the storage capacity which is or can reasonably be made available within the areas and the transport capacity which is or can reasonably be made available, the proportion of its CO₂ reduction obligations pursuant to international legal instruments and to Community legislation that it intends to meet through capture and geological storage of CO₂, the need to refuse access where there is an incompatibility of technical specifications which cannot be reasonably overcome, the need to respect the duly substantiated reasonable needs of the owner or operator of the storage site or of the transport network and the interests of all other users of the storage or the network or relevant processing or handling facilities who may be affected.

Article 21 goes on, that the transport network operators and operators of storage sites may refuse access on the grounds of lack of capacity. Duly substantiated reasons shall be given for any refusal. The Member States shall take the measures necessary to ensure that the operator refusing access on the grounds of lack of capacity or a lack of connection makes any necessary enhancements as far as it is economic to do so or when a potential customer is willing to pay for them, provided this would not negatively impact on the environmental security of transport and geological storage of CO₂.

There are also regulations concerning transboundary cooperation within the transportation of in the CCS-Directive, which may become an essential question as the pipelines may be built through different States and their legal frameworks. Article 24 of the Directive states, that in cases of transboundary transport of CO₂, transboundary storage sites or transboundary storage complexes, the competent authorities of the Member States concerned shall jointly meet the requirements of this Directive and of other relevant Community legislation.

3.2.2 *Regulation concerning the transportation of carbon via pipelines in Australia*

The Regulatory Guiding Principles for Carbon Dioxide and Geological Storage in Australia that were released in 2005¹⁹, also deal with the land use policy from the point of view of transportation of captured CO₂. Guiding Principles state, that in granting rights to inject the CCS stream into subsurface formations, governments should give due consideration to land use planning issues that may arise as a consequence

¹⁹ Ministerial Council on Mineral and Petroleum Resources, "Carbon Dioxide Capture and Geological Storage: Australian Regulatory Guiding Principles", November 2005, http://www.industry.gov.au/assets/documents/itrinternet/Regulatory_Guiding_Principles_for_CCS20051124145652.pdf.

transportation Issues. Regulation relating to the transport of a CCS stream should be considered where possible, using agreed national protocols and guidelines of monitoring and verification. Still, the situation may not be as complex in Australia, while the legislation concerning pipeline infrastructure is always national. On the other hand, yet the different territories' legislation may affect on the land use policies concerning CCS.

4 Conclusions

Due to CCS-Directive of the EU, the legal framework for CCS operations exists in the EU. On the other hand, it is not enough as CCS operations may have transboundary effects from the point of view of the land use policy. The pipeline networks may be built in transboundary way. When keeping in mind the protection of the environment, the storage sites permit policies and the evaluation if they are suitable or not should be similar everywhere, avoiding 'the emissions transporting' to another country.

The regulation created in the EU Member States may learn from example of Australian legislation and the concept of liability. Even though the CCS-Directive regulates the land use policy from the aspect of choosing the storage site properly, Australian legislation seems to highlight the safety and liability of the storage site even more. Evaluating the liability of the storage sites should include for example identification and accounting of leakage liability and post-closure responsibilities and the attitude is, that the current regulatory principles and common law should continue to apply to liability issues for all stages of CCS projects.

The effects of the whole CCS operation should not underestimated from the point of view of the land use and the definition that what is it that the storage site or the pipeline networks include. Pipelines constructed for the exclusive use of a single power plant for on-site or nearby CO₂ storage could be considered an extension of the plant. Alternatively, they could be considered a non-plant asset providing a transportation service for a fee, in which case the costs could still be recovered by the utility in its rates as an operating cost. This could raise questions about cost recovery for electric utilities under state utility regulations.²⁰

Especially in the areas that are dense populated, the land use policy plays an important role. In this case, there may be differences in the Australian and the EU legislation as in Australia there is more rural area to be used. Landowners of land above or near storage projects may be directly affected. The effects include environmental effects as well as property rights issues must be concerned as well. Ensuring that the operator has obtained the right to access a private landowner's land for monitoring purposes and/or to use the subsurface pore space under that land is crucial. A full-scale storage project could affect many landowners. Projects may also face greater challenges in establishing

²⁰ Parfomak, P., and P. Folger 2008: Carbon Dioxide (CO₂) Pipelines for Carbon Sequestration: Emerging Policy Issues. RL33971. Washington, DC: Congressional Research Service at <http://openncrs.com/document/RL33971>.

agreements with landowners. These concerns should be addressed by policymakers through legislative clarity concerning key property rights issues.

So, even though the legal frame of the EU concerning CCS legislation is about to find its place, there are still issues waiting to be solved in the field of land use policies, especially concerning creating rules for determining suitable storage places and regulating transboundary CCS pipelines, whereas for example the legislation in Australia, that is more experienced with CCS-operations, could offer some models.

5 References

Department for Business Enterprise & Regulatory Reform (2007): “North Sea CCS Infrastructure Report to the North Sea Basin Task Force”, pp. 23-35.

DOE/NETL-2006/1236 (2006): “International Carbon Capture and Storage Projects Overcoming Legal Barriers”, <http://www.netl.doe.gov/energy-analyses/pubs/ccsregulatorypaperfinalreport.pdf>, viewed: 17/06/2010.

IPCC Special Report (2005): “Carbon Dioxide Capture and Storage- Summary for Policymakers: A Special Report of Working Group III of the Intergovernmental Panel on Climate Change”, http://www.ipcc.ch/pdf/special-reports/srccs/srccs_summaryforpolicymakers.pdf, viewed: 17/06/2010.

Mace, M. J., Hendriks, C., and Coenraads, R (2007): “Regulatory challenges to the implementation of carbon capture and geological storage within the European Union under EU and international law”, *International journal of greenhouse gas control*, <http://www.sciencedirect.com>, viewed: 27/06/2010.

Ministerial Council on Mineral and Petroleum Resources (2005): “Carbon Dioxide Capture and Geological Storage: Australian Regulatory Guiding Principles”, http://www.industry.gov.au/assets/documents/itinternet/Regulatory_Guiding_Principles_for_CCS20051124145652.pdf, viewed: 05/06/2010.

Parfomak, P., and P. Folger (2008): “Carbon Dioxide (CO₂) Pipelines for Carbon Sequestration: Emerging Policy Issues”, Congressional Research Service, <http://opencrs.com/document/RL33971>, viewed: 10/06/2010.

Sands, P. (2003): *Principles of international environmental law*, 2nd ed., pp. 143, 869-939.

Svensson, R., M. Odenberger, F. Johnsson, and L. Strömberg (2004): “Transportation Systems for CO₂—Application to Carbon Capture and Storage”, pp. 50-126.

World Resources Institute (2008): “Guidelines for Carbon Dioxide Capture, Transport, and Storage”, pp. 40-130.

Directive 2009/31/EC on the geological storage of carbon dioxide.

Directive 85/337/EEC of 27 June 1985 on the assessment of the effects of certain public and private projects on the environment.

The Barrow Island Act of 2003.

Offshore Petroleum Act 2006.

1972 Convention on the Prevention of Marine Pollution by Dumping of Wastes and Other Matter.

The 1996 London Protocol to the London Convention, entered into force in March 2006.

The Convention for the Protection of the Marine Environment of the North-East Atlantic (OSPAR) 10.12.1982.

<http://www.gorgongasproject.com/>. Viewed: 17/06/2010.

<http://eur-lex.europa.eu>. Viewed: 20/06/2010.

An Integrated Product Policy Approach to Products and their Impact on Energy

Rosalind Malcolm

School of Law, University of Surrey, Guildford, GU2 7XH
United Kingdom

Email: R.Malcolm@surrey.ac.uk

Abstract

A new piece of European legislation, the Ecodesign Directive (2009/125/EC) came into effect on 20 November 2009 and is to be implemented by Member States by November 20, 2010. This covers energy-using products under an earlier directive (the Energy using Products Directive) but also extends the range of products covered to include those which are related to energy use even if they do not actually use energy directly such as construction materials and fittings. So it covers 'any goods having an impact on energy consumption during use'. A product list is to be developed by the European Commission. Reduction in energy consumption is posited in the European Climate Change Programme (European Commission, 2006) and climate change is a priority in the Sixth Community Environment Programme (Decision No 1600/2002/EC of the European Parliament and of the Council). Climate change also now appears as an objective of the European Union in the Lisbon Treaty.

Ecodesign is a key element of the European Commission's Integrated Product Policy and appears in the Commission Communication of 18 June 2003 'Integrated Product Policy – Building on Environmental Life Cycle Thinking'. It is also a key innovative aspect appearing in the Sixth Community Environment Action Programme. Building legislation on a life cycle approach is radical and novel given that most environmental impacts are regulated on a vertical basis where legislation is linked to the process rather than the product. This paper argues that the best way forward for achieving sustainability is to revolutionise the regulatory process by adopting a life cycle approach to the environmental impacts of products as the basis for legislation. It examines the impact of the Ecodesign Directive and seeks to evaluate the pros and cons of such an approach

Keywords: Ecodesign, energy related products, EU, integrated product policy, legislation.

1. Introduction

The industrialised world has moved into an era of consumption¹. Our ability to consume; our desire for new and improved products, seems to know no end.² All products come from the earth in some direct or indirect form and our consumption frenzy further despoils natural resources.³ The inequity of this in relation both to the developing world and for future generations makes this pattern of behaviour the antithesis of sustainable development.⁴ At the same time as the west is the consumer paramount, production has tended to move to the developing world for the same reasons as it did in colonial eras – labour is cheap and plentiful and regulatory controls on public health and environmental protection are weak. The balance between the industrialised and the developing worlds becomes increasingly distorted.

This state of affairs has delineated the political agenda for the development of environmental policy along with concerns over climate change and its ability to affect the long-term health of the planet - and sustainable consumption and production policies abound. From injunctions to consume (and breed) less, to exhortations to buy 'green', and live 'sustainable lifestyles', the pressure is to change behaviour. Research from sociologists and psychologists focuses on behavioural issues – how to persuade people to change their ways. But a more subtle change which is barely vaunted in the public media is taking place. The development of fundamental changes to the way we view our products is changing the way in which law is impacting on this area and, it is arguable, has the capacity to make significant inroads in the achievement of some concept of sustainable development. New environmental laws have become the drivers of environmental and market change.

2. Integrated Product Policy

A little known policy established at European Union level nearly a decade ago is behind this change and has been creeping inexorably through the annals of legislative reform. On 18 June 2003 the Commission adopted its Communication 'Integrated Product Policy – building on environmental Life-Cycle Thinking'⁵ which was followed preparatory studies in 1998 and a Green Paper on Integrated Product Policy ('IPP')⁶ in 2001. The Communication established five key principles: 'life-cycle thinking; working with the market; wide involvement of stakeholders; continuous improvement of products; and co-ordinated use of policy instruments.' Three key strategic documents cemented IPP into

¹ Commission (EC) 'The State of Implementation of Integrated Product Policy' (Report) COM (2009)693 final, 21 December 2009 (hereinafter Report on the State of Implementation of Integrated Product Policy)

² Malcolm, R., *Environmental Product Policy: a new regulatory paradigm for a consumer society* European Environmental Law Review [2005] EELR 134-144

³ De Benedetto, L. and Klemeš, J. (2010), 'The environmental bill of material and technology routing: an integrated LCA approach' *Journal Clean Technologies and Environmental Policy* 12(2), pp. 191-196

⁴ Our Common Future, Report of the World Commission on Environment and Development, 1987. Published as Annex to General Assembly document A/42/427.

⁵ Commission (EC) 'Integrated Product Policy – building on environmental Life-Cycle Thinking' (Communication) COM (2003)302 final, 18 June 2003

⁶ Commission (EC) 'Integrated Product Policy' (Green Paper) COM (2001)68, 7 February 2001

environmental policy: the Strategy on Sustainable Development,⁷ the EU's Sixth Environmental Action Programme⁸ and the UN Ten-year Framework of programmes on Sustainable Consumption and Production.⁹

Integrated product policy is revolutionary in that it seeks to develop a coherent and holistic approach to minimising environmental impacts across the life cycle of a product by using a variety of tools so as to achieve “‘greener’ products that combine lower environmental impacts with enhanced service to consumers”.¹⁰ It takes a longitudinal view across the lifetime of the product setting out to achieve products which are greener and better – the idea is to improve the product overall not just its environmental performance. Amongst the policy tools advocated to achieve these objectives is improvement in product manufacturing and design. A single measure is not possible across the range of products and services so IPP looks to a ‘variety of policy tools and appropriate incentives for the market to move in a more sustainable direction by encouraging the supply and demand of greener products’.¹¹ Such tools as ‘product design, emission ceilings, labelling, fiscal measures and self-regulatory initiatives’ by industries are all part of IPP and it is anticipated that many different groups of the public need to be involved from all corners of industry, government and the consumers themselves.¹²

The longitudinal approach which underpins IPP is the life cycle approach¹³ which addresses the environmental impacts of the product from cradle to grave.¹⁴ Product development is a highly developed activity at the core of the free market economy and issues such as cost, performance and quality are all routinely considered.¹⁵ But the environmental impacts of a product have been less central to the concern of an industrialist in the past. Content to ignore the external costs of his product and to let those costs fall on society or the individual, this aspect has historically been ignored. But life cycle thinking has brought a scientific approach to the mechanism for addressing the impacts that a product will have on the environment during its lifetime.¹⁶ It aims to

⁷ Council of the European Union (2006) Review of the EU sustainable development strategy (EU SDS)—Renewed strategy. Brussels, 9 June 2006 (Document No 10117/06)

⁸ Commission (EC) ‘The sixth environment action programme of the European Community - ‘Environment 2010: Our future, Our choice’ (Communication) COM (2001) 31 final, 24 January 2001. In particular, the Sixth Environmental Action Programme called for the application of IPP to natural resources and waste policy.

⁹ See UN Department of Economic and Social Affairs, Division for Sustainable Development http://www.un.org/esa/dsd/dsd_aofw_scpp/scpp_tenyearframprog.shtml viewed 12/06/2010

¹⁰ Report on the State of Implementation of Integrated Product Policy

¹¹ Commission Staff Working Document Accompanying the Report from the Commission to the Council, the European Parliament, the European and Social Committee and the Committee of the regions On the State of Implementation of Integrated Product Policy COM(2009)693 final 21 December 2009 (hereinafter Commission Staff Working Document Accompanying the Report on the State of Implementation of Integrated Product Policy)

¹² Report on the State of Implementation of Integrated Product Policy

¹³ Baumann and Tillman 2004 as discussed in Vieira, P., Tenreiro, A. and Oliveira, T. (2010), ‘The increase of sustainability in cylinder manufacturing’ *Clean Technologies and Environmental Policy* 12(1), pp. 83-86

¹⁴ Baumann, H. och Tillman, A-M., *The hitchhiker's guide to LCA : an orientation in life cycle assessment methodology and application* (Lund, 2004)

¹⁵ Vieira, P., Tenreiro, A., and Oliveira, T. (2010) *The increase of sustainability in cylinder manufacturing* *Clean Technologies and Environmental Policy* 12(1), pp. 83-86.

¹⁶ European Commission, DG Enterprise and Industry ‘Ecodesign Your Future: How Ecodesign can help the environment by making products smart’

undertake an inventory of the impacts during the various stages of the life cycle from production to use and to disposal¹⁷ which are measured, using existing data.¹⁸ The object is to identify all the impacts so as to avoid moving the problem along the life cycle. The life of a product can be extensive and invasive – both geographically as well as in terms of the nature of the impact.¹⁹ The components of a product may be won, quarried and manufactured in far-flung places around the world far distant from the places where the product is used; and then its disposal may take place in yet another country or region. To identify the impacts across this chain, means that it is less likely that the country of production will be forced to bear the ills (whether to public health or to environment) created by the profligacy of the country (or people) of consumption. At the very least, there will be a better understanding of the life cycle of a product.

2.1 Practical application of life-cycle thinking within the EU

As a conceptual framework, IPP is necessarily vague establishing an idea rather than specific tools which have emerged subsequently in precise instruments which have started to implement life cycle thinking. In particular, it has been carried further as a central component of the Sustainable Consumption and Production and Sustainable Industrial Policy (SCP/SIP) Action Plan²⁰ which provides a coherent framework integrating a ‘mutually-enforcing set of existing and new policy instruments and initiatives (both mandatory and voluntary) to promote resource efficient and eco-friendly products and raise consumer awareness’.²¹ This includes Directive 2009/125/EC²² (the Ecodesign Directive), which is a direct transposition of IPP principles into product design legislation and which includes as one of its objectives: to combat climate change in a cost-effective manner. The term ecodesign aims at improving the environmental performance of products throughout their entire life cycle by systematically integrating environmental aspects at the earliest stage of product design,²³ an approach which avoids

http://ec.europa.eu/enterprise/policies/sustainable-business/ecodesign/files/brochure_ecodesign_en.pdf viewed 10/06/2010

¹⁷ Commission Staff Working Document Accompanying the Report on the State of Implementation of Integrated Product Policy.

¹⁸ Vieira and others (2010). Note the standardised LCA tool contained in ISO14040/44: *Environmental management - Life cycle assessment - Principles and framework*, International Organisation for Standardisation (ISO).

¹⁹ Commission Staff Working Document Accompanying the Report on the State of Implementation of Integrated Product Policy

²⁰ Commission (EC) ‘Sustainable Consumption and Production and Sustainable Industrial Policy Action Plan (Communication) COM (2008) 397 final, 25 June 2008

²¹ Commission Staff Working Document Accompanying the Report on the State of Implementation of Integrated Product Policy

²² Directive 2009/125/EC of the European Parliament and of the Council of 21 October 2009 establishing a framework for the setting of ecodesign requirements for energy-related products (‘Ecodesign Directive’) [2009] OJ L 285/10 (hereinafter ‘Ecodesign Directive’)

²³ Commission (EC) ‘Establishment of the Working Plan for 2009-2011 under the Ecodesign Directive’ (Communication) COM (2008)660 final, 21 October 2008; Sutherland, E. (2009), ‘Regulating the energy efficiency of ICT equipment’, *International Energy Law Review*, **8**, pp.282-284; European Commission ‘Energy efficiency – Ecodesign of Energy using Products’ (Directorate-General for Enterprise and Industry http://ec.europa.eu/energy/efficiency/ecodesign/eco_design_en.htm viewed 09/06/2010

uncoordinated product planning which could have a negative impact on the environment.²⁴ The preventive principle²⁵ lies solidly behind this.

Further legal instruments at EU level implementing IPP include the revised Ecolabel Regulation,²⁶ the revised Directive on Energy Performance of Buildings,²⁷ the revised Waste Framework Directive,²⁸ and the EMAS Regulation.²⁹ These are to be distinguished from another set of laws which impose duties and responsibilities on producers, namely, the Directive on Packaging and Packaging Waste;³⁰ the Waste Electrical and Electronic Equipment (WEEE) Directive; the End-of-Life Vehicles Directive,³¹ the Batteries Directive³² and the RoHS Directive³³ known as extended producer responsibility laws. In many cases, these laws focus on ‘end-of-life’ and deal only with the disposal aspect of the product. IPP is to be distinguished in that it deals with ‘whole life’.

Beside this legislation, various industry sectors such as packaging, chemical, plastics, metal and the construction industry are now applying life-cycle thinking to inform their approaches towards product policy and communication with customers and authorities.³⁴ Industry associations have also been involved in this work through the Consultation Forum. This participation is important as, particularly for SMEs, the practical application of life cycle assessment can be particularly challenging.

2.2 What are energy related products?

The Ecodesign Directive is concerned with energy-related products i.e. those that have an impact on energy use, and it grew out of its earlier incarnation as a Directive³⁵ which was limited to the consideration of products which *used* energy (the Energy using Products Directive). Energy related products are a significant contributor to the overall

²⁴ *Your Future: How Ecodesign can help the environment by making products smart* (EC Directorate-General for Enterprise and Industry) n. 4

²⁵ Article 191.2 Treaty on the Functioning of the European Union.

²⁶ Regulation (EC) No 1980/2000 of the European Parliament and of the Council of 17 July 2000 on a revised Community eco-label award scheme [2000] OJ L 237/1

²⁷ Directive 2002/91/EC of the European parliament and of the Council of 16 December 2002, on the energy performance of buildings [2002] OJ L 1/65

²⁸ Directive 2006/12/EC of the European Parliament and of the Council of 5 April 2006 on waste [2006] OJ L 114/9; Commission Staff Working Document Accompanying the Report on the State of Implementation of Integrated Product Policy

²⁹ Regulation (EC) No 1221/2009 of the European Parliament and of the Council of 25 November 2009 on the voluntary participation by organisations in a Community eco-management and audit scheme (EMAS), repealing Regulation (EC) No 761/2001 and Commission Decisions 2001/681/EC and 2006/193/EC [2009] OJ L 342/1

³⁰ Directive 94/62/EC of the European Parliament and of the Council of 20 December 1994 on packaging and packaging waste [1994] OJ L 365/10

³¹ Directive 2000/53/EC of the European Parliament and of the Council of 18 September 2000 on end-of life vehicles (the EVS) [2000] OJ L 269/34

³² Council Directive 91/157/EEC of 18 March 1991 on batteries and accumulators containing certain dangerous substances [1991] OJ L 78/38 (repealed by Directive 2006/66/EC of the European Parliament and of the Council of 6 September 2006 on batteries and accumulators and waste batteries and accumulators [2006] OJ L 266/1)

³³ Directive 2002/95/EC of the European Parliament and of the Council of 27 January 2003 on the restriction of the use of certain hazardous substances in electrical and electronic equipment (the RoHS Directive) [2003] OJ L 37/19
Report on the State of Implementation of Integrated Product Policy

³⁵ Directive 2005/32/EC of the European Parliament and of the Council of 6 July 2005 establishing a framework for the setting of ecodesign requirements for energy-using products.

consumption of energy and other natural resources and to control their design would have a potentially high impact for limiting the exploitation of resources and the use of energy.³⁶ As has been described, energy-related products (ERPs) are a broader category than energy-using products (EuPs) (which they subsume) and include products which ‘use, generate, transfer or measure energy (e.g. electricity, gas, fossil fuel and renewable energy sources), including consumer goods such as boilers, water heaters, computers, televisions, and industrial products such as transformers, industrial fans and industrial furnaces’. They also include other energy-related products which do not necessarily use energy but have an impact on energy and can therefore contribute to saving energy, such as windows, insulation material or bathroom devices (e.g. shower heads, taps).³⁷ The Directive is therefore a very significant driver in the construction industry and the built environment and, in this context, the inducement of a trade off between higher purchase price and lower running costs³⁸ can be available. As ERPs, some building products such as wall insulation or double glazing can immediately reduce heating costs³⁹ thus demonstrating local benefits (cheaper costs for the householder) and global benefits (reduced greenhouse gas emissions).⁴⁰ ‘The life-cycle of energy-related products is associated with a considerable number of important impacts on the environment, namely the consequences of energy consumption, consumption of other materials/resources, waste generation and release of hazardous substances to the environment’.⁴¹

2.3 Benefits of ecodesign

The key benefit of the Ecodesign Directive is the power to set harmonised energy efficiency requirements for all energy using products in the residential, tertiary and industrial sectors across the single market benefiting both trade and the environment.⁴² So, if, for instance, ecodesign requirements for household refrigerators have been set under the Ecodesign Directive, any refrigerator which does not comply with these requirements will not qualify for the CE-marking, and therefore cannot be sold in the European Union. The energy consumption standards set under the Directive lead to considerable energy savings: ‘by 2020, the first ecodesign measures on 9 product groups will allow energy savings equivalent to 12% of the electricity consumption of the EU in 2007 (compared to a ‘business as usual’ scenario)’.⁴³ Consistent with this approach, is an emphasis on the commercial market nature of the Directive and its EU focus. It only applies to products which have significant sales and trade in the EU and the ecodesign

³⁶ Establishment of the Working Plan for 2009-2011 under the Ecodesign Directive

³⁷ European Commission, Enterprise and Industry ‘Sustainable Business’

http://ec.europa.eu/enterprise/policies/sustainable-business/ecodesign/index_en.htm viewed 10/06/2010

³⁸ Kemna, R., van Elburg, M., Li, W. and van Holsteijn, R. (2005), ‘Methodology Study Eco-design of Energy-using Products’ http://ec.europa.eu/energy/demand/legislation/doc/2005_11_28_finalreport1_en.pdf viewed 02/07/2010

³⁹ Note also that in the UK, the British Standard BS ISO 15686-5:2008 Buildings and constructed assets, has been in effect since 2008. This covers planning for the service life of the construction and covers economic and environmental matters.

⁴⁰ Nishioka, Y., Levy, J.I. and Norris, G. A. (2006), ‘Integrating Air Pollution, Climate Change and Economics in a Risk based Life-Cycle Analysis: A Case Study of Residential Insulation’ *Human and Ecological Risk Assessment*, **12**, pp.552-571

⁴¹ European Commission ‘Energy efficiency – Ecodesign of Energy using Products’

http://ec.europa.eu/energy/efficiency/ecodesign/eco_design_en.htm viewed 09/06/2010

⁴² Energy efficiency – Ecodesign of Energy using Products; Sutherland, 2009

⁴³ Your Future: How Ecodesign can help the environment by making products smart

requirements must be cost-effective.⁴⁴ As Boardman (2004) demonstrates this is hugely significant since the EU represents almost 200 million households,⁴⁵ making the European market larger than that of North America and trades worldwide including Turkey, Malaysia, China, Korea and frequently has its policies used as a template by other countries (not just candidate countries) this is an enormous market and one which is highly influential.⁴⁶

The challenge for policy makers is to ensure that more efficient appliances are available so that the purchase of eco-friendly products becomes the norm. But some aspects of this policy can be counterintuitive. For instance, design for longevity would, at first sight seem desirable – less materials, less waste and so on. But, even though there are undoubted desirables in longevity, it may be counter-effective if it slows down the adoption of more resources-efficient new models by the market.⁴⁷ Further, as Boardman (2004) argues, consumer knowledge is critical – for example, the classification of electricity consumption for refrigerators is replicated in a large number of energy-using pieces of equipment: cars, houses, boilers, light bulbs, televisions. However, surveys show that consumers are not aware of this. As a result, the market for energy efficient products is not yet driven by informed consumer choice. Arguably many consumers would buy environmentally-friendly goods, if they had the information and knew they were available.⁴⁸

It is the case though that affordability and cost can be factors which outweigh ecodesign as a buying factor for the purchaser and there is concern that stress on ecodesign should not make the product unaffordable, or, indeed, have a negative impact on the competitiveness of the industry. The reverse is intended and the objective of the measures is to give manufacturers a competitive edge with the improved performance of their products.⁴⁹ Above all, nothing must interfere with the principles of the free movement of goods.

Trade and economic benefit is also viewed as one of the pluses of the Directive as the pressure for innovation can mean new products and therefore opportunities for profit. Just as one of the disadvantages of new developments is that this causes products to become out-of-date and not ‘the latest’ in technological enhancement, this can be an economic and an environmental advantage. The new ‘better’ product may mean that the shelf life of the old product is reduced but the enhanced environmental performance of the new product may outweigh any disadvantage in this. In the current worldwide recession, the European Commission has welcomed any drive towards ecological innovation as needful not only to fight climate change but also to provide welcome opportunities for economic advantage. It is anticipated that the extension of the scope of

⁴⁴ The Directive only applies to products where more than 200,000 units are sold per annum.

⁴⁵ European Environment Agency

⁴⁶ Boardman, 2004

⁴⁷ Kemna and others, 2005

⁴⁸ Boardman, 2004

⁴⁹ Your Future: How Ecodesign can help the environment by making products smart

the Ecodesign Directive will open up new business opportunities in particular for small and medium sized enterprises (SMEs) and for further cost efficient energy savings.⁵⁰

2.4 Ecodesign Directive

The Working Plan⁵¹ for the original Energy using Products Directive 2005/32/EC⁵² (the EuP Directive) specified that the Directive was a key component of the EU's policy for improving the energy and environmental performance of products on the internal market. Its potential to cover other environmentally significant products, namely all energy-related products, was highlighted in the Action Plan on Sustainable Consumption and Production and Sustainable Industrial Policy⁵³ (the SCP/SIP Action plan).⁵⁴ As a result, the scope of the EuP Directive was widened in 2009 to include energy-related products and it was recast as Directive 2009/125/EC⁵⁵ (the Ecodesign Directive) in order to include all energy-related products into the legal framework.⁵⁶ According to the Commission, the Ecodesign Directive can be considered the legislative cornerstone of the SCP/SIP Action Plan.⁵⁷ The Directive ensures free movement of products across Europe and encourages integration of ecodesign in small and medium-sized enterprises (SMEs).⁵⁸

The key principles of the amended and extended Ecodesign Directive have not changed and continue to focus on the identification of environmental aspects at the product design stage – the very embodiment of the preventive principle. Now compulsory minimum ecodesign requirements for energy-using products (such as different types of electrical appliances including boilers, water heaters, computers, televisions or industrial fans), can be applied to all energy-related products (except vehicles for transport) and covers all energy sources, 'including those products that do not consume energy during use, but have an indirect impact on energy consumption, and therefore they can be designed in such a manner as to save energy such as shower heads and other bathroom fittings, insulation materials and double-glazing'.⁵⁹ In an example provided by the Report, it is estimated that if the share of replacement windows being double-glazed would increase

⁵⁰ Europa, Rapid Press Releases 'More eco-friendly products for a low carbon future' (Brussels 24 April 2009) <http://europa.eu/rapid/pressReleasesAction.do?reference=IP/09/638> viewed 14/06/2010

⁵¹ Establishment of the Working Plan for 2009-2011 under the Ecodesign Directive

⁵² Directive 2005/32/EC of the European Parliament and of the Council of 6 July 2005 establishing a framework for the setting of ecodesign requirements for energy-using products and amending Council Directive 92/42/EEC and Directives 96/57/EC and 2000/55/EC of the European Parliament and of the Council [2005] OJ L 191/29 amended by Directive 2008/28/EC of the European Parliament and of the Council of 11 March 2008 amending Directive 2005/32/EC establishing a framework for the setting of ecodesign requirements for energy-using products, as well as Council Directive 92/42/EEC and Directives 96/57/EC and 2000/55/EC, as regards the implementing powers conferred on the Commission [2008] OJ L 81/48

⁵³ Report on the State of Implementation of Integrated Product Policy

⁵⁴ Communication on the Sustainable Consumption and Production and Sustainable Industrial Policy Action Plan

⁵⁵ Ecodesign Directive, n. 39

⁵⁶ Report on the State of Implementation of Integrated Product Policy; More eco-friendly products for a low carbon future

⁵⁷ Commission Staff Working Document Accompanying the Report on the State of Implementation of Integrated Product Policy

⁵⁸ Establishment of the Working Plan for 2009-2011 under the Ecodesign Directive

⁵⁹ European Commission, Enterprise and Industry 'Sustainable Business' http://ec.europa.eu/enterprise/policies/sustainable-business/ecodesign/index_en.htm viewed 10/06/2010

by 30%, it would allow additional energy savings of 55,000 GWh (27 Mt CO₂, or 2 to 3 nuclear power plants) by 2020.⁶⁰

2.5 The scope of the Ecodesign Directive

Although all energy-related products are potentially within its scope,⁶¹ as a framework directive, the Ecodesign Directive does not establish binding requirements, but provides rules and criteria for setting such requirements through daughter directives.⁶² There are two methods by which the implementing regulations establish mandatory specific requirements: either by setting limit values or by establishing generic requirements which may require, for example, that the energy consumption of the product be as low as possible or conforms to a set standard. There may also be a requirement for the provision of information such as guidance on best usage from the manufacturer or they may require a lifecycle analysis of the product in order to identify ‘alternative design options and solutions for improvement’.⁶³

A large number of implementing measures have already been made under the Directive which apply to individual products or groups of products.⁶⁴ The Commission initially identified and tendered for 14 studies of products which are likely to be significantly affected by potential implementation measures⁶⁵ and further candidate products have been identified.⁶⁶ The focus is on those domestic appliances which are traded extensively within the EU,⁶⁷ which have a significant environmental impact,⁶⁸ and significant potential for improvement in terms of environmental impact.⁶⁹ Where a product meets the above criteria, it is covered by implementing measures or by a self-regulation measure in accordance with Article 15(3)(b).⁷⁰ By July 2010, the Directive has been applied to, various energy-using products including, *inter alia*, fluorescent lamps,⁷¹ non-

⁶⁰ More eco-friendly products for a low carbon future

⁶¹ Forbes, R. (2009), ‘New Ecodesign Directive Replaces Energy using Products Directive’ <http://www.internationallawoffice.com/newsletters/detail.aspx?g=bb020ec2-2ba8-4238-a40d-fc3ade3640d3> viewed 02/07/2010

⁶² European Commission, Enterprise and Industry ‘Sustainable Business’ http://ec.europa.eu/enterprise/policies/sustainable-business/ecodesign/index_en.htm viewed 10/06/2010

⁶³ Your Future: How Ecodesign can help the environment by making products smart

⁶⁴ Department for Business Innovation and Skills <http://www.berr.gov.uk/policies/business-sectors/environmental-and-technical-regulations/eco-design-directive> viewed 14/06/2010

⁶⁵ Garrett, P., Aumônier, S., Madsen J. and Wallen, E. ‘Environmental Tradeoffs of the Energy-using Products (EuP) Directive and Product Policy’ <http://www.lcm2007.org/paper/509.pdf> viewed 02/06/2010

⁶⁶ Department for Business Innovation and Skills <http://www.berr.gov.uk/policies/business-sectors/environmental-and-technical-regulations/eco-design-directive> viewed 14/06/2010

⁶⁷ Ecodesign Directive, Article 15(2)(a) ‘the product shall present a significant volume of sales and trade, indicatively more than 200 000 units a year’

⁶⁸ *Ibid*, Article 15(2)(b) ‘... as specified in the Community strategic properties as set out in Decision No 1600/2002/EC’

⁶⁹ *Ibid*, Article 16(2)

⁷⁰ *Ibid*, Article 15(1)

⁷¹ Commission Regulation (EU) No 347/2010 of 21 April 2010 amending Commission Regulation (EC) No 245/2009 as regards the ecodesign requirements for fluorescent lamps without integrated ballast, for high intensity discharge lamps, and for ballasts and luminaires able to operate such lamps [2010] OJ L 104/20. It is amended in order to avoid unintended impacts on the availability and performance of the products covered by Regulation No 245/2009 and to improve coherence, as regards the requirements on product information, between Regulation (EC) No 245/2009 and Regulation (EC) No 244/2009

directional household light bulbs,⁷² with the phasing out of traditional incandescent lamps.⁷³

Article 2 defines the term ‘energy-related product’ as ‘any good that has an impact on energy consumption during use which is placed on the market and/or put into service, and including parts intended to be incorporated into energy-related products covered by this Directive which are placed on the market and/or put into service as individual parts for end-users and of which the environmental performance can be assessed independently’.⁷⁴ When a product is covered by a measure under the Ecodesign Directive, Article 3(1) also requires that it shall bear the CE (European conformity) mark.⁷⁵ Article 5 requires a declaration of conformity with the measures to be issued by the manufacturer or authorised representative in the EU. Before placing a product on the EU market, a manufacturer or authorised representative must also carry out a conformity assessment procedure in order to ensure the product's conformity according to one of the systems mentioned in Annex IV or V of the Directive.⁷⁶ Once CE marked, a product can be placed anywhere on the EU market.⁷⁷ Furthermore, in accordance with the implementing measures, manufacturers will be obliged to provide certain information to consumers,⁷⁸ such as an ecological profile of the product, describing the materials, emissions and waste associated with the product throughout its lifecycle and statement of the benefits of ecodesign and the role that consumers can play in the sustainable use of the product.⁷⁹ If the product is being imported into the EU then Article 4 requires that these obligations fall on the importer.

3. Ecodesign Working Plan and Preparatory Studies

A Commission Communication: *Establishment of the Working Plan for 2009 – 2011 under the Ecodesign Directive*⁸⁰ sets out an indicative list of product groups which are considered as priorities for the adoption of implementing measures.⁸¹ The process for regulating a product under the Directive is to undertake a preparatory study and an impact assessment conducted by external experts and the Commission with the aim of identifying cost-effective solutions to improve the overall environmental performance of the product. Implementing measures are eventually adopted by the Commission under the regulatory procedure with scrutiny.⁸²

⁷² Commission Regulation (EC) No 859/2009 of 18 September 2009 amending Regulation (EC) No 244/2009 as regards the ecodesign requirements on ultraviolet radiation of non-directional household lamps [2009] OJ L 247/3

⁷³ Sutherland, 2009

⁷⁴ Ecodesign Directive, Article 2.

⁷⁵ Forbes, 2009

⁷⁶ Ecodesign Directive, Article 8

⁷⁷ Ecodesign Directive, Article 9

⁷⁸ Ecodesign Directive, Article 14

⁷⁹ New Ecodesign Directive Replaces Energy using Products Directive

⁸⁰ COM(2008) 660 final

⁸¹ European Commission ‘Energy’ http://ec.europa.eu/energy/efficiency/ecodesign/working_plan_en.htm viewed 09/06/2010

⁸² Establishment of the Working Plan for 2009-2011 under the Ecodesign Directive

The indicative list of products to be investigated in priority by October 2011 includes products like commercial refrigerating equipment, sound and imaging equipment (such as DVD players and game consoles), food preparing equipment (such as ovens, hobs and grills, coffee machines) and machine tools.⁸³ A second Working Plan for the period of 2011- 2014 is expected to be drawn up in 2011 and there will be a review of the effectiveness of the Ecodesign Directive and its implementing measures no later than 2012.⁸⁴ Article 21 specifies that the review will evaluate the methodology⁸⁵ for the identification and coverage of significant environmental parameters such as resource efficiency, taking into consideration the experience gained with the first implementing measures⁸⁶ and considering the whole life cycle of products; the threshold for implementing measures; market surveillance mechanisms; and self-regulation.⁸⁷ Based on this review, the Commission will assess the appropriateness of further extending the scope of the Directive beyond energy related products, in order to reduce environmental impacts throughout the whole life cycle of these products.⁸⁸ Extending its range to building materials is an obvious next step.⁸⁹

3.1 Transitional period 2005-2008

The original Energy using Products Directive was adopted in 2005 and implementing measures were to be introduced for the product groups mentioned in Article 16(2) of the original Directive. This made a link directly to the European Climate Change Programme where a number of energy-using products were identified as priorities, such as 'heating and water-heating equipment, electric motor systems, lighting in both the domestic and tertiary sectors, domestic appliances, office equipment in both the domestic and tertiary sectors, consumer electronics and heating/ventilating/air conditioning (HVAC) systems'. The product groups in this transitional phase were: tertiary-sector lighting equipment (covering both public street lighting and office lighting), stand-by and off-mode electricity losses, external power supplies and simple set-top boxes for digital reception of television signals, televisions, domestic lighting, domestic refrigerators and freezers, washing machines, dishwashers, boilers and water heaters, commercial refrigerators, electric motors, pumps, fans, circulators and room air-conditioners.⁹⁰ (Annex II lists the product groups covered in the transitional period.)

3.2 Indicative list of product groups

Under the Working Plan 2009-2011 an indicative list of product groups, taking into account the work done in the transitional period were listed as indicative priorities for preparatory studies and implementing measures for the years 2009-2012. These include:

⁸³ Ecodesign Your Future: How Ecodesign can help the environment by making products smart

⁸⁴ Ecodesign Directive, Article 21; More eco-friendly products for a low carbon future; How Ecodesign can help the environment by making products smart

⁸⁵ Methodology for the Ecodesign of Energy using Products (MEEUP)

⁸⁶ Ecodesign Your Future: How Ecodesign can help the environment by making products smart

⁸⁷ Ecodesign Directive, Article 21(a)-(d)

⁸⁸ Ecodesign Directive, Article 21; More eco-friendly products for a low carbon future; Ecodesign Your Future: How Ecodesign can help the environment by making products smart

⁸⁹ D M Gann, Y Wang and R Hawkins 'Do regulations encourage innovation? the case of energy efficiency in housing' (1998) Building Research & Information 26(4), 280-296

⁹⁰ Establishment of the Working Plan for 2009-2011 under the Ecodesign Directive

⁹¹ ‘Air-conditioning and ventilation systems; Electric and fossil-fuelled heating equipment; Food-preparing equipment; Industrial and laboratory furnaces and ovens; Machine tools; Network, data processing and data storing equipment; Refrigerating and freezing equipment; Sound and imaging equipment; Transformers; and Water-using equipment’.⁹²

Initially a Commission study⁹³ for preparing the working plan identified 57 product groups which fall within the scope of the Ecodesign Directive which were not covered in the transitional period.⁹⁴ In order to identify the product groups with the highest potential for reduction of greenhouse gas emissions the study screened the 57 product groups against the principal environmental impact – primary energy consumption in the use phase.⁹⁵ This resulted in 25 product groups ranked A and 9 ranked B.⁹⁶

The Commission further assessed the 25 A-ranked product groups with a view to prioritisation based on the set of criteria laid down in Article 15 of the Ecodesign Directive, namely that: (1) the product group represents a significant volume of sales and trade within the EU; (2) the product group has a significant environmental impact within the Community resulting from the energy-using products during their life cycle; and (3) the product group presents significant potential for improvement in terms of its environmental impact without entailing excessive costs.⁹⁷

3.3 Conclusions

One of the problems for a product policy is where there is no natural turnover – no reason to buy new. A product policy works most effectively if the product was going to be purchased anyway. Boardman (2004) describes the problem of a situation where there is no natural turnover rate, as with cavity wall insulation where there is no natural replacement cycle, and therefore ‘householders have to be inspired to take action’.⁹⁸ Once consumers are stimulated, the response can occur quickly since they are able to take action soon, as they do not have to wait for a product to break down.⁹⁹ In addition, product policy does not have an immediate negative impact on those on lower incomes in the same way fuel price rises do.¹⁰⁰ There may be a problem in that a product policy is aimed at the rich – you need to have money to replace goods. So there may need to be

⁹¹ Excluding energy-using products covered in the transitional period.

⁹² Establishment of the Working Plan for 2009-2011 under the Ecodesign Directive

⁹³ EPTA Ltd, Greece; PE International, Germany; NTUA, Greece: Study for preparing the first Working Plan of the Ecodesign Directive, Report for tender No: ENTR/06/026, Revised Final Report: 06/12/2007: http://ec.europa.eu/enterprise/eco_design/workingplan.htm.

⁹⁴ Study for preparing the first Working Plan of the Ecodesign Directive, p. 29

⁹⁵ Establishment of the Working Plan for 2009-2011 under the Ecodesign Directive

⁹⁶ Study for preparing the first Working Plan of the Ecodesign Directive, p. 33

⁹⁷ Establishment of the Working Plan for 2009-2011 under the Ecodesign Directive

⁹⁸ Boardman, 2004

⁹⁹ Boardman, 2004

¹⁰⁰ Boardman and Fawcett, 2002 cited in Boardman, 2004

initiatives or public sector funding to enable the poor to buy equipment which is more energy efficient.¹⁰¹

Further, clear development of product policy may be necessary to ensure that the bonus of environmental benefit is not just a by-product of product development. Markets on their own may not be catalysts for environmental gains and it is precisely in that situation where integrated product policy and some manifestation of its tools become essential. The serendipitous nature of the achievement of energy efficiency is too chancy to be relied upon. As Boardman (2004) argues: it is the role of manufacturers to persuade us to consume their products whereas it is the role of policy makers and consumers to achieve the 'wider social goal of reducing environmental impact'.

But the key change identified here is the markedly different approach to environmental regulation which the introduction of product policy has shifted from a vertical approach which tackled key issues of pollution along the chain to a preventive approach which rests on a longitudinal analysis across the whole life of a product of its impacts on the environment. This is truly a major shift in the direction of the achievement of a logical and rational approach to development which can be sustained environmentally for future generations and between different regions of the world today.

Environmental product policy can be a powerful tool, but depends upon political will, at both the Member State and the EU level, to ensure that manufacturers are given clear direction to ensure they bring on board environmental life cycle thinking in the design and development of their products otherwise new products will continue to prioritise the best interest of the manufacturer which may not include environmental issues. Weak policies will not deliver reduced energy consumption¹⁰².

4. Acknowledgments

I wish to gratefully acknowledge the assistance given by Sanna Elfving, PGR candidate in the School of Law, University of Surrey in the preparation of this paper.

5. References

Boardman, B. (2004), 'Achieving energy efficiency through product policy: the UK experience' *Environmental Science & Policy* 7, pp. 165–176

COCIR (2009), 'Self-Regulatory Initiative - EcoDesign of Energy using Products for Medical Imaging Equipment' http://www.ecodesign-info.eu/documents/Medical_imaging_equipment_VA_20Oct09.pdf viewed 15/06/2010

¹⁰¹ In the UK, low-income households received a new efficient fridge for the price of a second-hand one provided they relinquished a working, old appliance. This was funded by the UK electricity utilities under a regulatory requirement to achieve energy savings.

¹⁰² Boardman, 2004

Commission (EC) 'Action Plan for Energy Efficiency: Realising the Potential' (Communication) COM (2006) 545 final, 19 October 2006

Commission (EC) 'Establishment of the Working Plan for 2009-2011 under the Ecodesign Directive' (Communication) COM (2008)660 final, 21 October 2008

Commission (EC) 'Integrated Product Policy – building on environmental Life-Cycle Thinking' (Communication) COM (2003)302 final, 18 June 2003

Commission (EC) 'Integrated Product Policy' (Green Paper) COM (2001)68, 7 February 2001

Commission (EC) 'Policy Instruments to Reduce Stand-by Losses of Consumer Electronic Equipment' (Communication) COM (1999)120, 15 March 1999

Commission (EC) 'Proposal for a Directive of the European Parliament and of the Council establishing a framework for the setting of ecodesign requirements for energy-related products' COM (2008) 399 final, 16 July 2008

Commission (EC) 'Sustainable Consumption and Production and Sustainable Industrial Policy Action Plan (Communication) COM (2008) 397 final, 25 June 2008

Commission (EC) 'The sixth environment action programme of the European Community - 'Environment 2010: Our future, Our choice' (Communication) COM (2001) 31 final, 24 January 2001

Commission (EC) 'The State of Implementation of Integrated Product Policy' (Report) COM (2009)693 final, 21 December 2009

Commission Regulation (EC) No 107/2009 of 4 February 2009 implementing Directive 2005/32/EC of the European Parliament and of the Council with regard to ecodesign requirements for simple set-top boxes [2009] OJ L 36/8

Commission Regulation (EC) No 1275/2008 of 17 December 2008 implementing Directive 2005/32/EC of the European Parliament and of the Council with regard to ecodesign requirements for standby and off mode electric power consumption of electrical and electronic household and office equipment [2008] OJ L 339/45

Commission Regulation (EC) No 244/2009 of 18 March 2009 implementing Directive 2005/32/EC of the European Parliament and of the Council with regard to ecodesign requirements for non-directional household lamps [2009] OJ L 76/3

Commission Regulation (EC) No 245/2009 of 18 March 2009 implementing Directive 2005/32/EC of the European Parliament and of the Council with regard to ecodesign requirements for fluorescent lamps without integrated ballast, for high intensity discharge

lamps, and for ballasts and luminaires able to operate such lamps, and repealing Directive 2000/55/EC of the European Parliament and of the Council [2009] OJ L 76/17

Commission Regulation (EC) No 278/2009 of 6 April 2009 implementing Directive 2005/32/EC of the European Parliament and of the Council with regard to ecodesign requirements for no-load condition electric power consumption and average active efficiency of external power supplies [2009] OJ L 93/3

Commission Regulation (EC) No 640/2009 of 22 July 2009 implementing Directive 2005/32/EC of the European Parliament and of the Council with regard to ecodesign requirements for electric motors [2009] OJ L 191/26

Commission Regulation (EC) No 641/2009 of 22 July 2009 implementing Directive 2005/32/EC of the European Parliament and of the Council with regard to ecodesign requirements for glandless standalone circulators and glandless circulators integrated in products [2009] OJ L 191/35

Commission Regulation (EC) No 642/2009 of 22nd July 2009 implementing Directive 2005/32/EC of the European Parliament and of the Council with regard to ecodesign requirements for televisions [2009] OJ L 191/42

Commission Regulation (EC) No 643/2009 of 22 July 2009 implementing Directive 2005/32/EC of the European Parliament and of the Council with regard to ecodesign requirements for household refrigerating appliances [2009] OJ L 191/53

Commission Regulation (EC) No 859/2009 of 18 September 2009 amending Regulation (EC) No 244/2009 as regards the ecodesign requirements on ultraviolet radiation of non-directional household lamps [2009] OJ L 247/3

Commission Regulation (EU) No 347/2010 of 21 April 2010 amending Commission Regulation (EC) No 245/2009 as regards the ecodesign requirements for fluorescent lamps without integrated ballast, for high intensity discharge lamps, and for ballasts and luminaires able to operate such lamps [2010] OJ L 104/20

Commission Staff Working Document Accompanying the Report from the Commission to the Council, the European Parliament, the European and Social Committee and the Committee of the regions On the State of Implementation of Integrated Product Policy COM(2009)693 final, 21 December 2009

Council Directive 1999/31/EC of 26 April 1999 on the landfill of waste [1999] OJ L 182/1

Council Directive 91/157/EEC of 18 March 1991 on batteries and accumulators containing certain dangerous substances [1991] OJ L 78/38 (repealed by Directive 2006/66/EC of the European Parliament and of the Council of 6 September 2006 on batteries and accumulators and waste batteries and accumulators [2006] OJ L 266/1)

Council Directive 91/157/EEC of 18 March 1991 on batteries and accumulators containing certain dangerous substances [1991] OJ L 78/38

Council Directive 92/75/EEC of 22 September 1992 on the indication by labelling and standard product information of the consumption of energy and other resources by household appliances [1992] OJ L 297/16

Darby, S. (2008), 'Energy feedback in buildings: improving the infrastructure for demand reduction' *Building Research & Information* **36**(5), 499-508

De Benedetto, L. and Klemeš, J. (2010), 'The environmental bill of material and technology routing: an integrated LCA approach' *Journal Clean Technologies and Environmental Policy* **12**(2), pp. 191-196

Department for Business Innovation and Skills 'Eco-design of Energy-related products (ERP) Directive' <http://www.berr.gov.uk/policies/business-sectors/environmental-and-technical-regulations/eco-design-directive> viewed 14/06/2010

Directive 2000/53/EC of the European Parliament and of the Council of 18 September 2000 on end-of life vehicles (the EVS) [2000] OJ L 269/34

Directive 2002/91/EC of the European parliament and of the Council of 16 December 2002, on the energy performance of buildings [2002] OJ L 1/65

Directive 2002/95/EC of the European Parliament and of the Council of 27 January 2003 on the restriction of the use of certain hazardous substances in electrical and electronic equipment (the RoHS Directive) [2003] OJ L 37/19

Directive 2002/96/EC of the European Parliament and of the Council of 27 January 2003 on waste electrical and electronic equipment (the WEEE Directive) [2003] OJ L 37/24

Directive 2005/32/EC of the European Parliament and of the Council of 6 July 2005 establishing a framework for the setting of ecodesign requirements for energy-using products and amending Council Directive 92/42/EEC and Directives 96/57/EC and 2000/55/EC of the European Parliament and of the Council (the EuP Directive) [2005] OJ L 191/29

Directive 2006/12/EC of the European Parliament and of the Council of 5 April 2006 on waste [2006] OJ L 114/9

Directive 2006/32/EC of the European Parliament and of the Council of 5 April 2006 on energy end-use efficiency and energy services and repealing Council Directive 93/76/EEC [2006] OJ L 114/64

Directive 2006/42/EC of the European Parliament and of the Council of 17 May 2006 on machinery, and amending Directive 95/16/EC (recast) [2006] OJ L 157/24

Directive 2006/66/EC of the European Parliament and of the Council of 6 September 2006 on batteries and accumulators and waste batteries and accumulators [2006] OJ L 266/1

Directive 2008/1/EC of the European Parliament and the Council of 15 January 2008 concerning integrated pollution prevention and control, codified version [2008] OJ L 24/8

Directive 2008/12/EC of the European Parliament and the Council of 11 March 2008 amending Directive 2006/66/EC on batteries and accumulators and waste batteries and accumulators [2008] OJ L 76/39

Directive 2008/28/EC of the European Parliament and of the Council of 11 March 2008 amending Directive 2005/32/EC establishing a framework for the setting of ecodesign requirements for energy-using products, as well as Council Directive 92/42/EEC and Directives 96/57/EC and 2000/55/EC, as regards the implementing powers conferred on the Commission [2008] OJ L 81/48

Directive 2008/98/EC of the European Parliament and the Council of 19 November 2008 on waste and repealing certain Directives [2008] OJ L 312/3

Directive 2009/125/EC of the European Parliament and of the Council of 21 October 2009 establishing a framework for the setting of ecodesign requirements for energy-related products (the Eco-Design Directive) [2009] OJ L 285/10

Directive 2010/30/EU of the European Parliament and of the Council of 19 May 2010 on the indication by labelling and standard product information of the consumption of energy and other resources by energy-related products [2010] OJ L 153/1

Directive 94/62/EC of the European Parliament and of the Council of 20 December 1994 on packaging and packaging waste [1994] OJ L 365/10

Directive 96/57/EC of the European Parliament and of the Council of 3 September 1996 on energy efficiency requirements for household electric refrigerators, freezers and combinations thereof [1996] OJ L 236/36

Draft Commission Regulation implementing Directive 2009/125/EC of the European Parliament and of the Council with regard to ecodesign requirements for household dishwashers (not published in the Official Journal yet)

EurActive 'EU boiler efficiency negotiations drag on' (Brussels 17 May 2010)
<http://www.euractiv.com/en/energy-efficiency/eu-boiler-efficiency-negotiations-drag-news-494228> viewed 11/06/2010

Europa 'Energy efficiency for the 2020 goal'

http://europa.eu/legislation_summaries/energy/energy_efficiency/en0002_en.htm viewed 15/06/2010

Europa, Rapid Press Releases 'More eco-friendly products for a low carbon future' (Brussels 24 April 2009)

<http://europa.eu/rapid/pressReleasesAction.do?reference=IP/09/638> viewed 14/06/2010

European Commission, 'Energy efficiency – Eco-design of Energy-Using Products'

http://ec.europa.eu/energy/efficiency/ecodesign/eco_design_en.htm viewed 09/06/2010

European Commission, DG Enterprise and Industry 'Ecodesign Your Future: How Ecodesign can help the environment by making products smart'

http://ec.europa.eu/enterprise/policies/sustainable-business/ecodesign/files/brochure_ecodesign_en.pdf viewed 10/06/2010

European Commission, Enterprise and Industry 'Sustainable Business'

http://ec.europa.eu/enterprise/policies/sustainable-business/ecodesign/index_en.htm viewed 10/06/2010

European Commission, 'Manage Energy'

<http://www.managenergy.net/products/R535.htm> viewed 01/07/2010

European Commission, 'Guide to the Implementation of Directives Based on New Approach and Global Approach'

<http://ec.europa.eu/enterprise/newapproach/legislation/guide/index.htm> viewed 11/06/2010

European Environment Agency 'EN18 Electricity Consumption'

http://themes.eea.europa.eu/Sectors_and_activities/energy/indicators/EN18%2C2008.11/EN18_Electricity_Consumption.pdf viewed 02/07/2010

Gann, D. M., Wang, Y. and Hawkins, R. (1998) 'Do regulations encourage innovation? the case of energy efficiency in housing' *Building Research & Information* **26**(4), 280–296

Garrett, P., Aumônier, S., Madsen J. and Wallen, E. 'Environmental Tradeoffs of the Energy-using Products (EuP) Directive and Product Policy'

<http://www.lcm2007.org/paper/509.pdf> viewed 02/07/2010

Forbes, R. (2009), 'New Eco-Design Directive Replaces Energy-Using Products Directive'

<http://www.internationallawoffice.com/newsletters/detail.aspx?g=bb020ec2-2ba8-4238-a40d-fc3ade3640d3> viewed 02/07/2010

Nishioka, Y., Levy, J.I. and Norris, G. A. (2006), 'Integrating Air Pollution, Climate Change and Economics in a Risk based Life-Cycle Analysis: A Case Study of Residential Insulation' *Human and Ecological Risk Assessment*, **12**, pp.552-571

Commission Directive 97/17/EC of 16 April 1997 implementing Council Directive 92/75/EEC with regard to energy labelling of household dishwashers [1997] OJ L 118/1

Regulation (EC) No 106/2008 on a Community energy-efficiency labelling programme for office equipment [2008] OJ L 39/1

Regulation (EC) No 1221/2009 of the European Parliament and of the Council of 25 November 2009 on the voluntary participation by organisations in a Community eco-management and audit scheme (EMAS), repealing Regulation (EC) No 761/2001 and Commission Decisions 2001/681/EC and 2006/193/EC [2009] OJ L 342/1

Regulation (EC) No 1275/2008 implementing ecodesign requirements for standby and off mode electric power consumption of electrical and electronic household and office equipment [2008] OJ L 339/45

Regulation (EC) No 1907/2006 concerning the Registration, Evaluation, Authorisation and Restriction of Chemicals (REACH) and establishing a European Chemicals Agency [2006] OJ L 396/1

Regulation (EC) No 1980/2000 of the European Parliament and of the Council of 17 July 2000 on a revised Community eco-label award scheme [2000] OJ L 237/1

Regulation (EC) No 2422/2001 of the European Parliament and of the Council of 6 November 2001 on a Community energy efficiency labelling programme for office equipment [2001] OJ L332/1

Council of the European Union (2006) Review of the EU sustainable development strategy (EU SDS)—Renewed strategy. Brussels, 9 June 2006 (Document No 10117/06)

Steenblik, R., Vaughan, S. and Waide P. (2006), 'Can Energy-Efficient Electrical Appliances be considered "Environmental Goods"?' International Energy Agency/OECD Trade and Environment Working Paper No. 2006-04
<http://www.oecd.org/dataOECD/20/5/37321363.pdf> viewed 02/07/2010

Sutherland, E. (2009), 'Regulating the energy efficiency of ICT equipment', *International Energy Law Review*, **8**, pp.282-284

The Department for the Environment, Food and Rural Affairs (Defra) <http://efficient-products.defra.gov.uk/eup-implementing-measures/> viewed 11/06/ 2010

Kemna, R., van Elburg, M., Li, W. and van Holsteijn, R. (2005), 'Methodology Study Eco-design of Energy-using Products'

http://ec.europa.eu/energy/demand/legislation/doc/2005_11_28_finalreport1_en.pdf
viewed 02/07/2010

Vieira, P., Tenreiro, A. and Oliveira, T. (2010), 'The increase of sustainability in cylinder manufacturing' *Clean Technologies and Environmental Policy* **12**(1), pp. 83-86

Working Document For Voluntary Agreement Lot 4; Imaging Equipment
http://ec.europa.eu/energy/efficiency/ecodesign/doc/forum/2009_10_12/2009_10_12_voluntary_industry_agreement_eup_lot4.pdf viewed 11/06/2010

Retrofits/ renewals in buildings for energy efficiency and the building control regulations in local authorities in South Africa

Wafula J.E.¹ and Talukhaba A.A.¹

¹School of Construction Economics and Management,
University of the Witwatersrand, Johannesburg

Abstract:

More than one-third of energy is consumed in buildings worldwide, accounting for about 15 percent of global greenhouse gas emissions. In cities, buildings can account for up to 80 percent of CO₂ emissions. The built environment is therefore a critical part of the climate change problem – and solution. Most existing buildings were not designed for energy efficiency, but by retrofitting with up-to-date products, technologies and systems, a typical building can realize significant energy savings. Improving the energy efficiency of buildings is a priority for reducing both greenhouse gas emissions and energy costs.

In South Africa, the developmental Local Government paradigm makes local authorities the focal points of implementing the various regulations, standards and codes for energy efficiency in buildings set by the national and provincial governments. This borrows heavily from the often repeated phrase, “Think globally, and act locally”. Local authorities provide a necessary and practical supplement to national government efforts and deliver a substantial impact by just changing their regulations. Besides, they are mandated by statute, and have put in place mechanisms whereby all new buildings and retrofits/refurbishments/renewals to buildings under their jurisdictions have to go through their building control and approval departments before they are implemented.

This paper investigates the current building control and approvals regulations in local authorities in relation to energy efficiency policies for buildings in general, and for retrofits/refurbishments/renewals in particular. The paper highlights the building regulations & control mechanisms, with regard to energy efficiency for retrofits/refurbishments, in the key Metropolis of South Africa. The National policy proposal for energy efficiency in buildings, with emphasis on retrofits/refurbishment was reviewed too. It compared this with promising and innovative regulations and controls which deliver higher energy efficiency in retrofits/refurbishments in buildings in other countries to make deductive conclusions.

It has been found that City governments, commercial buildings/property portfolio owners, schools, universities, and public housing authorities can identify, design, and implement energy efficiency retrofits in all renewal building projects. Different mechanisms for increasing building efficiency include retro-commissioning, building technology replacement, and energy performance contracting.

These mechanisms can be perfected by the building regulations control departments of the various local authorities to be best practices and can be widely shared and easily adopted among them.

In conclusion, Building on the experience gained through initial retrofit projects along with awareness & education and developing sustainable financial instruments, Local Authorities can partner other organizations like the Green Building Council of South Africa in developing standardized procedures in the building regulations and control process, to accelerate the implementation of energy efficient retrofits nationwide.

Key words: Energy Efficiency, Retrofits/Refurbishment/Renewals, Local Authorities, Buildings, Building Regulations and control, Approvals, Policy.

1.0 Introduction

Buildings consume energy throughout their lifecycle - from construction to operation and then demolition (City of Johannesburg, et al, 2008). Buildings in the US are responsible for 71% of the country's total electricity consumption, and 33% of emissions (Clinton Climate Initiative, 2006). In the quest for stability in electricity supply and eventual energy security, South Africa needs to emphasise energy/electricity demand reduction via energy efficiency and real-time energy scheduling apart from increasing the supply sources.

In 2004, there were 11.2 million dwelling units in the country, and 70.6 million sq m of non-residential building space. In Pretoria, cooling and lighting alone was responsible for 75% of energy use in office buildings (CSIR 2004). Cost effective energy efficiency retrofits to this existing building stock can result in major energy savings.

Ward and Schäffler (2008) in a trends paper for Gauteng Province in South Africa, clearly outline energy security, carbon mitigation and environmental impacts, and economic development and competitiveness among other things as the main motivation behind the urge to use renewable energy and implement energy efficiency policies. Further on, Ward and Schäffler (2008) report that energy consumptions in Johannesburg and surrounding metropolis indicate that industry accounts for 16% and transport for 60% of energy demand, Domestic use 16% and government, commerce and local authority 7% (refer to Figure 1). The combined energy use by the domestic, government, commerce and local authority sectors contribute to a total of 22%; the energy in this sector is mainly used in buildings thus emphasising the need to implement energy efficiency policies in that area.

Retrofits of buildings is a sure way of improving energy efficiency in buildings, this is supported by Maisely and Beverly (2007) and Reeves (2008) who reiterate that it reduces energy consumption by the elimination of aged components and also by ensuring that the technology used is not obsolete.

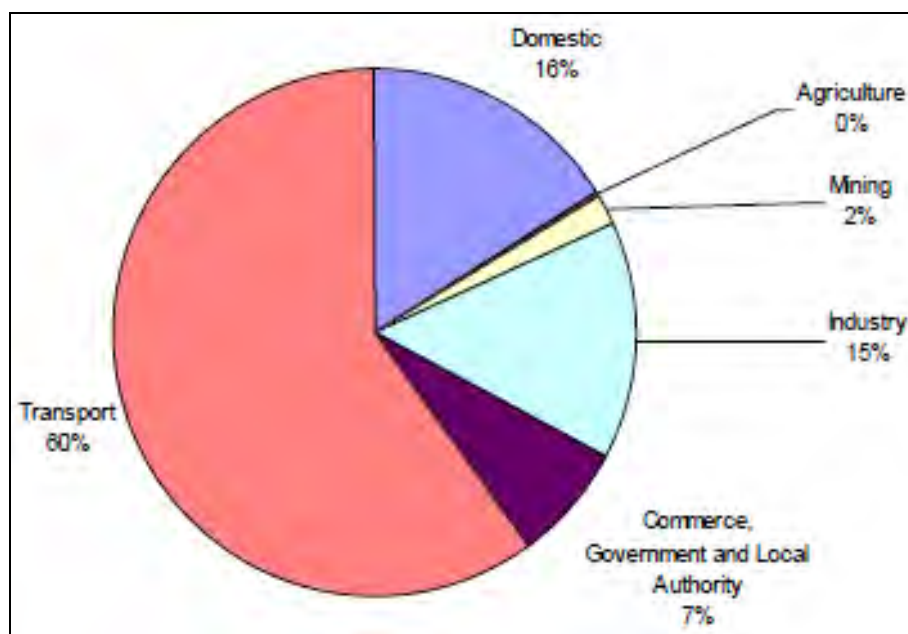


Figure 1: Energy consumption by sector in metropolitan areas around Gauteng Province; Source: Ward and Schäffler, 2008

Ürge-Vorsatz et al (2006) report that the largest savings in gross electricity demand can be delivered by applying electricity efficiency measures such as efficient lighting, ventilation and cooling; in the case of building retrofits accounts for up to a third of these savings. In support of this the Energy Efficiency Strategy suggested by the Republic of South Africa (2008) proposes retrofits as one of the higher-cost interventions measures with short payback periods; this is in addition to a cocktail of measures that are low cost and can be carried out with immediacy. As part of the strategy, the following are proposed (Republic of South Africa, 2008): (1) mandatory standards and audits for energy efficient HVAC, lighting and non-electric appliances and (2) energy management systems in the buildings sector for retrofits. These proposals could be effectively led and implemented by the local authorities through the building regulation control function. In specific, the energy efficiency mandatory audits should be carried out alongside the schedule of others like fire, health and safety, which the local authorities do periodically.

Bennet (2001), du Toit (2006) and Reinink (2007) demonstrated that developing national regulations and standards for building's energy efficiency and/or sustainable buildings is critical in achieving greater energy efficiency and provides a great potential towards reduction in green house gas (GHG) emissions, apart from being a sound investment. South Africa's national building regulations and building standards act (Act no 103 of 1977)¹ is administered by local authorities and is the enabling legislation for the national building regulations (Holden 2004, du Toit 2006). This act details the process of enforcing the national building regulations by local authorities. As a result, all building plans (including major alterations/retrofits) must be approved by the 'building control and approval' departments in the local authorities before

¹ National Building Regulations and Building Standards Act 103 of 1977. Available at <http://www.capetown.gov.za> Accessed on 2009/06/10

implementation. This provides an important opportunity for the local authorities to intervene by insisting on mandatory or voluntary regulations for energy efficiency in their building control and approval process. This concept is further aided by the fact that the local authorities are permitted to make by-laws within their areas of jurisdiction. As such the introduction of energy efficiency by-laws by the local authorities could be used to effect regulations for retrofits in these areas. However, it is important to note that South Africa's national building regulations and building standards act (Act no 103 of 1977) is silent on the idea of retrofits for energy efficiency. This therefore makes it a necessity for the local authorities to enact by-laws which are specific to energy efficiency retrofits so as to ensure controlled implementation and regulation to derive maximum benefits. South Africa is not alone on this, it has been written that former US Vice President, Nobel Laureate and Global warming activist Al Gore once got blocked from installing solar panels on his house because the local authority by-laws did not allow it. It took several months to change the building codes, which were unrelated to solar power, to grant him permission for the installation.

It is encouraging to note that energy efficiency based retrofit regulations are affordable, this is demonstrated by UNEP (2008) and WGBC (2009) who are of the opinion that 'available commercial technologies' make it possible to halve energy consumption in both new and old buildings without significant investments. These could be done at least cost by incorporating measures like improved ventilation and insulation, increased use of natural lighting, the use of energy efficiency appliances and lighting alongside the use of renewable energy sources in the building control and approvals regulations.

There is general acceptance that building energy codes in building regulations as effected by local authorities are an important energy policy instrument (IEA2008, EC 2006). Usually, Building energy regulations impose mandatory minimum energy performance requirements on new buildings and, deemed to include large extensions. This being true, it's therefore relatively a small step to apply the same to refurbishments/renovations/renewals, especially major refurbishments of entire buildings (building research establishment, 2008). This can be through mechanisms like energy labelling and periodic mandatory energy audits by the building control departments of the local authorities. For existing buildings, energy upgrading regulations would demand replacement of components after a certain number of years, (Klinckenburg and sunikka, 2006).

2.0 Methodology

The paper focused on literature review on the use of regulatory and control mechanisms as a policy strategy towards achieving greater energy efficiency in existing building stock with the local authorities' department of building control and approvals process as the enforcing agent. A general scenario on the regulatory framework on retrofits in the world is presented followed by specific situation studies/cases in South Africa. It must be acknowledged that this area is relatively new in South Africa and in many developing countries around the world and as such, much is yet to be published. This paper is therefore limited by scope because of its over-reliance on literature review. It should be pointed out that this forms a preliminary stage of an ongoing research and a field based study is scheduled later this year.

3.0 Energy efficiency regulations in buildings in general and for retrofits/refurbishments/renewals in particular in local authorities

Several instruments or options are available for the promotion of energy efficiency in buildings. Van Egm ond (2001), OECD (2003) , Kuijsters (2004) and Reinink (2007) separately affirm these options as regulatory, fiscal, economic, or communications instruments. These instruments may take the form of direct regulatory instruments, indirect regulation or information instruments (refer Figure 2). It is worth noting that the choice of these instruments by governments is determined by the level of success that is achievable when applied. This is however variant on the socio-economic and political context of the locality.

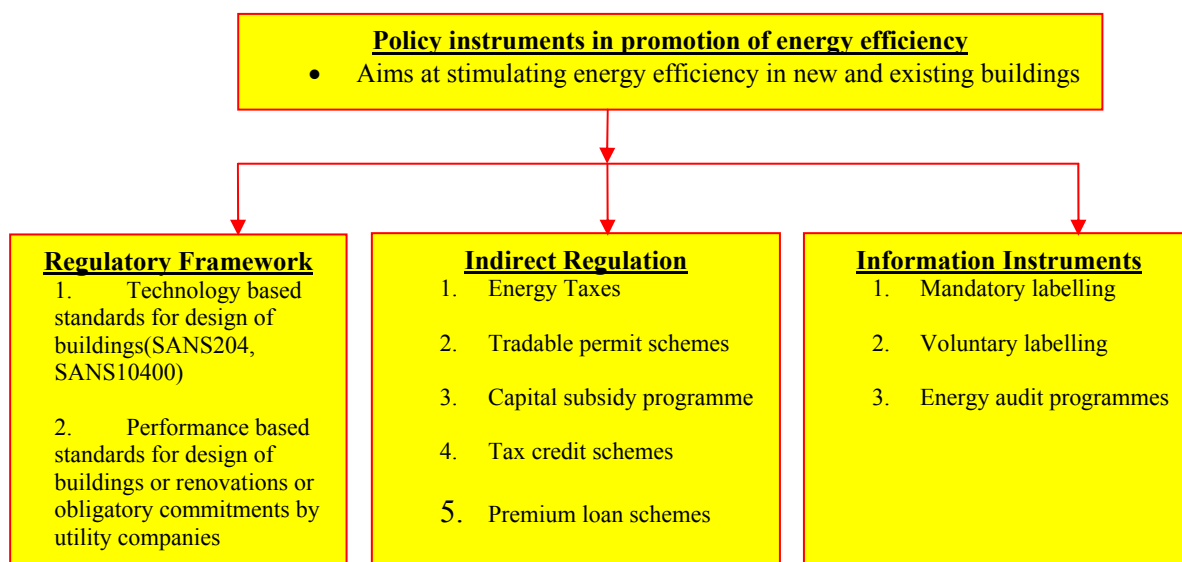


Figure 2 Instruments for energy efficiency implementation in buildings; Adapted from OECD, 2003)

It is noted as an example that implementation of the building codes have reduced energy consumption of dwellings in the U SA by about 30%. The local authorities here have taken leadership in energy efficiency legislations in buildings and through the National Association of Counties (NACo) and the Mayors Conference have achieved 97% use of more efficient lighting in buildings, about 90% require, or anticipate requiring in the next year that new city buildings and large retrofits be energy efficient. The Chinese government have established an energy consumption target in buildings which is 65% less than the current practice in existing buildings, this is via an 'energy consumption standard' for the construction sector whose compliance is encouraged by tax and fees rebate system for low energy buildings to encourage their construction and retrofits (UNEP, 2007). On the other hand energy regulations in buildings in Netherlands have reduced the consumption rates by 15% on its introduction in 1995 and later by 27% on

tightening of requirements (Ecofys, 2004). Other countries where the introduction of energy regulations have been successful are Denmark where the act to promote energy and water savings in buildings (DEA, 1996)² is in use and India where the energy conservation act was enacted in 2001³.

In South Africa, regulations governing energy efficiency in buildings is still at an infancy stage. The specific ones dealing with retrofits/renewals or refurbishments are almost non-existent. This is mainly covered under the electricity regulation act⁴ that stipulates that the regulator shall take into account the energy efficiency measures undertaken by the client while deciding on tariff structure and the Electricity regulations for compulsory norms and standards for reticulation services⁵.

It is expected that these regulations would stimulate increased drive towards energy efficiency in new buildings and existing stock. In addition, it should be noted that the draft standard for energy efficiency in buildings was unveiled in 2008 and if adopted as the national standards would contribute greatly to energy efficiency in buildings. These are SANS 204⁶ Parts 1, 2 and 3, Edition 1.

The newly published building standard is expected to eventually legislate on insulation levels, solar water heaters and energy-efficient lighting and will be prescriptive. While intended for mandatory application in new buildings, SANS 204 can also be used for voluntary energy efficient retrofits of existing buildings, as the owners strive to reduce their electricity and energy accounts. At the same time, the SABS 0400 (the Building Code) is in the process of being rewritten to be SANS 10400 to take into account the energy efficiency standards of existing buildings (Reynolds, 2007, du Toit, 2007).

SANS 204 specifies the general requirements for design and operation of energy efficient buildings with both natural and artificial environmental control and subsystems. One key issue in SANS204 which targets retrofits (Standards South Africa, 2008)⁶, is the emphasis on purpose driven planned maintenance of the mechanical/electrical components; which is advised to be in line with broader economic and energy efficiency agenda.

Enforcement of energy efficiency regulations in buildings by Central or Local government department/agency charged with reviewing designs and performing pre, in-situ and post construction inspections is the most common situation in the world

² Danish Energy Authority (DEA) 1996, Act to promote energy and water savings in buildings, no. 485 of 12 June 1996, legislative document

³ Government of India, 2001, Energy Conservation Act, 2001, September 2001

⁴ Republic of South Africa, 2009, Electricity Regulation Act, 2006 Notice 139 of 2009, Government Printers, Pretoria.

⁵ Republic of South Africa, 2008, Electricity regulations for compulsory norms and standards for reticulation services, Government Printer, Pretoria.

⁶ Standards South Africa (2008) SANS 204-1:2008 Edition 1, Pretoria, South Africa.

(Building Research Establishment, 2008). In South Africa, it is envisaged that the implementation of SANS 204 will fall under the docket of local authorities. It is however noted that it will take considerable time for South Africa to effectively start applying energy efficiency requirements in building regulations in new buildings, besides the retrofits. Holden (2004) approximates this to take a minimum of five years.

At the moment, it is clear that local authorities are ill equipped in implementing and enacting local building regulations due to legal and technical incapacity (du Toit 2007). In fact, the National and Provincial governments are equally handicapped. As a result it is envisaged that the problem be sorted out by rewriting their bylaws with energy efficient regulations such as limiting the quantity of energy consumption in buildings through setting maximum energy/m² caps and making the same to be part of the building applications and approval regulations. This will then subsequently cause all existing stock to be retrofitted to meet this requirement. In the absence of local regulatory powers, cities can consider developing local guidelines or standards. For uniformity and acceptability, these should ideally be based on existing standards and norms and best practices from similar jurisdictions.

The approval process for new buildings or alterations as prescribed by the National Building Regulations and Building Standards Act 103 of 1977 and effected by the local authorities revolves around Administrative matters, protection of property and public health, safety and convenience for the users and occupiers of the buildings. During submission for approvals, key items required are as follows (City of Johannesburg, undated):

- i. Completed application form, signed by the owner of the property or his/her authorised representative (proof of authorisation is required)
- ii. A copy of the registered title deed
- iii. A copy of the approved Site Development Plan is also needed, if this is required in terms of the zoning regulations that apply to the stand.
- iv. A separate form also needs to be completed by a professional engineer or technologist registered with the Engineering Council of South Africa when structural work such as reinforced concrete floor and roof slabs, special reinforced foundations, and so forth are part of the proposed building.
- v. Different certificates or designs need to be submitted depending on the technical aspects of the plan. You should consult an architect or engineer in this regard.
- vi. In the case of non-residential developments (for example, offices, factories, shops, institutional buildings and so on), a zoning certificate with a copy of the most recent Amendment Scheme is to be furnished.

There is no requirement for energy efficiency codes certification to deal with envelope energy issues, HVAC systems, lighting & hot water systems in the building, or for retrofits to be energy efficient in the case of alterations.

A preliminary evaluation of an energy efficient retrofit project in Ekurhuleni Metropolitan Municipality, in greater Johannesburg is presented below.

4.0 Results of energy efficient retrofits: The case of Ekurhuleni Metropolitan Municipality, Greater Johannesburg

In a case to retrofit the buildings at the Germiston Civil Centre at Ekurhuleni Metropolitan Municipality in 2005, 23 zip hydroboils, 2003 CFL lights, 90 LED lights, 15 Geyser Timers and 96 8ft. fluorescent tubes were used (Parra, 2005). Results on the case study undertaken at the at Ekurhuleni Metropolitan Municipality (EMM) indicate that close to 53% of energy savings could be reached in buildings in the greater Johannesburg area (Parra, 2005). (Refer Figure 3 and 4)

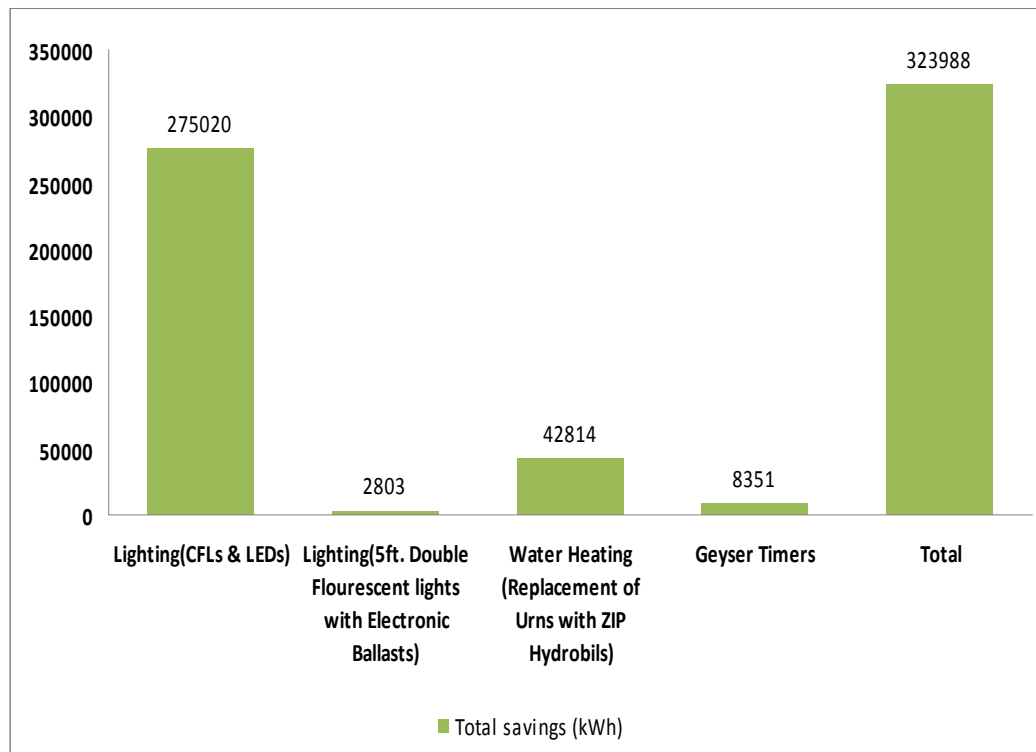


Figure 3: Electricity Savings in kWh for retrofits at Ekurhuleni Metropolitan Municipality

Considering that EMM represents about 5.6% of South Africa's electricity demand savings realized in this case study is reminiscent of the possibilities that could be achieved with properly regulated and controlled energy based retrofits in buildings (Ekurhuleni Metropolitan Municipality, 2004).

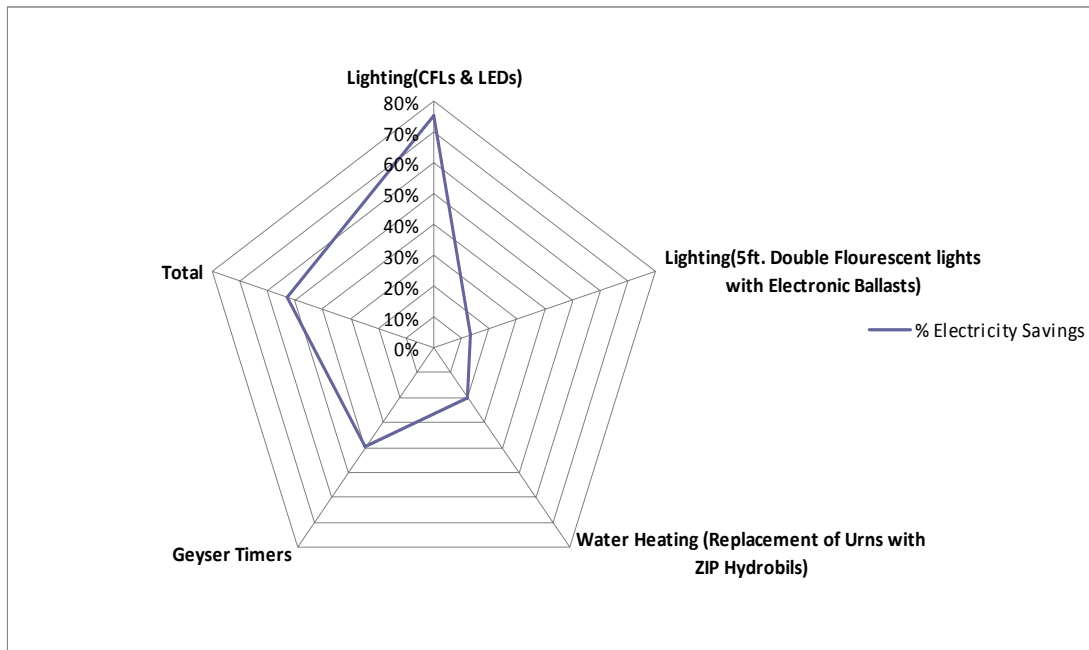


Figure 4: Radar Plot showing % Electricity Savings due to retrofits at Ekurhuleni Metropolitan Municipality. Source; Parra, 2005

5.0 Possibilities towards compliance to Energy Codes

The Building Research Establishment (BRE, 2008) proposes mandatory compliance with minimum energy performance requirements for both new and existing buildings. Thus the buildings are required to be formally certified by private assessors paid for by the building owners. It is further suggested that the process be audited by the authority under which the code is issued (Building Research Establishment, 2008). Thus the following enforcement model is idealized by RICS (refer Figure 4). In South Africa however, the local authorities have a well established buildings control and regulations framework and structure and there is a deliberate intention by government to retain and create employment at local levels. In consideration of this the model proposed by Wafula *et al* (2009) is recommended; under this model the enforcement and the certification be done by these local authorities (refer Figure 5)

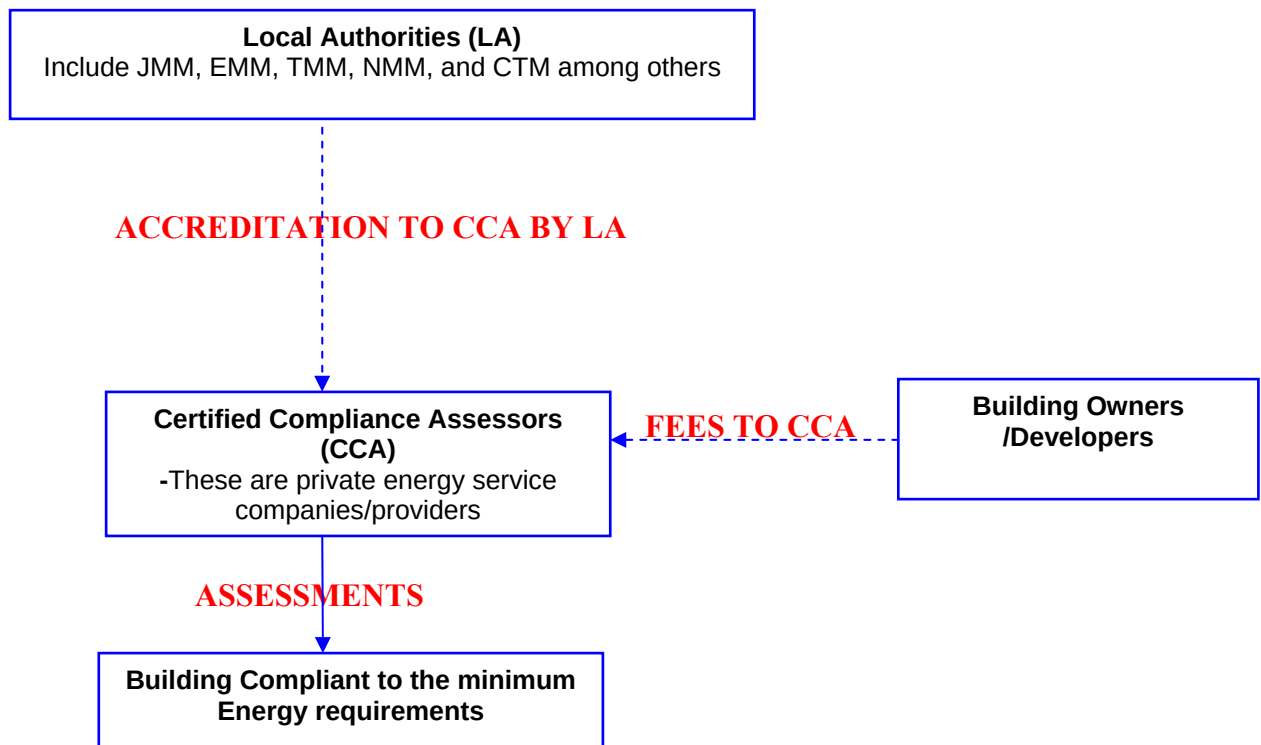


Figure 5: RICS compliance enforcement model, Source: Building Research Establishment, 2008

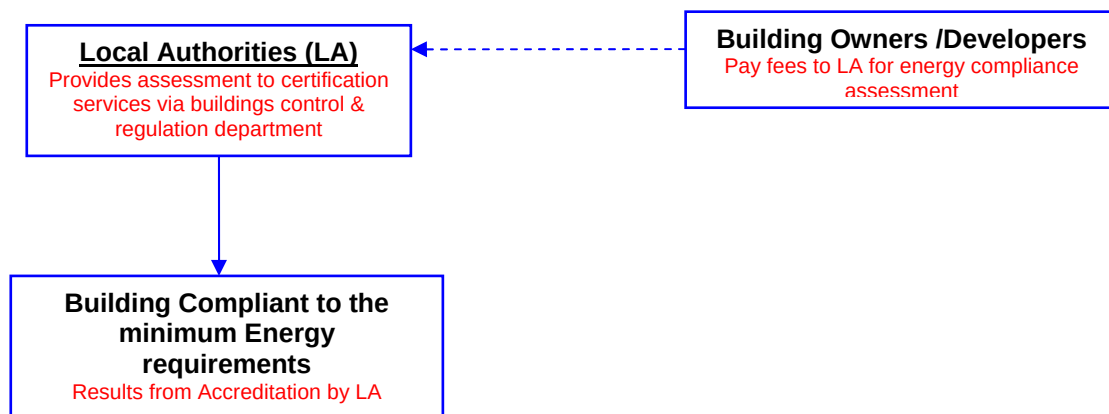


Figure 6: Proposed compliance enforcement for the Local Authorities in South Africa, Wafula *et al*, 2009

As indicated earlier in this paper, local authorities are mandated to enforce building regulations within their areas of jurisdiction, as per the Act (103 of 1977). Some cities have explored this possibility by developing local energy efficiency building regulations, but have since abandoned this process given that building regulations are

established nationally in order to promote uniformity within a vast sector, and the regulations for retrofits are even more difficult to enforce because of the high financial costs involved. However, few building tenants in South Africa are starting to show signs of demanding better energy performance, and some landlords have discovered the financial benefits of improving efficiencies, although landlord / tenant relationships and building management systems in many cases work against change. The Green building council of South Africa (GBCSA) and the South African Property Owners Association have partnered to bring into the market an energy efficiency rating tool especially for commercial new and large retrofit projects. This is supposed to be a requirement as part of the sell and transfer documents for properties among their members. Currently, this is a voluntary mechanism targeting large property holding companies and developers but it is hoped that eventually, when the above start to show monetary savings in their operations, other small scale property owners will scale up their activities and retrofit for energy efficiency. Municipalities are also actively and continuously looking at new by-laws to address these issues, like Johannesburg has energy efficiency guidelines for new buildings. It is hoped that the next step will be the publication of requirements for retrofits and making the same mandatory. It is therefore urgent that property owners/developers smarten up quickly.

6.0 Worldwide Scenarios

Around the world, building controls and regulations with regards to energy efficiency for retrofits are evolving and are being implemented significantly by voluntary standards (refer to Figure 7).

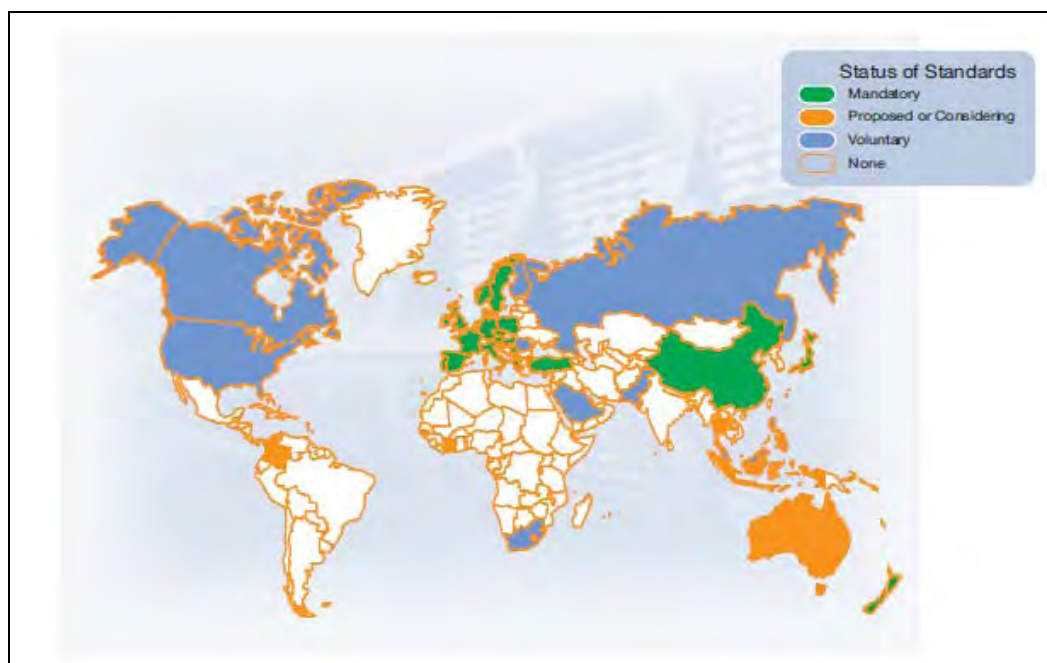


Figure 7: Status of energy building standards around the world; Source: Busch 2000

In the Scandinavian region, there is a general norm of using national building codes and standards, which regulate physical, thermal and electrical requirements of building

components, service systems, indoor conditions, health and safety standards, operation and maintenance procedures and energy calculation methods (UNEP 2007b). A number of building codes currently include energy performance standards, limiting the amount of energy that buildings can consume. The building codes in Scandinavia are enforced by the various building development control agencies in the respective countries.

It is reported that currently there is the renewable energy and energy efficiency partnership (REEP) project aimed at promoting the concepts of low-energy buildings in China, the results will be incorporated in proposed new legislation regarding low-energy buildings by Chinese government in order to ensure that appropriate policies and building codes are implemented to encourage and deliver the required reduction in energy consumption (UNEP, 2007). This targets retrofits as well as new buildings. The implementation of elements of these codes are to be piloted by the four major municipality cities of Beijing, Shanghai, Tianjin and Chongqing and other economically developed big cities such as Shenzhen (Ling et al, 2007).

The European Union is the one area which has a specific regulatory requirement for energy efficiency for retrofits. It often uses directives for regulating various environmental themes whose implementing agents are various government departments in the respective member states. In most cases, this falls in the ambit of local authorities in the respective countries. The directive on the energy performance of buildings was enacted in January 2003 and its main elements are specified as follows (UNEP 2007b):

- i. Minimum energy performance requirements, for new buildings and for major renovation of existing buildings larger than 1000m²;
- ii. Energy performance certificates to be made available when buildings are constructed sold or rented out;
- iii. The year 2010 is the reference year after which the rules will be extended to apply to all buildings and renovations (at the moment buildings below 1000 m² are not covered)

In the United States Green Building council (USGBC) has been central in developing the energy efficiency standards for both new buildings and existing stock. Using a member's consensus system, they have developed a LEED rating tool for commercial construction and major renovations/retrofits. It ranks buildings as Certified, Silver, Gold or Platinum depending on the level of sustainability achieved by construction or retrofit (Sussman, 2008). It is a credible and universally applicable system which has gained wide acceptance.

7.0 Discussions, Recommendations and Conclusions

From the foregoing literature, it is evident that several developing countries have already enacted legislation on energy efficiency in buildings. These regulations and guidelines, including the ones issued by the city of Johannesburg, target new buildings. However, only a few evaluations or studies are available to show the best way to apply this legislation to achieve energy efficiency goals in buildings and especially in retrofits. Due to late entry into energy efficiency practice, most developing countries lack quantitative data and are mainly reliant on data from the developed nations; this poses a

contextual problem taken that their programmes may not be applicable to the prevailing local condition. This can be seen in the case for South Africa which has had to borrow quite heavily from the Nordic countries and Australia in formulating its energy efficiency codes in buildings (Reynolds, 2007).

Developed countries find solace in the fact that increasing energy prices will continue to be the catalytic driver for improved energy efficiency policies in developing countries. In South Africa, the main electricity utility company 'Eskom' successfully lobbied to raise tariffs by nearly 32% and currently is seeking 35% increase per year for the next three years (Engineering News, 2010). As a result it is expected that the South African government will in response empower the local authorities as the electricity distributing agencies to incorporate energy efficiency measures in their building control processes and bylaws, besides engaging in training of the built environment practitioners and advocacy to the general citizenry.

The end of an Era of cheap energy in South Africa, and the need to create more construction work in the building industry following the end of world cup associated works, makes it natural that retrofits are going to be an important driver of the construction work in the economy. It is therefore strategic and necessary for South Africa to capitalise on this and use the local authorities as the enforcing agents for energy efficiency codes in buildings. This is made easier by the fact that in South Africa, the local authorities are considered the most important party in service delivery (Republic of South Africa, 1995). It is noted that in South Africa, energy efficiency policies are in various government agencies and departments (these are Department of Energy, Department of Minerals, ESKOM, NERCSA, Department of Public Works, Roads and Transport), the policy implementation details is therefore a challenge. For the building sector, it is recommended that the local authorities should take the initiative and follow the lead provided by the city of Potchefstroom and effect energy efficiency regulations for retrofits through their building control and approvals regulations processes and systems.

The most direct and comprehensive way to achieve energy efficiency in buildings in general and existing building stock in particular is through the change of building regulations. Local authorities find themselves in positions of developing and implementing the building regulations in areas under their jurisdiction. Through this authority, they can get direct and assured energy efficiency benefits in retrofits by enacting and implementing high yield specific design regulations which require no additional cost. This may include a planning tool like addition of vegetation requirement in an already built up area. This will lower temperatures during summer and reduce heating costs. As new technologies develop, many other specific cost effective design elements will present opportunities for targeted mandates in retrofits. Changing from conventional hot water heating system to solar water heaters in domestic buildings is another key regulation which should be implemented urgently in retrofits. Heating water can take up to 15-25% of the energy use in a home and a solar water heater can cut annual operating costs by up to 50-80%. In 2000, Barcelona implemented such a requirement and was subsequently followed by other cities all over Spain successfully.

In South African cities, several former Apartheid Era 'Hostels' and many other former office blocks are being converted into residential use. This gives local authorities an opportunity to require mandatory incorporation of solar water heaters with every retrofit application submitted for approval. This can then be rolled out to all other residential applications retrofits made to the respective local authorities.

It should be emphasized that the success of the Building Regulations is highly dependent on compliance by the construction industry hence the need for Local Authorities to invest in continuous training and advocacy of Building Control Officers and Built Environment Professionals on the energy efficiency regulations and their application. Getting voluntary action is slow and sometimes difficult to measure results. Local authorities should make it mandatory for the professional practitioners submitting retrofit/refurbishment plans for developers to show an energy efficiency commitment statement in their briefs as a precondition for approval for redevelopment. Similarly, such mandates should also be required upon sale of commercial and residential real estate before a transfer can be effected. This provides a unique attractive opportunity for enhancing energy efficiency in both new buildings and retrofits.

For the case of the existing stock owned by the municipalities and the local authorities themselves, operational practicality requires that they examine them first for energy efficiency and retrofit them appropriately as pilot/case projects. There are many obvious resources which can be explored in detail to deal with this. One stand out practice is the employment of Energy Service Company (ESCO). The ESCO develop, design and finance the energy efficiency project/program on behalf of the local authority on its premises. They install and maintain the energy efficiency equipment involved, measure, & monitor and verify the program's energy savings. They also assume the financial risk that the program will deliver the amount of energy savings guaranteed. This is appropriate for the local authorities since they do not have upfront costs and the ESCO's get paid from the savings generated by the program. This allows the local authorities to deliver their normal operations without additional financial burden for energy efficiency programs on their building stock. Significantly, results of these programs can subsequently be rolled out to all other large building stock holders in their jurisdiction. The challenge to the local authorities in this arrangement lies in establishing appropriate procurement legislation which protects its interests at all times. For South African local authorities, this is an attractive and fiscally sound means of financing an energy efficiency upgrade/retrofit in local authority buildings.

Energy efficient retrofits/renewals and refurbishments do involve a high financial outlay. Getting appropriate finance for it is therefore important to its success. In South Africa, these interventions in refurbishments could be viewed as reducing demand. This should then be eligible for funding from sources like Eskom DSM Fund. Appropriate partnerships should be created between the local authorities building regulations control entities and the above funds administration systems to make it mandatory for the recommendations of the local authorities to be part of the criterion used for developers to draw from it. However, the South African local authorities can develop additional revenue streams to support energy efficiency new and renewal/retrofit programs in their jurisdictions and complement the Eskom DSM programs. In a scheme that can serve both to raise funds and promote installation of renewable energy, the local authorities

can introduce a fee, of say R5000 for all new developments and large retrofits of over 2000 square feet if they fail to include the installation of a two kilowatt solar photovoltaic system or equivalent renewable energy system. Similarly, homeowners who consume energy beyond reasonably forecasted budgets, especially on energy intensive activities such as heating outdoor pools or spas can be charged a mitigation fee of say R5000 if they fail to install energy efficiency or renewable energy systems. The funds collected can be used to promote energy efficiency and renewable energy policies in buildings. Also, the South African policymakers should look at the Property Assessed Clean Energy (PACE) financing mechanism as a tool in overcoming barriers to commercialised building retrofits (Pike Research, 2010). This way, the goal of energy efficient retrofits will have made a giant step forward.

In terms of the mode of policy implementation, South Africa can borrow a leaf from the EU and set a minimum space (in area terms) for retrofits which must be certified by the local authorities. It should also be noted that although best practices and experiences can be shared and regional cooperation is useful, building code specifications especially for retrofits cannot be uniform for all parts of a country due to climatic, financial and other peculiar differences. This further fosters the idea of local authorities as the ideal enactor and enforcer for energy efficiency codes for retrofits in buildings.

An appropriate Education and Outreach programme to raise awareness is key to implementing successful energy efficiency retrofits regulations through the building regulations and control mechanism, in the local authorities. This should target developers on the economic/financial benefits of energy efficient buildings, and demonstrate the technical options for retrofitting of existing buildings to the technical professionals in the industry.

Research gaps exist in several areas of energy efficiency in buildings in South Africa, an example is the lack of cost estimation data for energy-efficient retrofits vis-à-vis ordinary business as usual retrofits. Financial specialists in organisations do not look at energy efficient retrofits as Investments, but rather as facilities Expenses measured on straight line depreciation scale just like building consumables; hence there is no efficient market guidance for energy-efficient retrofits in buildings. This causes low sensitivity to energy-efficient retrofits issues in the maintenance of buildings. Research and studies in these areas will enhance learning and help make improvements in program designs for energy-efficiency retrofits in buildings in future.

South African Local authorities play a significant role in the energy efficiency agenda. This is because much of the implementation is done at the local level. They can move forward and enact and implement proactive and innovative building regulations which address the challenges of energy efficiency in existing building stock.

8.0 Acknowledgement

The authors acknowledge financial support for the study from the South African National Energy Research institute (SANERI).

9.0 References

Bennett K.F. (2001) Energy efficiency in Africa for sustainable development: A South African perspective, UNESCO Workshop on sustainable development, Nairobi

Building Research Establishment (2008) Can building codes deliver energy efficiency? Defining a best practice approach. A report for the Royal Institution of Chartered Surveyors, RICS, London, United Kingdom.

City of Johannesburg, Gange and CSIR (2008), Design guidelines for energy efficient buildings in Johannesburg. Available at <http://www.joburg.org.za> on 25/01/2010.

City of Johannesburg(Undated) Development planning-Building control. Available at <http://www.joburg.org.za/content/view/553/9/>

Department of Energy and Minerals (DME) (1998) SAEDES guideline: Energy and demand efficiency standard for existing and new commercial buildings, Pretoria.

Derringer, J. Iyer, M. and Yu Jo Huang, Y. J. (2004) Transferred just on paper? Why Doesn't the Reality of Transferring /Adapting Energy efficiency codes and Standards Come Close to the Potential? Proc. 2000 ACEE Summer study on Energy Efficiency in Buildings, Pacific Grove, CA, August 2004.

du Toit E. (2007) How to Implement Renewable Energy and Energy Efficiency Options. Sustainable energy Africa, Westlake.

Ekurhuleni Metropolitan Municipality (2004) State of Energy Report, November 2004. Available at www.cityenergy.org.za/files/resources/soe/EMMSOE.pdf

Holden R.M. (2004) Towards Energy Efficient Building Regulations – The City of Johannesburg's Building Bylaws Review Process, City of Johannesburg, Johannesburg.

Eberhard A. (2004) Electricity reform in South Africa is at a critical juncture, business day South Africa, 29 July 2004. Available from: <http://cesp.stanford.edu/news/321/> on 2009/05/20

Ecofys (2004) Evaluatie van het klimaatbeleid in the gebouwde omgeving 1995-2002, prepared for department of housing, spatial planning and environment. Available from: www.ecofys.nl/nl/publicaties/documents/samenvattingrapport_20040623.pdf on 2009/04/22

Engineering News, Nersa publishes municipal electricity tariff increase details, 8th July 2008. Available at <http://www.engineeringnews.co.za/article>. Accessed on 31 July 2009

The European commission Action plan for energy, energy efficiency COM(2006)545 at ec.europa.eu/energy/action_plan_energy_efficiency/.../com_2006_0545 on 15/01/2010

The International Energy Agency (2008) Policy recommendations on building energy codes, Vienna.

Klinckenburg F. and Sunikka M. (2006) Better Buildings through Energy Efficiency, Eurima, 2006.

Kuijsters A. (2004) 'Environmental response of the Chilean building sector: Efforts and Constraints towards environmental building practices in the Santiago Metropolitan Region', MSc Thesis, Eindhoven University of Technology. Available from : <http://www.tue.nl/bib> on 2009/05/08.

Liang Jing, Li Baizhan, Wu Yong, Yao Runming (2007) An investigation of the existing situation and trends in building energy efficiency management in China. Energy and buildings, Elsevier Science, Beijing.

Maisely G. E. and Beverly M. (2007) Energy master planning for HVAC systems in new and existing buildings, National Institute of Building Sciences, Available at <http://www.wbdg.org>.

Organisation for Economic Co-operation and Development (OECD) (2003) Environmentally sustainable buildings: Challenges and policies, Paris, OECD. Available from: http://www.oecd.org/publications/pol_brief on 2009/04/15

Organisation for Economic Co-operation and Development (OECD) (1989) Economic Instruments for environmental protection. Paris, OECD.

Parsons S.A. (2004) International building energy standards (codes): Their relevance for the proposed SA national energy efficiency standard SANS 0204, unpublished document, CSIR, Pretoria.

Parra M.M.(2005)Improving Energy Efficiency in Ekurhuleni Metropolitan Municipal Buildings, Report for I.C.L.E.I. Cities for Climate Protection Campaign, Available at www.iclei.org/ccp

Reeves L. (2008) The changing world of maintenance and energy in pre-loved buildings, EcoLibrium, April 2008. Available at <http://www.airah.org.au/eco>.

Reinink M.W. (2007) Towards an effective energy labelling programme for commercial buildings: A Comparative evaluation of the green buildings for Africa programme in relation to international experience', MSc Thesis, University of the Witwatersrand.

Republic of South Africa (1995) White paper on transformation of the public service, Notice No. 1954 OF 1994, Government Printer

Reynolds, L K (2007) The South African Energy Efficiency Standards For Buildings, at <http://www.cpcut.ac.za> on 2009/05/20

UNEP (2008) The Kyoto Protocol, The CDM, and the Building and Construction Sector, Nairobi, Kenya.

Sussman E. (2008) Reshaping Municipal and County Laws To foster Green Building, Energy Efficiency, and renewable. N.Y.U Environmental Law Journal. Vol. 16. New York, USA.

UNEP (2007) Assessment of Policy Instruments For Reducing Greenhouse Gas Emissions From Buildings, Central European University, Budapest.

United Nations Environment Programme (2007b) Buildings and climate change; status, challenges and opportunities. Nairobi.

Ürge-Vorsatz D., Koeppel S. and Novikova A. (2006) Improving Energy-efficiency in Buildings: opportunities, promises, challenges and solutions . Conference proceedings for " Using energy and materials more efficiently: a precondition for sustainable development", Seoul, September 21-22, 2006

Van Egm ond- de Wilde de Ligny, ELC (2001) Technology policies in developing Countries. Lecture Notes course ON470, Eindhoven University of Technology, Eindhoven.

Wafula J., Aduda K. and Talukhaba A.A.(2009) Energy efficiency in buildings and building regulations and control in local authorities in South Africa, COBRA 2009 -The Construction and Building Research Conference of the Royal Institution of Chartered Surveyors, University of Cape Town, South Africa, 10-11 September 2009. Available at www.rics.org/site/download_feed.aspx?fileID=5060...PDF

Ward S. and Schäffler J. (2008) Gauteng 2034 development strategy: Energy trend paper Gauteng Department of Economic Development. Available at [www.nano.co.za/Gauteng2034EnergyTrends28Oct08%20v5%20\(2\).pdf](http://www.nano.co.za/Gauteng2034EnergyTrends28Oct08%20v5%20(2).pdf)

World Green Building Council(2009) Boosting the green house effect – CDM Reform Key to Climate-Friendly Building and construction sector, at <http://worldgbc.org/home/107> on 2009/04/01

The Quest for Sustainable Homes and Reduced Emissions from Buildings: Getting it Right at the Planning Stage

Julie Adshead¹

¹Salford Law School,
University of Salford, Salford, M5 4WT,
United Kingdom

Email: J.D.Adshead@salford.ac.uk

Abstract:

Some 40% of carbon dioxide emissions in the United Kingdom (UK) derive from the construction and occupation of buildings and around 40% of energy useage is accounted for by the built environment. The law and policy surrounding emissions from buildings is multi-layered and complex. Taking into account the global and regional context, this paper focuses upon the legislative measures in place in the UK to promote improved energy performance and to reduce carbon emissions from the built environment. In particular it focuses upon the role of the planning regime in this regard. The paper examines elements of planning law and practice that can serve to improve standards of sustainability in buildings at the planning stage. Planning conditions are one such route and two recent case studies are considered in this paper in order to illustrate how such conditions requiring higher standards than current legislation can operate in practice. Some conclusions are then drawn as to how best the planning regime can serve to help the goal of reduced emissions and the direction that this may take following the recent change in government.

Keywords:

Buildings, Carbon Dioxide, Emissions, Legislation, Planning.

1 Introduction

Climate change and security of energy supply are key drivers of policy and legislation in current times. At the same time, government is also focusing upon those households subject to ‘fuel poverty’¹. There is an impressive number of initiatives in place to secure reduction in green house gas emissions and to promote energy efficiency. These measures range from legally binding international agreements to local voluntary community schemes. The complexity of the policy and legal frameworks is accentuated by the fact that there are multiple goals to achieve and this results in some incidences of paradox. Take micro generation, for example. Certainly domestic schemes will result in reduced carbon dioxide emissions and should provide a reliable source of energy for the future. As evidenced by the Climate Change and Sustainable Energy Act 2006, the government also sees micro generation as one of the hopes for reducing fuel poverty.

¹ Following the Warm Homes and Energy Conservation Act 2000, ‘fuel poverty’ is generally accepted as a consumer spend of more than 10% per week of income on energy.

Micro generation does, however, tend to be expensive (certainly in comparison with current energy prices) and if it is to be successful in the elevation of ‘fuel poverty’ then substantial financial support will be needed. Ultimately, whether the economics of micro generation make sense will depend upon whether the era of cheap centralised energy is really at an end (Dow, 2007). The array of measures in place to improve energy performance and thus reduce carbon emissions, will also serve to ease the burden of energy expenses on poor households. However, their role in addressing the goal of reducing carbon emissions is based on the premise that a large proportion of energy provision is from fossil fuel sources. A switch to a mix of nuclear and renewable sources (although this may raise other entirely different issues in relation to the environment and sustainability) would arguably be far more efficient in reducing carbon emissions from buildings.

This paper focuses upon the legal provisions in place to reduce carbon emissions from buildings. In particular, it reviews the role and the potential of the UK planning regime to this end. The first section sets the context in terms of global and regional commitments to counter climate change. Some of the legislative provisions in the UK are then explored. This includes overarching measures such as the Climate Change Act, 2008 as well as specific legislation that targets the energy performance of buildings. This second section of the paper also gives some brief attention to the non-legally binding ‘Code for Sustainable Buildings’ as this is integral to the way in which planning law can operate to improve standards. The third and final part of the paper considers the role of the UK planning regime, including the development and future of the ‘Merton Rule’, incentives for micro generation through the ‘Permitted Development Order’ route and the use of model planning conditions. In order to illustrate how the use of planning conditions can succeed (or not), two case studies are considered. In one of the case studies, a planning condition was upheld whereas the relevant condition in the other case study was subject to a successful appeal. Some tentative conclusions are then drawn as to the future direction of law and policy in the UK relating to emissions from buildings in the context of the coalition government’s ‘localism’ agenda.

2 The International and Regional Context

2.1 The International Regime

The international climate change regime comprises the United Nations Framework Convention on Climate Change² and its Kyoto Protocol³. Both of these instruments are in force and are legally binding. The Kyoto Protocol commits many of the industrial nations to a reduction in the annual average of green house gas emissions. In the ‘first commitment period’ from 2008 – 2012 the reduction is to an average level of 95% of 1990 emissions. Fundamental problems have been identified with the Kyoto Protocol. Not least of these is the fact that the United States (US), responsible for 20% of the overall output of green house gases, is not a party (Human Development Report, 2007/2008). It has also been suggested that the commitments made so far are inadequate and have not been successfully implemented (Barker *et al*, 2007, den Elzen,

² United Nations Framework Convention on Climate Change (adopted 29 May 1992, entered into force 21 March 1994) 1771 UNTS 107.

³ Kyoto Protocol to the United Nations Framework Convention on Climate Change (adopted 10 December 1997, entered into force 16 February 2005) 37 ILM 22.

2008). With the first commitment period drawing to a close in 2012, it was hoped that, following the Bali Action Plan, adopted by the international community in December, 2007, an 'agreed outcome' on long term cooperative action on climate change would be reached in Copenhagen in December 2009. Despite much debate as to the possible legal form of the 'agreed outcome' (Rajamani, 2009) and high hopes of a legally binding agreement (Thomas & Woodward, 2010) there was no such agreement. The result of the Copenhagen Conference is an accord, led by the US, between China, India, Brazil, South Africa and the US to tackle global warming and deliver aid to developing nations. Despite criticisms of the Kyoto Protocol and the failure to reach any kind of multi-lateral agreement at Copenhagen, there is no doubt that the Kyoto goals have been a powerful driver for governments. The Copenhagen accord does provide for nations to commit to implement emissions targets for 2020 and a number of world leaders have signified their intention to introduce yet further more stringent, legally binding targets (Thomas & Woodward, 2010).

2.2 Law and Policy of the European Union

The European Union (EU) has the quota of a reduction to 92% of 1990 levels by 2012 under the Kyoto agreement. However, the EU is committed to even more stringent targets than those provided for under the international regime. The EU objective is to reduce overall greenhouse gas emissions by at least 20% below 1990 levels by 2020 and by 30% in the event of an international agreement being reached. The Union has also set a binding target for energy from renewable sources of 20% of total EU energy consumption by 2020. A range of measures exist to achieve these goals including a directive on energy end-use efficiency and energy services⁴, with an overall objective of saving 9% of energy by 2012 and a directive on the promotion of the use of energy from renewable sources⁵ that provides for the improvement of energy efficiency in the context of the binding target for energy from renewable sources. The key legislative instrument applying to the control of emissions from buildings is the directive on the energy performance of buildings (2002/91)⁶.

Under directive 2002/91, member states are required to establish a methodology for determining the energy performance of buildings and set minimum energy performance standards. New buildings above 1000 square metres are to meet these standards as are buildings above this limit that undergo major renovation. On construction, sale or rent, an energy performance certificate for the building must be made available and this is to be no more than 10 years old. Public buildings exceeding the 1000 square metres limit are required to display their certificates. The directive also provides for inspection regimes for boilers and air conditioning systems and for inspections to be carried out by independent experts.

Over recent years there have been calls from the European Council and the European Parliament⁷ for the Commission's priorities, established in its *Action Plan for Energy*

⁴ Dir 2006/32 on energy end-use efficiency and energy services OJL 114 27.4.2006 p64-85.

⁵ Dir 2009/28 on the promotion of the use of energy from renewable sources OJL 140 5.6.2009 p16-62.

⁶ Dir 2002/91 on the energy performance of buildings OJL 1 4.1.2003 p65-71.

⁷ European Council of March, 2007, European Parliament Resolutions of 31 January, 2008 and 3 February 2009.

Efficiency: Realising the Potential published in 2006⁸ to be comprehensively and swiftly implemented. The action plan identified the significant potential for cost-effective energy savings in the buildings sector and, as part of the package of measures to achieve the priorities of the action plan, a new directive on the energy performance of buildings was published in May 2010. When in force (2012), the directive will expand upon the provisions of the 2002 directive significantly.

Under Article 4 of the new directive, member states will have to set minimum energy performance requirements for buildings or building units ‘with a view to achieving cost-optimal levels’. This requirement will also apply to ‘building elements that form part of the building envelope and that have a significant impact on the energy performance of the building envelope when they are replaced or retrofitted’. The requirement to meet minimum energy performance requirements will apply to all new and existing buildings, regardless of size⁹. Member states will also be required to set energy performance for technical building systems installed in existing buildings¹⁰. The directive includes a new binding requirement upon member states to ensure that by 2020 all new buildings are nearly zero-energy¹¹. The display requirements for public buildings are extended to include those ‘frequently visited by the public’ and the threshold size is lowered to 500 square metres (to be reduced further to 250 square metres in 2015)¹². In addition, inspection requirements are extended to apply to all elements of heating systems (not just boilers)¹³.

3 UK Legislative Provisions

3.1 The Climate Change Act 2008

The UK is also committed to more ambitious targets than those set in the Kyoto Protocol. In fact, it is the first nation worldwide to adopt a legally binding long term framework to cut carbon emissions. The controversial Climate Change Act, 2008 imposes a statutory duty upon the Secretary of State of 2050 ‘to ensure that the net UK carbon account for the year 2050 is at least 80 per cent lower than the 1990 baseline’¹⁴. The imposition of statutory duties on government is a novel approach in UK law (Stallworthy, 2010) and some might doubt the meaningfulness, in particular, of imposing a legal duty on an individual whose identity is, as yet, unknown. It is also difficult to see how legally binding this target can be, when it is unlikely to be legally enforceable (Townsend, 2009; Stallworthy, 2010). However, proponents of the Act (Grekos, 2009; Townsend, 2009) recognise its possibilities in terms of improving carbon management, moving the UK towards a low carbon economy and providing strong leadership and commitment to shouldering an equitable burden in reducing global emissions. It may also provide some certainty and encouragement for industry and business.

⁸ Communication from the Commission, *Action Plan for Energy Efficiency: Realising the Potential*, Brussels, 19.10.2006 COM (2006) 545 final.

⁹ Articles 6 and 7.

¹⁰ Article 8.

¹¹ Article 9.

¹² Article 13.

¹³ Article 14.

¹⁴ Section 1.

The Act requires that a series of five-yearly carbon budgets are set by order of the Secretary of State. The first three budgets (2008-22) have already been set by the Carbon Budgets Order¹⁵ with a view to meeting the 2050 target. Thus 2018 – 2022 is 34% lower than the 1990 baseline. There is a more ambitious figure of 42% by 2020, which will only be adopted if a global agreement is reached. The crediting of carbon units is going to be crucial if these objectives are to be met. This is clearly illustrated by the scale of net reduction from 2007 to 2008, which was a mere 2% (DECC, 2009). The Carbon Accounting Regulations, which define carbon units and set out how carbon can be credited to the account, came into force on 31st May, 2009¹⁶. Greenhouse gas allowances (under trading schemes) can also act as credits, but only credits under the EU emissions trading scheme can be credited to the UK carbon account.

The Secretary of State is subject to a duty ‘to ensure that the net UK carbon account for a budgetary period does not exceed the carbon budget’¹⁷. Some indication as to how this will be achieved is provided in measures recommended by the Climate Change Committee set up under the Act to advise the Secretary of State¹⁸. The key short term recommendations of the Committee are:

- Energy efficiency improvements in building and industry
- Fuel efficiency improvement in road vehicles
- A significant shift towards renewable and nuclear power generation and renewable heat

In order to achieve the extremely ambitious targets provided for in the 2008 Act, a range of policy and legislative initiatives have been put in place. The two primary mechanisms for the delivery of energy efficiency improvements in buildings are building regulations and the planning regime and these are considered below.

3.2 Powers under the Building Act 1984

The Building Act 1984 places certain aspects of building under statutory control and empowers the Secretary of State to make regulations that provide details of exactly how that control is exercised¹⁹. The scope of the Building Act in terms of regulating the conservation of fuel and power was significantly widened by two recent pieces of legislation: The Sustainable and Secure Buildings Act 2004 and the Climate Change and Sustainable Energy Act 2006. The Sustainable and Secure Buildings Act, 2004 enables the making of regulations for ‘furthering the conservation of fuel and power’²⁰ and extends the range of matters in respect of which regulations can be made²¹. Significantly it allows for the regulation of existing buildings in matters relating to

¹⁵ SI 2009, No. 1259.

¹⁶ SI 2009 No. 1257.

¹⁷ Climate Change Act 2008, Section 4 (1).

¹⁸ *Building a low-carbon economy - the UK's contribution to tackling climate change* published on 1st December 2008.

¹⁹ Building Act 1984, Sections 1, 2, 2A and Schedule 1.

²⁰ Sustainable and Secure Buildings Act 2004, Section 1 (1) (b).

²¹ Building Act 1984, Section 1A as inserted by Sustainable and Secure Buildings Act 2004, Section 1 (3).

energy conservation and carbon emissions²². It also inserts a new Section 2A into the Building Act 1984, which allows for regulations to be made that impose ‘continuing requirements’ on building owners and occupiers regardless of when the building was erected or whether other building works are ongoing. These powers are potentially far-reaching and would permit, for example, the making of regulations requiring all lofts to be insulated (McAdam, 2007). The Climate Change and Sustainable Energy Act 2006 further extended the powers under the Building Act 1984 by enabling regulations to be made under the Act relating to the installation of micro generation technologies in buildings²³. The Act also extends the time limit for prosecution of those in breach of regulations specifically relating to the conservation of fuel and power²⁴.

The Building Act 1984 allows for guidance documents to be approved and compliance with approved guidance creates a presumption that the works in question comply with the requirements of the Act as provided for in The Building Regulations. Part L guidance documents deal with the conservation of fuel and power²⁵ and these cover new dwellings, existing dwellings and new and existing buildings other than dwellings. Under The Building Regulations, target emission rates for buildings have to be set for new dwellings and buildings over 1000 square metres must be brought up to Part L standard when renovated²⁶. For new dwellings the government has committed to a programme by which regulations will demand 25% lower carbon emissions by 2010, 44% lower by 2013 and by 2016 all new dwellings should be zero carbon²⁷. Thus, the Part L standard in building regulations will be incrementally raised over forthcoming years. New Part L guidance implementing the 25% lower emissions requirement comes into force in October of this year. For existing housing stock, it is arguable that the full potential of the Building Act and regulations made under it have not as yet been fully realised. The government has turned its attention to the problem of tackling emissions from existing housing and a House of Commons Report was published on the subject in 2008²⁸. The Report sets out recommendations for improving energy efficiency in existing housing with a ‘shopping list’ of recommended measures for government (Grekos, 2008). These include measures to encourage the take-up of home micro generation, requirements for consequential energy efficiency improvements in planning consent on extension of homes, new requirements for Energy Performance Certificates and the production of a ‘Code for Existing Homes’ along the lines of the ‘Code for Sustainable Homes’ (see below).

4 The Code for Sustainable Homes

The Sustainable Buildings Task Group first proposed a Code for Sustainable Buildings in 2004. The idea was that the voluntary code would be a catalyst for low carbon, low impact building and set vanguard eco-standards for the government to follow. The Code for Sustainable Homes was finally launched in December 2006 and the technical

²² Sustainable and Secure Buildings Act 2004, Section 3 (7) (5).

²³ Climate Change and Sustainable Energy Act 2006, Section 11.

²⁴ Ibid., Section 13.

²⁵ Building Regulations 2000, Schedule 1.

²⁶ Implementing Directive 2002/91 (Number 6, above).

²⁷ Meaning that, during the course of a year, the net carbon emissions from all energy use in the building is zero.

²⁸ *Existing Housing and Climate Change*, House of Commons Communities and Local Government Committee, Seventh Report of Session 2007/8, 2, April 2008.

guide followed in April 2007. The code sets six standards of increasing rigour against which the whole home can be measured. A whole range of factors are considered alongside carbon dioxide emissions. There are nine design categories and the levels are rated from one to six stars. Level 6 (six star) of the code, in terms of emissions, is 'true zero carbon'. The lowest one star level is currently more demanding than minimum standards for building regulations. However, this will change once the new Part L standards come into force in October 2010. The Part L standard for emissions from buildings will then be equivalent to level 3 (three star) of the code.

At its inception, the code was a voluntary mechanism, but since May 2008, sellers of new properties have been required to provide information to the purchaser on the rating of the building, either in the form of a code certificate or a statement of non-assessment. Also from 2008, achievement of Code level 3 became mandatory for all publicly supported developments. The first proposals for a code were intended to embrace both new homes and non-domestic buildings and calls continue for the code to be expanded in this respect (UKGBC, 2009) as well as for a similar code to be adopted for existing homes (see above). Although, essentially still a voluntary system, as shall be demonstrated below, when used in tandem with the UK planning system it has the potential to drive up standards across the entirety of the nine design categories, and thus serve to increase the chances of reaching the overarching emission targets set by government in the Climate Change Act.

5 The UK Planning Regime

A full account of the UK planning regime is beyond the remit of this paper, but an overview of certain elements of UK planning law is helpful in understanding how the planning system may operate to drive up emissions performance.

5.1 Development

The UK system is centred upon 'development'. The key statute, the Town and Country Planning Act 1990, states that 'Planning permission is required for the carrying out of any development of land'²⁹. The definition of 'development' in the Act is extremely broad and encompasses both building operations and the change of use of buildings.

'Development..... means the carrying out of building, engineering, mining or other operations in, on, over or under land, or the making of any material change in the use of any buildings or other land'³⁰.

5.2 The Development Plan

UK planning is led by the development plan and reference to the plan will be the starting point in determining a planning application. Development plans are broad, giving general policy, aims, objectives and goals are generally permissive in nature. The Planning and Compulsory Purchase Act 2004 introduced a new range of strategies and plans. The 'Regional Spatial Strategy' introduced under the 2004 Act has since been suspended by the current government, but the local development framework

²⁹ Town and Country Planning Act, 1990, Section 57 (1).

³⁰ Ibid., Section 55 (1).

(including the development plan) continues to operate. The Planning and Compulsory Purchase Act states that

‘regard is to be had to the development plan, the determination shall be made in accordance with the development plan unless material considerations indicate otherwise’³¹.

5.3 Material Considerations

Although the development plan provides the starting point, it will not necessarily be the dominant determinant in the decision³². Material considerations can be taken into account and on occasion they can win out over the development plan³³. There is no statutory definition or guidance as to what constitutes a material consideration. Certainly planning guidance, representations, as well as many other matters determined by case law, will be material considerations. Environmental considerations are just one of many possible elements to be brought into the balancing act in the determination of a planning application.

5.4 Planning Guidance

Planning guidance plays a pivotal role in the UK planning system. There are a range of guidance documents in place, which, as noted above, will constitute material considerations in the determination of a planning application. The statutory requirement in the Planning and Compulsory Purchase Act 2004 for all plan making bodies to exercise their functions ‘with the objective of contributing to the achievement of sustainable development’ is reflected in the government’s Planning Policy Statement (PPS)1: Delivering Sustainable Development and its supplement, Planning Policy Statement: Planning and Climate Change as well as Planning Policy Statement 22: Renewable Energy. A new draft climate change planning policy statement has recently been released³⁴, which combines the policies currently set out in the supplement to PPS1 and PPS22.

5.5 Planning Conditions

Almost all planning determinations include conditions. The Town and Country Planning Act, 1990 allows the local planning authority to impose such conditions ‘as it thinks fit’³⁵. Guidance is given in the Act³⁶ and in the Secretary of State’s policy on conditions³⁷. The courts have taken quite a restrictive view on planning conditions and, in the case of *Newbury District Council v Secretary of State for the Environment*³⁸, it was held that conditions must

- be imposed for a planning purpose and not for an ulterior motive

³¹ Planning and Compulsory Purchase Act 2004, Section 38 (6).

³² *City of Edinburgh Council v Secretary of State for Scotland* [1998] JPL 224; *R v Leominster District Council ex parte Potheary* [1998] JPL 335.

³³ *R (on the application of the Council for National Parks Ltd.) v Pembrokeshire Coast NPA and ors.* [2005] EWCA Civ 888.

³⁴ Planning for a Low Carbon Future in a Changing Climate, 9 March 2010.

³⁵ Town and Country Planning Act 1990, Section 70 (1).

³⁶ *Ibid.*, Sections 72 and 75.

³⁷ Circular 11/95.

³⁸ [1981] AC 578.

- fairly and reasonably relate to the development permitted
- not be perverse (so unreasonable that no reasonable authority could have imposed them)

5.6 General Permitted Development Order

There are certain ways in which development can be permitted under statute and thus the need for a planning application is obviated. One such route is available under the General Permitted Development Order³⁹. There are a large number of development types listed in Schedule 2 that cover minor developments, developments carried out by public services and favoured activities (such as agriculture and forestry). Development consent may still be needed when projects exceed certain thresholds and the right to development can be withdrawn, for example, if an environmental impact assessment is required.

6 Planning and Micro Generation

The 2007 Government White Paper ‘Meeting the Energy Challenge’⁴⁰ recognised that planning consents are a major constraint on the implementation of a future energy strategy. Indeed the Scottish Executive has acknowledged that the planning system has the potential to act as a tool in opposing nuclear power (Dow, 2007). There is certainly little doubt that barriers have been encountered in the development of micro generation, but conversely the planning system has also been a driver behind the adoption of on-site renewable energy, which will have a significant role to play in delivering the Government’s 2016 zero carbon homes agenda (Sustainable Energy Partnership, 2007).

6.1 The Merton Rule

The Merton Rule takes its name from Merton Council. It amounts to a borough wide prescriptive planning policy for all buildings, which was developed and adopted by the Council in 2003. The policy requires new developments to generate at least ten percent of energy needs from on-site renewable technology. The normal threshold for application of the rule is ten homes or 1,000 square metres of non-residential development. The Merton policy has had a significant impact and was subsequently adopted by the Mayor of London and the majority of local authorities nationwide. Planning Policy Statement 22 on renewable energy expressly acknowledges the Merton Rule and advocates its adoption by local planning authorities and the encouragement of renewable energy projects through local planning documents.

However, the future of the rule is uncertain. Whilst Merton Council intends to extend the policy to cover all development in Merton and is considering whether it is appropriate to increase the percentage of the policy up to a 20% requirement (Merton Council, 2010), it has been suggested (Sustainable Energy Partnership, 2007) that the rule was watered down in the 2007 Climate Change supplement to PPS1 by the removal of the requirement to consider renewable energy projects ‘in **all** new developments’.

³⁹ Town and Country Planning (General Permitted Development Order) 1995, SI 1995/418.

⁴⁰ *Meeting the Energy Challenge: A White Paper on Nuclear Power*, Cm 7296, 10 January 2008.

The emphasis also seems to have shifted from a requirement for a percentage of on-site renewable energy to consideration of the possibility of utilising off-site renewable energy supplies. This latter development is subject to criticism as, it is suggested that, linking new housing to off-site renewable energy developments provides no additional cut to carbon dioxide emissions (Sustainable Energy Partnership, 2007). The recent draft supplement to PPS 1 essentially prohibits the adoption of the Merton Rule in the future, stating that;

‘Targets for application across a whole local authority area which are designed to secure a minimum level of decentralised energy use in new development will be unnecessary when the proposed 2013 revisions to Part L of the Building Regulations..... are implemented.’

However, it should be noted that currently a disclaimer appears on the Department for Communities and Local Government website, where the draft PPS appears, warning that all the content is subject to review in the light of the recent change of government.

6.2 Permitted Development Order

One of the problems with micro generation and the planning system traditionally lies with local authorities taking different approaches to small household projects. For example, some allow small turbines on houses whereas some do not and similarly with solar panels, whilst allowed in some areas, other local authorities view panels as damaging to conservation areas (Dow, 2007). The Climate Change and Sustainable Energy Act 2006 allows for such differences in practice to be reduced through amendment to the permitted development order. After due consultation, the Secretary of State made an order⁴¹ amending the General Permitted Development Order⁴² allowing permitted development for the installation of domestic micro generation equipment. Within the framework of the restrictions and conditions outlined in the Order, solar panels, heat pumps and biomass heating systems are all subject to the Order. This should make it easier for households to install micro generation equipment because, in many cases, it will no longer be necessary to apply for planning permission. It should also lead to greater consistency across local authority areas.

7 Planning Conditions and Sustainable Buildings

The supplement of Planning Policy Statement 1 on climate change urges planning authorities and developers to ‘engage constructively and imaginatively to encourage the delivery of sustainable buildings’. The Statement also acknowledges that ‘There will be situations where it could be appropriate for planning authorities to anticipate levels of building sustainability in advance of those set out nationally’. Planning authorities are advised to focus on development area or site-specific opportunities and specify requirements in terms of nationally recognised standards such as the Code for Sustainable Homes. Some local planning authorities have sought to achieve the goal of driving up standards of sustainability in new domestic dwellings in their areas by recommending enhanced levels of performance and certification under recognised codes

⁴¹ The Town and Country Planning (General Permitted Development) (Amendment) (England) Order 2008.

⁴² Town and Country Planning (General Permitted Development) Order 1995.

and schemes.

7.1 Brighton and Hove City Council

An example of this practice is to be found in the 'Model Planning Conditions and Informatives' of Brighton and Hove City Council. The model planning conditions state that, unless otherwise agreed in writing by the Local Planning Authority, no new build residential development can commence without evidence that it will achieve a minimum code level 3 under the Code for Sustainable Homes. Furthermore (again, unless agreed in writing), the approved units cannot be occupied until a final code certificate is issued confirming the minimum code level 3 performance. Similarly, the model conditions state that unless agreed with the Local Planning Authority, no residential development involving existing buildings can commence unless it is certified that the development will achieve an 'Ecohomes' rating and no occupation can take place until a post construction certificate to this effect has been submitted to and approved by the Authority. In a similar way, BREEAM registration, assessment, rating and confirmation are conditions of development and occupation for new build non-residential developments. Such conditions in planning determinations might, however, be open to challenge as is illustrated by the two following case studies.

7.2 The former New Penny Public House

Planning permission had been granted subject to conditions for the construction of 12 new build flats on the site of a former public house. An appeal was brought by the developers against the decision of Cheltenham Borough Council. One single condition to the planning permission was disputed. The condition in question stated that;

prior to the commencement of development a scheme to demonstrate a reduction in carbon emissions to achieve a minimum level of code 3 of the Code for Sustainable Homes shall be submitted to and approved in writing by the local planning authority. The development shall be carried out and maintained in accordance with the details so approved.

The reasons given for the condition were to ensure compliance with national and regional objectives and the aims of local plan policy CP1 regarding sustainable development. The latter policy stated that development would be permitted only where it took adequate account of the principles of sustainable development and it set out a number of criteria for this. The policy also stated a number of principles of sustainable development which might be taken into account as material considerations in the determination of planning applications.

The appeal was allowed, the Inspector concluding that the condition was imprecise and unreasonable and that it did not meet the tests set out in Circular 11/95. A new planning permission was granted without the disputed condition but retaining the relevant non-disputed conditions from the previous permission. The key reasons for allowing the appeal were as follows:

- The Supplement to PPS1 (the Supplement), states that councils wishing to proceed in advance of the Government's timetable must set out their policies for sustainable developments in development planning documents (DPDs) so as to

ensure examination by an independent Inspector. Neither policy CP1, *Sustainable Development of the Cheltenham Local Plan Second Review* nor the Council's supplementary planning guidance documents relating to policy CP1 referred to the Code for Sustainable Homes.

- Paragraph 42 of the Supplement allows non-compliance with adopted DPD policies if this is not feasible or viable. The appellant argued that the development had not been designed to attain code level 3 and that fundamental changes, involving a different design to that which had been approved, would be required to achieve this. The cost involved, it was argued, would make the proposed development unviable. Little evidence was proved by the council to counter these claims.
- The approach set out in the Supplement indicates that conditions requiring compliance with level 3 of the code or above should only be imposed if the developer has demonstrated a willingness to comply. Condition 13 was unreasonable without the developer's agreement to it.
- The developer had committed to construction of the flats in accordance with the Code to a level to be determined at the time of construction and to a number of measures of sustainable methods of construction. The Inspector concluded that the measures incorporated in the scheme met the sustainability objectives identified in policy CP1. It was also noted that it was normal to assess the code level of a development post-completion of the development rather than pre-commencement as required by condition 13.

(Appeal Decision, 2009)

7.3 Hut Cottage

Planning permission had been granted subject to conditions for a replacement bungalow at Hut Cottage, a single-storey dwelling with outbuilding on a small and irregularly shaped plot of land abutting listed buildings and in a conservation area. Planning permission had been granted for a replacement bungalow in 2005, which had expired. A further application of 2008 seeking renewal of the planning permission was for a proposed development identical to the earlier scheme. When planning permission was granted it was subject to conditions, one of which was disputed in an appeal against the decision of Chelmsford Borough Council. The condition in dispute (No. 5) stated that, unless otherwise agreed in writing by the Local Planning Authority:

- a) the development hereby permitted shall be built to a minimum of Level 3 of the Codes for Sustainable Homes (or its successor);
- b) No development shall take place until a Design Stage assessment (under the Code for Sustainable Homes or its successor) has been carried out and a copy of the summary score sheet and Interim Code Certificate have been submitted to and approved in writing by the Local Planning Authority;
- c) Prior to the first occupation of the dwelling, a copy of the summary score sheet and Post Construction Review Certificate (under the Code for Sustainable

Homes or its successor) shall be submitted to the Local Planning Authority verifying that the agreed standards have been met.

The reasons given for the condition were to achieve sustainable development in accordance with Policies CP11 and DC24 of the Adopted Core Strategy and Development Control Policies Development Plan Document (CSDCP) and the Sustainable Development Supplementary Planning Document (SPD). Policy CP11 provided guidance relating to energy and resource efficiency, renewable energy and recycling and Policy DC24 established criteria relating to energy efficient design and the use of materials, including a requirement that all new dwellings should attain a minimum rating of Level 3 of the Code for Sustainable Homes or its successor.

The appeal was dismissed and the disputed condition was found to be reasonable and necessary. The reasons given by the Inspector were as follows:

- There had been changes in policy in response to growing concerns surrounding global warming since the first permission was granted, which were reflected in the 2007 Supplement to PPS1. The Council's determination and conditions were guided by the policies within the CSDCP.
- It was noted that the Planning and Compulsory Purchase Act 2004⁴³ states that regard must be had to the development plan unless material considerations indicate otherwise.
- It was also noted that where planning permission expires, fresh applications should be judged against current planning considerations⁴⁴.
- The decision to reject the appeal was made notwithstanding the resultant financial implications and the voluntary nature of the Code for Sustainable Homes.

(Appeal Decision, 2009)

7.4 The Future of Requirements for Sustainable Buildings in Local Planning

The two appeals outlined above deliver some interesting lessons if planning conditions are to be successfully utilised to attain higher standards than currently required under building regulations. Clearly, the requirement of attaining a particular level of an accepted national code such as the Code for Sustainable Homes is, in principle, acceptable (at least under current planning guidance). The key message to be drawn from the different outcomes of the two case studies above, is that explicit reference to the code in question should be contained within the development plan document (DPD). The question as to whether requirement can be made for schemes of compliance to a level of performance pre-commencement rather than post-completion is less clear cut. To err on the side of caution local planning authorities might be best advised to stick to post-completion requirements in order to minimise the chances of successful appeal. It

⁴³ Planning and Compulsory Purchase Act 2004, Section 38 (6).

⁴⁴ Circular 08/2005, Guidance on Changes to the Development Control System.

is interesting that in the Hut Cottage appeal, the fact that the condition was not imposed with the agreement of the developer was given little weight. Also, although the Inspector noted the financial consequences of the imposition of condition No. 5 the issue of the feasibility and viability of compliance with level 3 of the code were not really explored in depth.

Although the future of the new draft supplement to PPS1 is uncertain, the draft policy on the local planning approach to setting requirements for sustainable buildings, as it is currently framed, would continue to allow requirements for a building's sustainability as long as they are set out in the DPD. The approach is, however, considerably more restrictive than that in the current supplement. Requirements should

‘relate to a development area or specific sites and not be applicable across a whole local authority area unless the justification for the requirement can be clearly shown to apply across the whole area;’⁴⁵

If this is retained in any new PPS, then it will be much harder for local authorities to adopt a requirement to meet elevated levels of a code across their whole area. Certainly they will have to provide clear and convincing justification within their DPD if this is to be the case.

8 Conclusions

The law surrounding climate change and energy efficiency is complex and multi-layered. This complexity is accentuated by the multiple aims involved and the fact that measures adopted do not always align to all of these. What is clear is that buildings are of central importance in realising international and local goals to reduce levels of green house gas emissions and thus help counter the threat of climate change. There are many ways of tackling the reduction of emissions from buildings. Alternative energy sources, renewables and energy efficiency are just some of these. Legislation will not provide the whole solution, nor will one single legislative route.

Regulations adopted under the Building Act 1984 are clearly of key importance in implementing the standards required of buildings in order to meet the targeted reductions in carbon dioxide emissions and the ultimate aim of carbon zero buildings. However, voluntary schemes, such as the Code for Sustainable Homes, also have an important role to play. This is clearly illustrated by the way in which the new Part L standards in the Building Regulations mirror level 3 of the Code as well as the way in which the Code has been used to drive up standards through the planning process.

New proposals for planning guidance appear to suggest that in some respects the role of the planning system is done. The proposal is to outlaw initiatives such as the ‘Merton Rule’ and restrict the requirements that can be imposed at the planning stage on the sustainability of buildings. Whilst there is doubtless some merit in having a single central level of control in the guise of building regulations, the building control system operates in a different way and at a different stage in the building life cycle to the

⁴⁵ PPS: Planning for a Low Carbon Future in a Changing Climate, Part 2; Consultation Draft, Policy LCF9.1 (i).

planning regime. Having a single central standard also ignores variations in local environments and stifles the drive to strive for better standards and develop new and affordable technologies, which have in the past been encouraged by voluntary code ratings and compulsory on-site renewable energy requirements.

There is great uncertainty at the moment as to the direction that planning law will take. The new government promises radical reform of the planning system and has pledged through its proposed Decentralisation and Localism Bill⁴⁶ to devolve greater powers to local authorities. Specifically, it plans to abolish Regional Spatial Strategies and ‘return decision-making powers on housing and planning to local councils’⁴⁷. It may, therefore, be that the trend outlined above will be reversed and local communities and their planning authorities will be able to lead the way in delivering alternative energy sources and sustainable buildings.

9 References

- Appeal Decision (2009) ‘Appeal Ref: APP/W1525/A/09/2096523 Hut Cottage, Colemans Lane, Danbury, Chelmsford CM3 4DN’ by Roger P Brown. Decision date 17 June 2009.
- Appeal Decision (2009) ‘Appeal Ref: APP/B1605/a/09/2097989 Former New Penny Public House, 84 Gloucester Road, Cheltenham, Glos. GL51 8NZ’ by Elaine Benson. Decision 11 June 2009.
- Barker, T. *et al* (2007) ‘Climate Change 2007: Mitigation of Climate Change: Contribution of Working Group III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change’, Cambridge University Press, Cambridge, Box 13.7, p 776.
- DECC (Department of Energy and Climate Change) ‘Final Report 2008’, http://www.decc.gov.uk/en/content/cms/statistics/climate_change/gg_emissions/uk_emissions/2008_final/2008_final.aspx, viewed 30.7.2010.
- den Elzen, M. (2008) ‘Emission Reduction Trade-Offs for Meeting Concentration Targets’, Bonn Climate Change Talks, IPCC workshop, UNFCCC SBSTA 28.
- Dow, S. (2007) ‘Climate Change and Sustainable Energy Act 2006’, *Environmental Law Review*, **Volume 9** (4), pp 279-284.
- Energy Saving Trust ‘Code for Sustainable Homes: New Build Housing’, Briefing Note, <http://www.energysavingtrust.org.uk/housingbuildings>, viewed 30.7.2010.
- Grekos, M. (2008) ‘House of Commons Report on Tackling Emissions from Existing Housing’, *Journal of Planning and Environment Law* (8), pp 1123-1124.
- Grekos, M. (2009) ‘Climate Change Act 2008’, *Journal of Planning and Environment Law* (4), pp 454-455.
- Hookins, D. and Stonehill, C. (2006) ‘Energy Performance of Buildings Directive – Time for an Energetic Approach’, *Landlord and Tenant Review*, **Volume 10** (6), pp 175 – 180.
- McAdam, W.B. (2007), *Built Environment Carbon Emissions: A Policy Review*, MSc Construction Law and Arbitration Dissertation, Leeds Metropolitan University.

⁴⁶ Queen’s Speech, 25/5/10.

⁴⁷ Department for Communities and Local Government Draft Structural Reform Plan, July 2010.

- Merton Council, <http://www.merton.gov.uk/living/planningpolicy/mertonrule>, viewed 30/7/2010
- Mittenthal, L. (2009) 'Update on Progress Towards Enactment of Revised Energy Performance of Buildings Directive', *Environmental Liability*, **Volume 17** (5), pp 78-79.
- Rajamani, L. (2009) 'Addressing the 'Post-Kyoto' Stress Disorder: Reflections on the Emerging Legal Architecture of the Climate Regime', *International and Comparative Law Quarterly*, **Volume 58** (4), pp 803-834.
- Stallworthy, M. (2010) 'UK Climate Change Law: Securing a Low Carbon Economy by Recourse to Experts and Government Obligations?', Proceedings of RICS COBRA Conference, Paris, 2nd & 3rd September, 2010.
- Sustainable Energy Partnership (2007) 'The Merton Rule: The Orwellian World of Sir Humphrey, Minus Means Plus', http://www.microgenscotland.org.uk/SEP_Broadsheet_2%5B2%5Dfinal16.10.pdf, viewed 30.7.2010.
- Thomas, M. and Woodward J. (2010) 'Copenhagen: Hopenhagen? Unfortunately not', *European Lawyer*, (92), p 7.
- Townsend, H. (2009) 'The Climate Change Act 2008: Something to be Proud of After All?', *Journal of Planning and Environment Law* (7), pp 842-848.
- UNDP (United Nations Development Programme) (2007/2008), 'Human Development Report, Fighting Climate Change: Human Solidarity in a Divided World', *UNDP*, http://hdr.undp.org/en/media/HDR_20072008_EN_Complete.pdf, viewed: 30.7.2010.
- UKGBC (UK Green Building Council) 'Making the Case for a Code for Sustainable Buildings' March 2009.
- Winter, G. (2010) 'The Climate is no Commodity: Taking Stock of the Emissions Trading System', *Journal of Environmental Law*, **Volume 22** (1), pp 1-25.

Equivalence in Dutch building regulations

Dr. N.P.M. Scholten PhD¹

Expertcentre Regulations in Building

Email: n.scholten@bouwregelwerk.org

Summary:

The Dutch Buildings Decree 2003 [*Bouwbesluit 2003*] and the Decree on fire safe use of structures [*Gebruiksbesluit*] are based upon the performance principle. They are a more detailed elaboration of concise functional definitions. To facilitate innovationproof of equivalent performance is allowed.

One can demonstrate that the aim of the regulations is being achieved to at least the same extent as required by the performance based clauses. Upon application for a permit or submitting a notification of fire safe use, the applicant has to demonstrate equivalence. In the case of a nonsuit the municipality should motivate to why equivalence has not been achieved.

Equivalence should be a purely technical issue. Local authorities are not permitted to impose requirements that are additional to, or that deviate from the performances specified. The many arising conflicts can be attributed to inadequacies in the official explanatory memorandum, to the inadequacy of the regulations themselves or to lack of knowledge by the participants in the decision process.

No authoritative technical institution however exists that has the statutory powers to resolve disputes in this particular area. The foundation Expertcentre Regulations in Building does fulfil this role on a *de facto* level. The Ministry of Housing, Spatial Planning and the Environment (VROM) recently set up an independent advisory committee for the practical application of fire safety regulations. That particular body has not been allotted any official authority.

It is illustrated that the applicant in practice often abides by the decision of the municipality, because instituting legal proceedings involve considerable cost and time loss. Formal objection and appeals proceedings under the General Administrative Act [*Algemene wet bestuursrecht*] take a long time to run their course. Mostly only procedure has judged without a judgment about the contested technical issue.

Improvements in the regulations and explanatory memorandum and greater coordination are proposed, while opinions are given regarding a central technical authority that issues decisions upon municipalities and citizens in disputes of a technical nature.

Keywords:

Building regulations, equivalence, technical disputes, court decisions, technical tribunal

Introduction

The technical building regulations in the Netherlands are based upon the Housing Act [*Woningwet*]. They are detailed in the Buildings Decree [*Bouwbesluit*] 2003 [1], the Building Decree Regulations [*Regeling Bouwbesluit*] 2003 and the Decree on fire safe use of structures [*Gebruiksbesluit*] [2]. On October 1st 2010, the General Stipulations Environment Act [*Wet algemene bepalingen omgevingsrecht – Wabo*] will enter into force. As a result, the legal basis of the building regulations will partly be dealt with under that particular Act, though the content of General Stipulations Environment Act shall not lose its power. The current situation is reproduced in Figure 1.

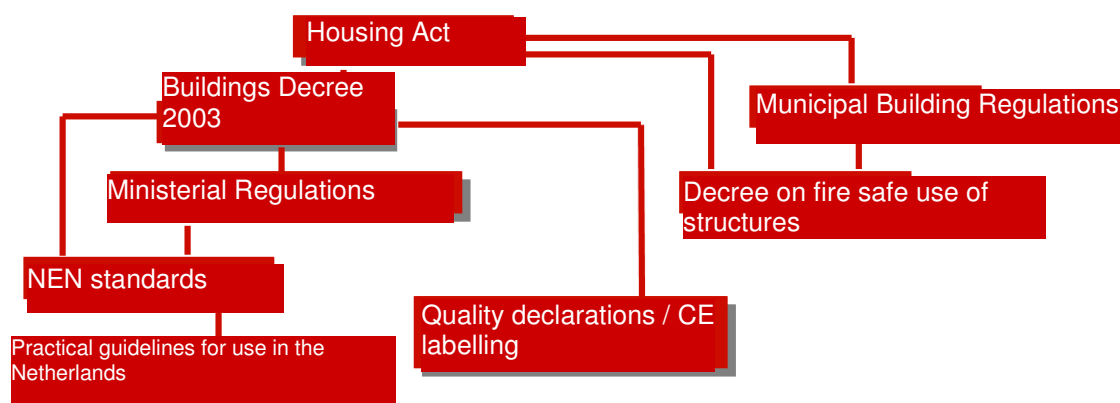


Figure 1 – Interrelationship between legislative documents up to 01.10.2010

The future situation is depicted in Figure 2.

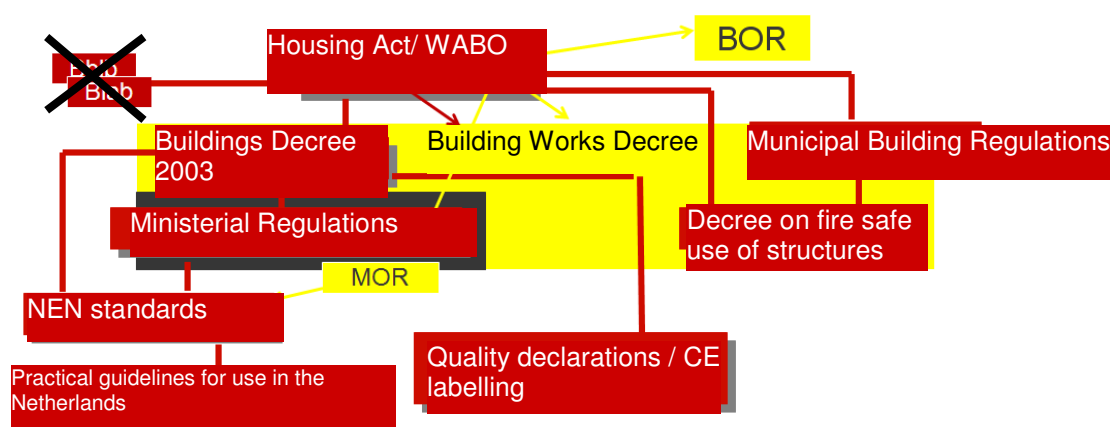


Figure 2 – The interrelationship of legislative documents from 01.10.2010 onwards (Building Works Decree [*Bouwwerkbesluit*] expected to take effect on 01.01.2012)

The building regulations contained within the Buildings Decree 2003 are designed to fulfil the purposes represented in Figure 3.

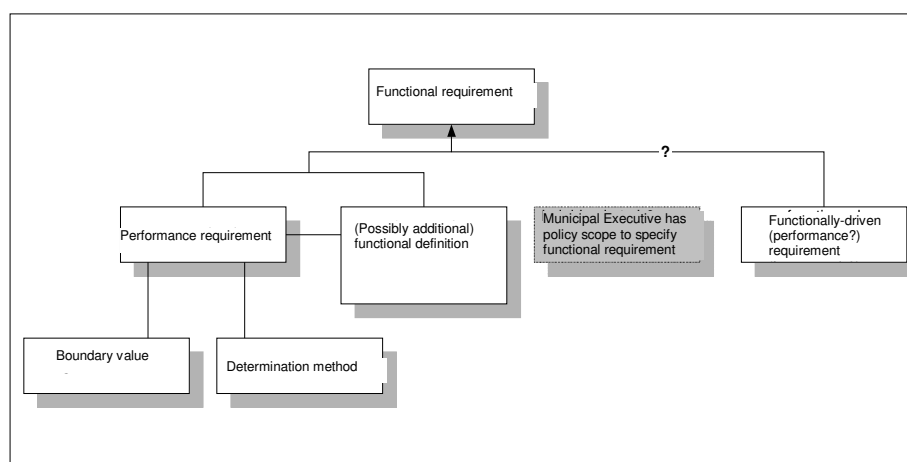


Figure 3 – Nature of the regulations

The functional requirements are, however very general in nature. The extent to which performance requirements can be related to the functional requirements is insufficient. Background documentation to the performance requirements is not available. The Explanatory Memorandum to the Buildings Decree 2003 provides insufficient detail with regard to the relationship between the performance requirements and the underlying objectives of the regulations.

This also applies in the case of the regulations contained within the Decree on fire safe use of structures [*Gebruiksbesluit*]. In many instances, this Decree fails to refer to the objectives and contains only the performance requirements, coupled with regulations that relate to possible solutions. The Explanatory Memorandum to the Decree reveals that the text of the Article fails to address the objective sufficiently and is based upon worst-case scenarios that frequently do not apply in practice. According to that explanation, it is then necessary to provide a tailor-made solution, through application of Article 1.4 (Equivalence). Due to the fact that no description is included of the objectives underlying the regulations, there is insufficient information available for any such tailor-made solutions to be provided. From a legal perspective, a tailor-made solution of that type represents dispensation and not equivalence.

Equivalence

In practice, situations may occur in which the performance requirements of the Buildings Decree 2003 are insufficiently, or not tailored at all. Situations such as those frequently involve structural solutions or methods of calculation that cannot automatically be assessed against the relevant performance

requirements. As far as the interior of a building is concerned, this could relate to a grey water system or ventilation that occurs via a conservatory. It may also involve an advanced method of calculation or a method of calculation imported from another country that has not been included in the standards documentation in the Netherlands, such as a sewerage system based upon working principles other than a primary venting system.

The legislators responsible for the Decree did not wish to prohibit any such solutions in advance. The Buildings Decree 2003 intends not to hinder the application of innovative or experimental building products or methods of construction. To this end, Section 1 of the Buildings Decree 2003 incorporates what is known as an equivalence clause, which reads as follows:

§ 1.3. Equivalence clause

Article 1.5

A regulation given in the second to sixth chapters which has to be applied in order to comply with a requirement regarding a construction or a part of it, does not have to be complied with, in so far as by applying that regulation the construction or the relevant part of it at least provide the same degree of safety, health protection, functionality, energy economy and environmental protection as intended with the relevant regulation.

A structural solution shall be deemed to apply if this takes the form of a structural provision, i.e. a provision that forms an irrevocable part of a building and cannot be changed by simple means. In the event that the Buildings Decree 2003 does not yet contain a functionally edited regulation, the application of the equivalence principle does not apply. In such cases, the applicant shall be required to give reasons why he/she believes that his/her solution fulfils the regulation and the Municipal Executive [B&W] will assess out that solution.

The same principle is also adopted by the Decree on fire safe use of structures, but only in so far as it relates to the interrelationship between the use of a building and fire safety.

If, when assessing an existing building or the actual use made thereof, a municipality is of the opinion that the use being made of the building is not in accordance with the law, the municipality will need to issue a motivated notice imposing an administrative penalty or incremental penalty payment in an attempt to bring the situation to an end by forcing the owner or user of the building to take appropriate measures. The municipality has also to examine whether equivalence solutions are used.

Objections and appeals

In accordance with Sections 7 and 8 of the General Administrative Law Act [*Algemene Wet Bestuursrecht*], the relevant party is entitled to submit an objection and appeal against any decision taken by the Municipal Executive [B&W]. This 'Relevant Party' is understood to refer to:

The applicant for a building permit, or in the case of an order - the owner or user.

In the first of these two cases, the decision may have been taken not to accept an application for a building permit or not to issue a building permit. The second case stated above could relate to the threat of closure or demolition.

The first step is to submit a written objection to the government body that took the decision, in this case the Municipal Executive [B&W]. It is subsequently possible to appeal against the decision taken with regard to the objection at the Administrative Chamber of the district court. It is then possible to appeal against the decision of the court by appealing to the Administrative Case Law department [*Afdeling Bestuursrechtspraak*] of the Council of State.

The judicial system is not equipped to reach a specialist technical judgment with regard to the factual correctness of a decision taken by the competent authority. In many cases, the most that the courts are able to do is to examine –whether the decision taken by the competent authority was reasonable? For the Plaintiff, however, that is often very unsatisfactory.

Another reason why this situation is often unsatisfactory can be attributed to the fact that no documentation outlining the background of the regulations is available to the administrative court, which means that the court has no information at its disposal against which it can evaluate the decision taken by the municipality. Partly due to a lack of technical knowledge, the administrative court is not in a position to judge whether a municipality has applied an incorrect interpretation of the rules or whether it has unjustifiably imposed additional or different requirements when applying the equivalence clause. The administrative court is also unable to judge whether a municipality was correct in ruling that equivalence had been satisfactorily achieved.

Authoritative judgments

The legislator stipulated that the final decision with regard to the technical assessment should be taken by the Municipal Executive [B&W] and the district court has the authority to hear any disputes that may arise. A municipality is not obliged to accept the opinion of an expert to be correct. If, however, a municipality does not accept a recommendation, drawn up on behalf of an applicant for a building permit, an owner, or a user, it shall be required to provide reasons for such non-acceptance.

Under Dutch law, no formal institution exists that possesses the statutory powers to issue judgments relating to equivalence, with the exception of the quality declarations drawn up under the recognised system of quality declarations. There are, however, a number of bodies that are regarded as authoritative within the market, such as TNO, Efectis Netherlands BV, which deals with fire safety,

and the Expertcentre Regulations in Building, which covers the full range of issues relating to regulations. Recently, the Ministry of Housing, Spatial Planning and the Environment set up an independent advisory committee for the practical application of fire safety regulations [1]. This committee also lacks any formal powers but acts as the government-financed competitor to the aforementioned private initiatives/institutions, which are working well in practice.

Case studies

Case study no. 1: The conversion of an unused farm into a restaurant

After the remodeling lasting more than 2 years is over (the works had been supervised by the municipality), a Prohibition of Use order is imposed upon the owner, just before the restaurant is about to open. According to the municipality, the upper floor must be capable of withstanding a representative floor load value of 5 kN/m² during a projected service life of 50 years. The municipality stated that this had not been demonstrated.

The owner commissions an expert report, which states that the municipality should not solely apply NEN 6702, which imposes requirement of 5 kN/m² in the case of new buildings, but that by virtue of the Buildings Decree 2003, the municipality should actually apply NEN 6700 (the underlying philosophy behind building regulations (comparable to NEN-EN 1990), taking into account the possibility under the law to exempt the building from the level stipulated in relation to new buildings and to allow it to be based upon a floor strength lower than 5 kN/m². To this end, the applicant availed himself of a report by TNO that is regarded as authoritative, which was mentioned in the explanatory notes relating to the national legislation that specifies safety levels and the planned service life that are applied to buildings undergoing conversion. The recommendation is to assume a value of 3.5 kN/m², which is already extremely high, and a planned service life of 25 years, in conjunction with the reliability level that TNO had established for buildings undergoing renovation. These therefore give rise to lower load factors. The design floor load is therefore reduced by 7.5 kN/m² to 4.28 kN/m². The municipality does not agree with this and in a preliminary injunction, the court rejects the owner's claim [3]. The independent committee, that advises the municipality, finds the opinion of the owner's adviser to be well-founded, but wishes to receive a second opinion from TNO. TNO confirmed the opinion expressed by the expert consultant. The municipality does not agree to this and it issues additional conditions, such as screwing tables to the floor etc. In view of the professional opinions obtained thus far, the objections committee recommends tolerating the status quo. The municipality is still in doubt, however. The alderman invites the owner's expert and TNO to attend a meeting. Three months later and based upon the discussion at that meeting, the municipality then proceeds to make a decision regarding the objection (the restaurant has meanwhile been unable to operate for 10 months), agreeing to 3.5 kN/m² and a projected design lifespan of 25 years, but imposing the safety factors that would apply to a service life of 50 years (the municipality is still applying the load factor of 1.5, as

opposed to a load factor of 1.3 that applies to renovated buildings). The municipality has not given permission for the actual use being made of the building, but requires a calculation, demonstrating that the floor of the upper storey is strong enough, even though the details are in fact already in the municipality's own archive. The municipality does not proceed to pay compensation to the owner. The owner then submits an appeal against the decision taken by the municipality.

This case illustrates the consequences that a lack of knowledge on the part of a municipality can have for an owner, in addition to the losses that such a lack of knowledge can cause. The amount of time taken by the municipality to reach a decision is also excessive.

Case study no.2: Construction of a non-compartmented, 20.000 m² office building with sprinkler system, but with limited fire load [4]

Despite the presence of a sprinkler system and despite the fact that the building complies with the pseudo-regulation developed by the fire brigade in order to determine whether a building is deemed to be safe, the municipality grants the building permit, on condition that the building be compartmented, each compartment not exceeding a total area of 5,000 m². Complying with this requirement would mean that the transparent nature of the proposed building would no longer exist. The municipality and the fire brigade justify their additional requirements by referring to a metro tunnel located in the vicinity of the building. The building regulations do not actually provide a facility for factors of that type to be taken into account.

A simple risk calculation can be used in order to demonstrate that the additional requirement imposed was excessive. A calculation of this type demonstrates how the various factors that contribute to the risk are made up of a combination of probabilities and the scope of any potential damage.

No sprinkler, 1000 m2	Probability	Probability	Probability	Probability	Damage	Probability	Risk
Fire peters out or is extinguished	0.5				1	0.5	0.5
Fire in 1 st compartment, separation OK	0.5	0.95			1000	0.475	475
2 compartments	0.5	0.05	0.2		2000	0.005	10
3 compartments	0.5	0.05	0.8	0.05	3000	0.001	3
>3 compartments	0.5	0.05	0.8	0.95	20000	0.019	380
					Total	1	868.5

Sprinkler, 5000 m2	Probability	Probability	Probability	Probability	Damage	Probability	Risk
Fire peters out or is extinguished	0.5				1	0.5	0.5
Sprinkler effective, fire in 4 sprinkler heads	0.5	0.96			9	0.48	4.3
Sprinkler ineffective, fire in 1st compartment	0.5	0.04	0.05		5000	0.001	5.0
2 compartments	0.5	0.04	0.95	0.01	10000	0.00019	1.9
>2 compartments	0.5	0.04	0.95	0.99	20000	0.01881	376.2
					Total	1	387.9

Sprinkler, 20,000 m2	Probability	Probability	Probability	Damage	Probability	Risk
Fire peters out or is extinguished	0.5			1	0.5	0.5
Sprinkler effective, fire in 4 sprinkler heads	0.5	0.96		9	0.48	4.3
Sprinkler ineffective, fire in entire building	0.5	0.04	1	20000	0.02	400.0
				Total	1	404.8

Discussion of the risks calculated:

The figures outlined above lead us to the following conclusions:

- The level of safety of a sprinklered building that is not subdivided into sub-compartments will hardly improve at all if it is subdivided into fire compartments of 5,000 m². The increase in safety is estimated to be around 4% (an area of 405 m² affected is reduced to 388 m²).
- This increase in safety is not proportional to the cost of the investments that would be required in order to subdivide the building into 5,000 m² fire compartments.
- The increase in safety is entirely negligible to the increase in safety that is provided, not by subdividing the building into fire compartments, but by installing a sprinkler system (area of 868 m² affected is reduced to 405 m²).
- 'Equivalence' only requires the building to comply with the Buildings Decree 2003, which stipulates an affected area of 868 m². In view of the fact that the relevant area affected would be 405 m², the solution to install a sprinkler system, whilst foregoing the subdivision of the building into fire compartments, provides a considerably higher level of safety.

The real estate company submitted an objection against this additional requirement.

In this particular case, the consultants involved adapted the building plans in order to illustrate the consequences of particular solutions, based upon a risk evaluation derived from expertise in the likelihood of failure. Armed with this information, they attempted to explain the underlying objective of the regulations to the administrative court.

Case study no. 3 : A large discotheque, incorporating two café-restaurants, relating to safe evacuation facilities[5]

The discotheque originally opened twenty-four years ago, with an occupancy permit stating that the building could contain a maximum of 1500 people at any one time. In 1998, without any changes having been made to the regulations, that number was reduced to 1100 people. In 2002, a further decision was issued, stating that only 938 people could be in the building at any one time. An objection was submitted against both decisions. The case is still being processed by the objections committee.

The building complex incorporate an extensive video surveillance system and an adequate First Aid system, but do not include a fire alarm system or evacuation system compliant with the NEN 2535 and NEN 2575 standards. Safety is assured based upon the existing systems. Detailed reports have since been drawn up by the three most reputable groups of expert consultants in the Netherlands. A calculation of the smoke filling time carried out by TNO and repeated by Efectis Netherlands bv has demonstrated that the time available for evacuation is considerably greater than the evacuation time

used by the fire brigade in its recommendations. Based upon a standardised evacuation flow model, it was possible to demonstrate that the discotheque could be safely used by 1538 people. Based upon its own rules, the Fire Brigade is maintaining that only 938 people may use the premises at any one time. The sub-district administration requested a second opinion from a well-known consultancy and that opinion supported the viewpoint adopted by the experts appointed by the owners.

The Fire Brigade informed the sub-district administration that if its advice is not followed, this may have repercussions in terms of the Fire Brigade's ability to extinguish any fire that might occur. The municipal council was also consulted, though it does not possess any formal powers in this regard. The outcome of that particular consultation was a requirement to conduct a full-scale evacuation exercise, on the basis of which the regulations of the Buildings Decree 2003 were clearly being questioned. The competent authority is not empowered to do this.

This case has been dragging on for more than 8 years. It appears to be a battle played out between the Fire Brigade on the one hand and the competent authority on the other, whilst the owner of the premises continues to be the victim of that battle. What is needed in this case is a decisive administrative response. The fact that this has not been forthcoming can be attributed to the fact that as an administrative body, the competent authority, frequently needs to hold discussions with the Fire Brigade and therefore does not wish to upset the mutual relationships unduly. For as long as the sub-district administration fails to take a decision in relation to the objection, the owner-operator of the premises has no legal remedies at his disposal to question this course of events.

Case study no. 4 Demand-driven ventilation [6]

The ventilation requirements included within the Buildings Decree 2003 are based upon the principle that a building can be ventilated at any time, whenever the user so desires. The performance requirements are based upon a definition of standard use and a set of boundary values recommended by the Dutch Health Council which assume that human activity will generate a much larger flow of air than is necessary from a scientific perspective. Thanks to a number of innovative developments, different ventilation systems are now available on the market, including demand-driven ventilation. Systems of that type provide ventilation, based upon the degree of contamination that is present in a specific room within the building. Declarations of equivalence have also been issued for systems of that type, based upon the Vencon software package produced by TNO. The fact that the air flows over the year are reduced means that buildings that incorporate systems of this type consume less energy than buildings ventilated in more traditional ways. This means the relevant TNO declaration also indicates the manner in which a ventilation system of this type has a positive effect upon the standardised calculation of the energy performance.

This declaration does however give rise to important discussions, due to the fact that at no point it indicates the foundations upon which the ventilation regulations and ventilation paragraph of the energy performance provisions are based.

Again this learns us that the background information underlying the regulations gives rise to a whole range of discussions that not only incur a high level of administrative costs, but also cause the construction process to be delayed. This case also demonstrates that those involved require a detailed knowledge of physics and of the background underlying the regulations, in order to be able to apply the principle of equivalence. From a practical point of view, a judge in an administrative court is not in a position to form a judgment upon such matters.

Case study no. 5 Beach volleyball centre in former factory premises[7]

Inside a steel building that is 9 metres high and that rests on a concrete floor, six sandy areas are in use for beach sports. Next to these sandy areas evening dance parties are held at weekends. The building is set up in such a way that more than 7.5 minutes are available after a fire alarm to evacuate it safely (thanks to the smoke and heat extractor, fire alarm system, evacuation alarm system and the building's external cladding etc). The exits are evenly distributed along the external walls of the building and any exit can be reached within 30 metres. A number of the emergency exits are immediately adjacent to the sandy areas. Calculations based upon worst-case scenarios show that the building can be evacuated within 2 minutes, even when taking into account that people can walk twice as slowly on sand as they could if they were walking on a hard floor. Key factor with regard to the evacuation are traffic jams that occur in front of each door and not the speed with which people can walk on a sandy surface.

Based upon these recommendations, the building was regarded as suitable to be used by 2800 people at any one time. The municipality is however sticking to a figure of 1536 people, because it believes and continues to believe that evacuating a building via the sandy areas is dangerous and is therefore forbidden. For some obscure reason, covering the sandy areas with boards is also not permitted, on the basis that this was contrary to the exemption that had been granted from the land-use plan. Incremental penalty payments have been imposed, due to alleged breaches of the permitted number of visitors. Permits for events attended by 2500 people have been granted in the past. This issue has been dragging on for more than 5 years. The first operator of the building actually became insolvent as a result of this whole affair and it is looking as if the second operator will share the same fate.

This case is yet a further example of a situation in which the opinion of the Fire Brigade is pitted against that of renowned experts. A variety of legal proceedings have already been conducted, but during the entire 5-year period, the administrative court has gone no further but to issue a judgment in relation to the legality of administrative issues and has not even started to look at the technical content

that forms part of the file. The municipality for its part, refuses to process the applications because the number of items of furniture are not indicated on the plans that form part of the application. The municipality is introducing spatial planning arguments, whilst the planning department that deals with such matters, asserts that the number of visitors is determined solely on the basis of fire safety regulations. During the five years in which proceedings have been underway, the operator has still not been successful in obtaining a judgment from the administrative court with regard to the legality of the position taken by the Fire Brigade and the municipality. Both operators have spent large amounts of money on legal and technical support, whilst the administrative court has not yet issued a judgment on the heart of the matter.

Case study no. 6 Daylight and approvals that are sometimes unjustified [8 till 11]

As far as ascertaining the equivalent surface area of daylight, the Buildings Decree 2003 refers to NEN 2057: 2001 Daylight openings of buildings – Determination method of the equivalent daylight area of a space, including the corrigendum C1: 2003.

Standard angles of limitation are applied:

- An angle of impedance α for a protruding obstacle of at least 25° and
- An angle of impedance β for overhangs to an angle of 30° to the left and 30° to the right, when viewed from the middle of the daylight opening.

A number of different evaluations of equivalence are currently circulating. These include evaluations based on the computer software package Andeline, SUPERLITE or the method made public in 2007 by the Equivalence Working Group of the Union of Municipalities. The question that then arises is whether or not these methods lead to the result the legislators intended. Moreover the question also arises as to whether the use of those methods means that users of buildings are content with the daylight that penetrates those buildings, or whether they start to suffer from health complaints, which is something that the lawmakers tried to prevent by implementing the legislation. This point is apparent from the foreword to NEN 2057.

If rooms have no view of the sky or virtually no view of the sky as a result of obstacles located within the same plot and therefore receive only a small amount of direct daylight, those rooms do not comply with the performance requirements of the Buildings Decree 2003, which was determined on the basis of NEN 2057. The NEN 2057 standard takes into account the fact that daylight can in fact be reflected to a certain degree and may enter the room via an indirect route. The method laid down in the computer program Andeline is based upon this.

The method with which equivalence is established is as follows. A computer model is created of the actual situation, including the obstructions that exist within the same plot. A computer model is then created of a reference situation, incorporating the same geometry and dimensions as the actual situation, but with the minimum required equivalent daylight opening that is prescribed for the room and with the minimum level of obstruction that needs to be taken into account in accordance with NEN 2057. This minimum obstruction consists of a surface in front of facade that screens the view of the sky up to an α angle of 25° and a β angle of 30° to the left and 30° to the right, when viewed from the middle of the daylight opening. If the average daylight factor in the actual situation is at least equivalent to the average daylight factor in the reference situation, we can then say that equivalence has been achieved.

In order to determine the minimum, average and maximum daylight factor, we use the SUPERLITE module of Adeline 3.0.

As far as this particular computer program is concerned, it is known that it is less accurate when calculating the daylight factor in a situation in which most of the daylight arriving in the space takes the form of reflected daylight. The reason for this is that the calculation algorithms used by Adeline assume that all reflection that occurs from a surface takes the form of Lambertian reflectance, in other words that all reflection that takes place is diffuse reflection (the light is reflected evenly in all directions). In the case of matt surfaces, this assumption is correct, but whenever glass surfaces or reflective surfaces are involved, such an assumption would be incorrect. The consequence of this is that the contribution to the indirect component of daylight is therefore underestimated, in some cases by as much as 50% (Galasiu 1998). Radiance, an advanced computer program used by TNO, carries out its calculations based upon the actual physical properties of the materials involved. What is more, SUPERLITE also has its limitations with regard to its ability to model specific situations. One example of this lies in the fact that the number of surfaces is limited (the maximum number is 100) and the surfaces must be flat and trapezoidal in shape. Radiance is not subject to limitations of this type.

Reflection factors

In situations in which the indirect component of daylight plays an important role, it is essential to select the correct reflection factors and glazing properties.

Method adopted by the Equivalence working group

The essence of this method is to take account of daylight entering from beyond the cone of 60 degrees that is applied in NEN 2057. The extent to which this gives rise the same average ingress of daylight as in a reference space has not however been considered. Nor has any account been taken of the distribution of daylight factors across the space, or of the ratio of direct and indirect light entering the

room. If the average daylight factor, the distribution of daylight factors across the space and the ratio of direct and indirect light are equal to the reference situation (based upon the application of NEN 2057), an identical daylight situation is deemed to apply. The justification of the method adopted by the Equivalence Working Group does not state this specifically, however. If that had been the case, we would have come to the conclusion that an equivalent situation did not apply, but that the actual situation was worse. The Equivalence Working Group, which has since been disbanded, determines a value in relation to the ingress of daylight, without stating which scientific sources it has used and in the full knowledge that the ingress of daylight is not established in that manner, which is why it is not taken into account as part of the NEN 2057 standard.

Analysis of calculation methods

The method used, which is based upon computer simulation, in which the ingress of daylight of a space to be built is compared with a room of the same geometry and dimensions that is compliant with the Buildings Decree 2003, is, in principle, an effective method of demonstrating equivalence, on condition that the distribution of the daylight factor across the space and the ratio of direct and indirect daylight is included in the comparison.

For that matter, if two spaces of the same geometry and dimensions have the same average daylight factor, the ingress of daylight will not necessarily be equal, as the distribution of the daylight factor across the space and the distribution between direct and indirect daylight play an important role. If the average daylight factor, the distribution of daylight factors across the space, the ratio of direct to indirect light are the same in each space, an identical daylight situation is deemed to apply.

In situations in which the ingress of daylight primarily consists of reflected daylight, Adeline, SUPERLITE is unsuitable as a means of determining the daylight factor. In other situations in which the direct component of daylight predominates, Adeline can be used highly effectively, however. The Radiance software package provides reliable results in both situations, as Radiance takes account of the actual physical properties of the materials involved. When it comes to determining the equivalence of buildings that are geometrically complex, only Radiance is sufficiently capable of generating plausible results.

The method used by the Equivalence Working Group does not enable this type of equivalent daylight status to be determined. For that reason, this cannot constitute the type of method such as the one referred to in Article 1.5 of the Buildings Decree 2003.

This case study demonstrates that the background behind the regulations is of importance, but also that anyone seeking to apply the regulations must also have knowledge of the physics involved. The

background of the regulations is not documented and in many cases, the requirement for knowledge of physics is too much to ask of those who are required to apply the regulations in practice. The result of this is that building plans are sometimes approved unjustifiably, even upon the recommendation of a working group that one would not expect to commit such an error.

Evaluation

These case studies have shown us that applying the principle of equivalence is something which, in practice, gives rise to a great deal of discussion and not infrequently to delays in the building process, endless legal proceedings and subsequent economical damage. The lesson that we can draw from all of these examples is that from a practical point of view, the administrative court is not in a position to determine which of the parties is in the right. Put simply, the administrative court lacks sufficient specialist knowledge. The problems arise on the one hand from inadequate regulations, in which the purpose and the background of the regulations are insufficiently documented and on the other hand from a lack of sufficiently in-depth knowledge of the regulations and the physical and usage models employed by regular users of the regulations.

It would seem to be desirable if, in addition to an arbitration process, it were possible to organise some form of specialist arbitration scheme whose mandate is not derived from legislation, but which is independent of both the legislature and the competent authority.

References:

1. Building Decree, Stb. 2001, 401, changed afterwards several times.
2. Decree on fire safe use of structures, Stb. 2008, 327
3. Court decision 09/767, Zwolle-Lelystad, 26 june 2009
4. Municipality decision, 2009/00-2717/83/01, 1 February 2010, Rotterdam
5. Municipality decision, BWT 23-01-0013, 12 December 2002, Amsterdam
6. Gelijkwaardige luchtkwaliteit en energiezuinigheid bij toepassing van Alusta passieve regelunit Pascalle en tijdgeregelde mechanische afzuiging (systeem Pascalle+), TNO 2005-BBE-R017, December 2005
7. Court decision Amsterdam, case number 10/704 GEMWT 01 G432
8. Galasiu, A.D. and M.R. Atif (1998) *Applicability of Daylighting Computer Modeling in Real Case Studies: comparison between Measured and Simulated Daylight Availability and Lighting Consumption*, IEA SHC Task 21: Daylight in Buildings Subtask C: Daylighting Design Tools, Report T21/C1-21/CAN/98-11.
9. NEN (1991) *NEN 2057 Daglichtopeningen van gebouwen - Verkorte bepalingmethode voor de equivalente daglichtoppervlakte van daglichtopeningen* [Daylight openings in buildings – Short method to determine the equivalent daylight surface area of daylight openings] Nederlands Normalisatie Instituut, Delft.
10. NEN (1997) *Wijzigingsblad NEN 2057/A1 Daglichtopeningen van gebouwen - Verkorte bepalingmethode voor de equivalente daglichtoppervlakte van daglichtopeningen* [Daylight openings in buildings – Short method to determine the equivalent daylight surface area of daylight openings] Nederlands Normalisatie Instituut, Delft.
11. TNO (2001) *Daglichtstudie Pakhuis Meesteren Rotterdam* TNO-rapport 2001-G&I-R020, Delft.

The validity of daylight calculations in rights to light cases

Dr Peter S Defoe PrD(BE) DipArb FRICS FCIArb MCQI CQP

Anglia Ruskin University

peter.defoe@btinternet.com

Ian Frame

Anglia Ruskin University

Ian.Frame@anglia.ac.uk

Abstract

The aim of this research was to discover whether it would be possible to justify an alternative measure of daylight sufficiency that would satisfy the requirement of a court and also be capable of being used in negotiations of compensation.

By testing the validity of the original research, which has underpinned the current methods without question since 1923 and, comparing this with alternative methods of measurement a new standard was developed which more accurately represents the value of daylight in a room. Further research and experimentation and testing confirmed the levels of illuminance required for sufficiency.

It is proposed that daylight calculation should, in future, be undertaken using a CIE overcast sky model, accounting for Lambert's formula. It is further proposed that the level of illuminance required should be equivalent to at least 25 Lux (0.5% Sky Factor) over at least half of the area of the room rather than the 1 foot-candle (10 Lux or 0.2% Sky Factor), which is currently used.

Using this new methodology, the practitioner will be better able to advise both the client and the court, where the degree of loss might be actionable.

Keywords

Easement, Prescription, Waldram, Rights to Light, Sky Factor, Illuminance, Calculation, CIE, Lamberts Rule, CIEL Diagram.

1. Introduction

This paper describes a completed research project towards the Professional Doctorate in the Built Environment at Anglia Ruskin University in which the aim was to reassess the basis upon which specialist surveyors advise both their clients and the courts when considering adequacy of daylight within a room.

The legal principles are clear in that a right to light can be acquired and that this right needs to be protected. The difficulty that has arisen is that the courts, rather than seeking to make the right absolute, have defined a standard of sufficiency, which was acceptable in the early twentieth century (Colls 1904) but now appears to bear little relationship to what is perceived to be an acceptable standard and they have accepted a methodology which has now been superseded by technology. Unfortunately, the experts who provide evidence in Court still rely upon this original methodology, despite the availability of alternative methods, in the sincere belief that this is what the Courts require as the standard of evidence and also because no-one has yet defined a more acceptable methodology, let alone one which might still be relevant in another hundred years.

What has to be appreciated and recognised when seeking to identify an appropriate alternative methodology is that the cases which reach the court are an insignificant percentage of the total number of cases being dealt with at any point in time and that the overwhelming majority of cases are dealt with through negotiated settlement of compensation. The decisions of the courts inform the processes used in establishing the likelihood of injunction and/ or the levels of compensation, the practitioners advise both the client in the first instance and the courts ultimately on the amount of daylight available, the degree of loss and the likely level of compensation that would be necessary to reduce or remove the risk of injunction or, alternatively, any changes that could be made to the design that would remove the risk.

The challenge therefore was to discover a scientific solution that will be acceptable to the courts, to assist their decision making process and will also be acceptable to other professionals in assessing the degree of loss, bearing in mind that they will have invested heavily in the technology based upon the original methodologies and will be reluctant to change unnecessarily.

From the foregoing, the four key objectives of this research can be defined as:

- To demonstrate that it is reasonable to assume a value of 500 foot-candles (5,000 Lux) for the illuminance provided by the unobstructed sky.
- To determine whether it is reasonable to state that a level of illuminance equivalent to 1 foot-candle (10 Lux) over half the area of the room is adequate for ordinary purposes.

- To test whether the Waldram diagram, in its present form, provides a realistic representation of the illuminance provided by the sky.
- To devise a better method of measurement either from existing alternatives or from a new paradigm.

It has long been accepted that the illuminance provided by the unobstructed sky, on all but the most overcast of days, will exceed 500 foot-candles or approximately 5,000 Lux. (Waldram 1925)

If the value should be significantly higher or lower then this would have a direct affect on the amount of illuminance that could be assumed to be available within a room by reference to its adequacy for ordinary purposes.

The BRE and others have undertaken much research into the amount of light available and for this part of the research it is proposed to compare the results of these previous researches to establish whether this level is achieved over a significant majority of the working year, comparable to the original assumptions.

If the required illuminance should be significantly higher or lower than 1 foot-candle then this would have a direct affect on the perceived adequacy for ordinary purposes, or if the requirement were for the whole of the room to benefit from at least this level of illuminance then this too would affect the perceived adequacy.

The jury experiment is used to assess the appropriate level of illuminance for ordinary use and these results are then related to the experiments using a physical room and theoretical models to determine whether the required levels can be achieved if at least 50% of the room receives at least that level of illuminance.

Whilst the Waldram diagram is used, most often, to determine relative change of illuminance, many situations rely upon the absolute definition of the well-lit area of a room. If the diagram is not sufficiently accurate, then the well-lit area may be misrepresented. The experiments in the physical room and using the theoretical models will compare reality with existing theory and with the revised diagram developed using an overcast sky model rather than a uniform sky model to assess whether the existing methodology approximates reality to an acceptable level or whether the new diagram is in fact more accurate.

Any new method must be usable and acceptable by the fellow professionals and ultimately by the Courts. By canvassing opinion of a focus group and careful consideration of the alternatives including the use of the overcast sky model, it is intended to identify whether a new method could

be adopted which provides both a more accurate representation of reality and, at the same time, remains easy to use as the present methodology.

2 Historical Context

A Right to Light is a negative easement which, commonly, can be acquired in one of three ways, by use since time immemorial (prescription under common law), prescription by lost modern grant, or under the Prescription Act 1832 (prescription by statute), the breach of which is considered to be a Nuisance and is governed by the law of tort, the remedies for which are Injunction and/or Damages

The basic requirements, for an easement to exist, were defined in *Re Ellenborough Park* [1956] Ch131

- There must be a dominant and servient tenement,
- The easement must accommodate the dominant tenement,
- The dominant and servient tenements must be owned by different persons; and
- The easement must be capable of forming the subject matter of a grant.

A tenement is a legal interest in the land or property such as the freehold or leasehold ownership and, in the context of Rights to Light, the owner of the dominant tenement has the right to the passage of light across the land of the owner of the servient tenement. Where the right exists, the servient owner may not cause any obstruction to the light passing over his land, to the extent that it deprives the dominant owner of sufficient light for ordinary use. In this way the benefit of the easement accommodate the dominant owner.

The definition of sufficient light is discussed later in this paper but at this point it is necessary to explain that the light has to pass through an aperture to a room forming part of the dominant tenement and not just to fall on the bare land. Thus the amount of light entering the room is what has to be measured for sufficiency.

If the tenements are owned by the same person then an easement cannot be created for the legal interest of that owner but may be created where there are other interests involved such as a leaseholder where the freeholder owns both properties but only occupies one of them.

A right to light might also be granted by one owner to another, without relying upon the forms identified above, in a legal agreement such as a deed of sale.

It is important, however, to recognise that the law evolves over time and in this respect the law dealing with 'Rights to Light' is no exception and there are many landmark cases to which practitioners refer for precedent. In addition to this, the Law Commission have previously issued their consultation paper No 186, which looked at Easements, Covenants and Profits à Prendre. This consultation process has

now concluded and the results are awaited but it is a possibility that they may lead to changes in connection with the prescriptive acquisition of a right to light.

The following are some of the relevant events/ cases in chronological order up to the present day.

Table 1: Significant Legal Milestones in Rights to Light Cases

Date	Event	Significance
1275	The Statute of Westminster 1189 Time immemorial	Established in the principle that a benefit that had been enjoyed for longer than could be remembered by a living person should be enjoyed as of right.
1832	The Prescription Act	Set the period of enjoyment as 20 years in rights to light cases.
1866	Dent v. Auction Mart Co	Scientific evidence used and judge referred to adverse affect of artificial light
1904	Colls v Home & Colonial Stores	Established principle of sufficiency according to ordinary use.
1921	Charles Semon and Company, Limited v. Bradford Corporation	Court decision consistent with Waldram methodology. Expert's advice accepted.
1927	Horton's Estate Ltd v James Beattie Ltd	Established no special measure to be used for tasks and consistent with Waldram methodology.
1930	Price v. Hilditch	Expert evidence accepted and consistent with Waldram methodology
1931	Sheffield Masonic Hall Company, Limited v. Sheffield Corporation	Judge did not like the evidence provided by Waldram
1935	Fishenden v Higgs and Hill Limited	Sky contour diagrams used. Generally followed Waldram
1954	Cory v The City of London Real Property Co	Walsh gave evidence and judge commented upon inconsistency of advice to the courts
1967	Ough v King	Court supported principle of 50/50 rule although accepted that this might change

		over time
1986	Carr-Saunders v. Dick McNeil Associates	Possible future use (sub division) considered when assessing damage but still consistent with Waldram
1994	Deakins v. Hookings	Followed Ough v King and consistent with Waldram methodology
2005	Midtown Ltd v. City of London Real Property Co Ltd	Argument that artificial light normally used was rejected
2006	Regan v Paul Properties DPF No1 Ltd and Others	Confirmed 50/50 rule as a useful guide when dealing with residential properties.

For the purposes of this paper, the Rights to Light Act 1959 has been ignored as the main effective part of this is the notional light obstruction notice.

The more recent case of, *Regan v Paul Properties*, heard on appeal late in 2006 was expected to produce a definitive assessment of what is sufficient, the well lit area remaining being estimated at 42% but in the event, the final settlement occurred outside of court and nothing further has been published but the offending building has been reduced by one bedroom.

We do however have the benefit of the summing up of Stephen Smith QC Sitting as a Deputy Judge of the Chancery Division in the lower Court in which, he recited the legal history of the use of calculations to demonstrate loss.

Until the early 1920's, the most sophisticated method of assessing the adequacy or otherwise of any daylight entering a window was to a method commonly used to assess the adequacy of the light entering a window was to measure the angle between the window sill and the top of the proposed building. This was sometimes referred as the 'cones of light' approach. If the angle was 45° or less, then it would be assumed that that there was sufficient light entering the window and if more, then there would not be enough.

During the 1920's, Percy Waldram, a well-known and self-proclaimed expert, devised a more sophisticated method of assessment, which required a knowledge of complex trigonometry and was used in this form until the late 1970's and even later by many practising surveyors. Few now continue to use the traditional methods as computers most often perform the calculation process and those who know how to do it in the traditional way are few and far between. This paper does not explore the actual methods by which computers produce their results as are subject to

copyright and there are no publicly available versions. Suffice to say that they do not employ the Waldram diagram except as a visual output. They do however still produce contour plans as will be explained later in this paper.

The Waldram method measures the value of light from the sky (not reflected light and not direct sunlight) and thus ignores the variations in the amount of light from time to time caused by differences in the cloud cover or between the seasons. (Bickford-Smith and Francis 2000)

A standard was set and the figure of 500 foot candles illumination was adopted by the National Physical Laboratory in 1928 as being the average condition of sky brightness found in towns in Great Britain over the greater part of winter days, over long periods in late autumn and early spring, over substantial but less lengthy periods in early autumn and late spring and on wet days in summer. It was felt that over these periods and, therefore, over a great part of the year, reasonable people would normally expect to have adequate light for ordinary purposes.

A little earlier than this, in "The Illuminating Engineer"(1923) Mr. Waldram had opined (supposedly following extensive fieldwork undertaken by himself and his father):

[... for ordinary purposes, comparable with clerical work, the natural illumination at which average reasonable persons would consistently grumble was that which represented 1/250 (0.4%) of the outside illumination which would fall on a window sill from an unobstructed quarter sphere of sky, of the same brightness as that of a patch of sky which illuminated the position under consideration. This grumble point is, of course, the same as 1/500 or 0.2% of the light which would fall from an unobstructed hemisphere of uniform sky onto a flat roof.]

Not long thereafter, the Commission Internationale de L'Eclairage, which met at Cambridge, resolved a number of issues, including that less than 0.2% daylight (aka 1 foot candle, '1 Lumen per square foot' or approximately 10 Lux) should be regarded as inadequate for work involving visual discrimination, the assessment being made at tabletop height (which was to be taken to be 85 cm (2ft 9 in)).

Waldram's methods of measuring light devised around this time translates the three dimensional reality of light flowing through a window into a room at tabletop height, into a two dimensional diagram. By comparing the diagrams of the extent of adequate light (i.e. 0.2%) available in a room, which are produced to represent the position before and after a proposed development, the effect of the development on the available light can be plotted over the floor area of a room on what is known as a "sky contour diagram". The area between the 'before' and 'after' contours is the area in which adequate light has been (or will be) lost because of the development.

In this way, an easement of light confers on the dominant tenement a right to a minimum level of light, but it does not protect all the light which may have been previously available. Thus it is entirely possible for a developer to construct a development, which has the effect of reducing the amount of light, which has been received by neighbouring buildings for very many years, without committing the tort of nuisance.

To this day the courts are still being advised by rights to light practitioners that the value of the light from the unobstructed sky should be assessed as 500 foot-candles of illuminance, that this value is modified in accordance with the formula provided by Waldram and others and that 0.2% of the skydome as measured at a height of 850 mm above floor level will provide sufficient light if half the room or more is lit to the same level.

When referring to the 50/50 rule, the Courts have made the assumption that 50% of the room would be 'well lit' and the remainder would not. Whether this implies, as suggested in *Cory v The City of London Real Property Co (1954)*, that the overall appearance would be of an adequately daylighted room or, as implied in *Ough v King (1967)*, that the remainder of the room might have a lesser use, is not clear.

This is important because on the one hand the so called 'grumble line' between well lit and not well lit areas would represent the overall illuminance of the whole room, not necessarily the average value, and, on the other, that the grumble line would represent the lowest value of illuminance of half the room thus ignoring the remainder of the room and assuming that only half the room needs to be useable. This latter interpretation is supported in the Judge's summary in *Regan v Paul Properties (2006)*, which makes clear that the 0.2% value represented the minimum level required for clerical work.

3. The Calculation Process

3.1 Assumptions

The calculations used to determine sufficiency are, as mentioned above, based upon a set of assumptions promulgated by Percy Waldram and others in the early part of the 20th Century. There are 6 main assumptions that have been challenged and which, if wrong, would have the potential to invalidate much of the advice given over recent years including that given in some high profile court cases.

1. The use of a Uniform Sky model
2. Ignoring the affect of Windows and Frames
3. Ignoring Internal Reflection

4. Taking measurements at 850 mm above floor level
5. Assuming that 500 Foot Candles (approximately 5000 Lux) of illuminance is available over the majority of the working day from unobstructed sky
6. Assuming that 1 Foot Candle (approximately 10 Lux) of illuminance is sufficient for ordinary needs.

3.2 *Use of the Waldram Diagram*

The physical manifestation of these assumptions is embodied in what is known as the Waldram Diagram. There are in fact several so called Waldram Diagrams including one that is used for assessing the vertical sky component for planning purposes but only one is actually used in Rights to Light cases.

4. Assessment Of Merits Of Assumptions

4.1 *The use of a Uniform Sky model*

A Uniform Sky is a theoretical but non-existent anomaly. There is no such thing as a uniform sky and never will be. Even when a sky is uniformly overcast as most people assume occurs frequently in the UK, the light from above tends to be stronger than the light from lower levels. When compared with the CIE model of overcast sky or the more recently proposed Useful Daylight Illuminance (UDI) sky, the traditional Waldram diagram misrepresents the value of light at higher levels. The Standard CIE overcast sky presents a theoretical model that values light from above at 3 times that from the horizon. $L=L_z \frac{1+2(\sin\alpha)}{3}$.

The Waldram diagram also ignores the affect of the sunpath but it has to be remembered that the exercise is, for the most part, a comparative one. That is to say it compares the amount of light as existing with that, which would be available after an obstruction is erected. It does have the advantage of eliminating all variables and thus making the calculation process somewhat easier, it is also assumed to be a ‘worst case’ evaluation as for the majority of the working year a higher level of daylight is available (Littlefair 1991).

Table 2 shows the comparison of the CIE predicted overcast sky against the assumption by Waldram of a Uniform sky model and the results produced by measurement during a single experiment at BRE Garston.

It can be seen that the experimental results, albeit a single set, are closer to the CIE model than the Uniform Sky Model assumption by Waldram.

Table 2: Comparison Uniform Sky With CIE Standard Overcast Sky

Altitude in Degrees	Test Average (Cd)	Predicted based on CIE (Cd)	Uniform Sky Model (Cd)
90	5.85	2.0	1.34
80	3.75	1.98	1.34
70	2.63	1.92	1.34
60	2.2	1.82	1.34
50	1.95	1.69	1.34
40	2.03	1.52	1.34
30	2.01	1.33	1.34
20	1.83	1.12	1.34
10	1.1	0.9	1.34
0	0.55	0.67	1.34

4.2 *Ignoring the affect of Windows and Frames*

Window glazing resists the passage of light. Common double-glazing can reduce the amount of light passing through the opening by 40%. Different glazing and dirt on glazing will also change the results as will the size of window frames which can vary from thin metal to large PVCu cross sections where a common reduction in the area of an opening is around 15 to 20%.

By ignoring the losses that occur through transmission through glazing and the fact that the window frames will reduce the area through which light can pass the actual amount of light passing through the opening appears to be over valued. Here too it might be argued that as the process is one of comparison with these factors as constants, that there is no purpose served in being more accurate. It also assists in making the calculation process that much simpler as it would be difficult to formulate a fair calculation of daylight availability that depended upon the type of glazing and the size of frame. It is a principle of law that the dominant owner cannot place an unfair burden on the servient owner and thus, if the dominant owner wished to have opaque un-cleaned double glazing in a large frame which restricted the amount of light passing into the room quite significantly, this should not be considered a reason for the proposed development to be significantly less than the situation where the dominant owner had clean windows and small window frames allowing a reasonable level of light to enter the room.

4.3 *Ignoring Internal Reflection*

It is acknowledged in most daylight calculations that internal reflectances adds to the overall ambience of a room and that whilst no direct daylight may reach the back of the room, the room itself may be still sufficiently daylit

However, internal reflection is also ignored, possibly for mathematical simplicity, although it should be noted that legal doctrine prescribes that the complainant may not make the situation worse than it would otherwise be and a dark room such as one where the walls are clad in timber will reflect less light than one where the walls are painted magnolia. Mathematically, in the process used for determining the Average Daylight Factor of a room for Planning purposes, it is noted that in some circumstances the benefit of internal reflectance when measured has an equal and opposite affect to the losses from glazing and the frame of a window.

It is possible therefore that a reasonable interpretation of the historical methodology is that these two variables over which the other party has no control (The window and the internal surfaces) are deemed to cancel each other and can therefore be ignored.

In an experiment that formed part of this research, a model under an artificial sky dome was tested with and without glazing and with the internal surfaces reflective and non-reflective. Under these conditions the central reading of daylight factor was 1.75% with glazing and internal reflectance and 1.5% without glazing and with non-reflective finishes. A small change to the window material or a slightly less reflective internal finish would bring these two results even closer together.

4.4 *Taking measurements at 850 mm above floor level*

The rationale for taking measurements at 850 mm above floor level has never been fully explained although it is often referred to as the work surface height. However most work surfaces are at around 700 to 750 mm in height.

Underlying this assumption is the fact that light striking a surface a shallow angle provides less illuminance than that from a steeper angle. This is referred to as Lamberts cosine rule, which is used to adjust value of light at lower levels according to Waldram. By taking a higher working plane for the calculations there will be a small change to the results obtained generally although this often depends upon the height of the window cill being considered as a surface close to the window and below the cill level would show an area as being unlit.

Waldram and others asserted that this adjustment was incorporated into the Waldram Diagram model but this has not been possible to justify this mathematically. In addition, to be valid, the reading must be taken at work surface height not some arbitrary height above floor level.

4.5 Assuming that 500 Foot Candles (approximately 5000 Lux) of illuminance is available over the majority of the working day from unobstructed sky

Many people have now measured the light from the sky, including the BRE (Littlefair 1993) and particularly across the south of England, and have shown that the value of the illuminance measured varies considerably across time but that, statistically, the illuminance provided exceeds 5000 Lux for more than 80% of the working year (Hunt 1979). Bearing in mind how gloomy it can get on winters evenings and the fact that at these times it is normal to use artificial lighting, it seems that this value provides a reasonable constant to be used in the calculations.

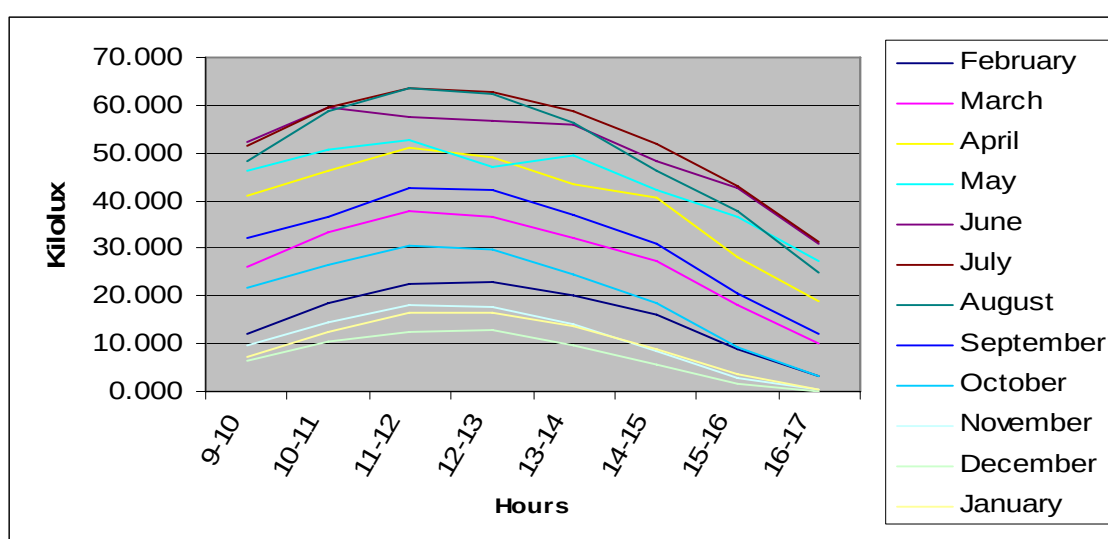


Figure 1: 5000 Lux Sky

4.6 Assuming that 1 Foot Candle (approximately 10 Lux) of illuminance is sufficient for ordinary needs.

There is no sound basis for this assumption and this is where there is the largest scope for challenge to Waldram in particular. Waldram carried out a very limited study in the first quarter of the Twentieth Century, using rooms in various offices in London where participants were asked to indicate where in the room they felt that they had sufficient daylight to work by. These responses were then charted and compared with a diagram where the predicted levels of daylight had been drawn (notwithstanding that these levels will have varied during the experiment itself). The results were then ‘interpreted’ by Waldram who had publicly acknowledged that he could work in lower light levels than his peers and resulted in his statement that an illuminance of 1 foot-candle was sufficient for ordinary needs.

It appears from available evidence that contours were plotted on each plan showing the 'daylight factor levels that correspond with the points at which the observers felt that the amount of light changed from being just sufficient to insufficient (what later became called the 'grumble line'). The results were then compared with the 'no sky line' contour for the rooms and in one classic case it shows the no sky line as being between the 0.1% and 0.25% daylight factor contours and nearer the window that at least five members of observers indicated that they had sufficient light to work by. Evidently this was achieved through internal reflectance.

The interpretation of these results is subject to serious questions but if Waldram's assertion that 0.2% of the sky dome would provide 1 foot candle of illuminance (approximately 10 Lux) is accepted then this must be compared with a controlled test undertaken by the authors, in 2007, with twice as many participants, that showed the minimum level of sufficiency for over 80% of the participants was around 25 Lux.



Figure 2: Room with 10 Lux illuminance

Figure 2 shows how a room would look if it were illuminated to the levels accepted as sufficient for ordinary purposes. In the experiment, undertaken at Anglia Ruskin University, less than 15 % of participants were able read properly at this level of illumination.

5. The Waldram Diagram

The traditional Waldram Diagram, illustrated in Figures 3, 4 and 5, commences with a series of horizontal lines representing altitude. The adjustment of spacing between these lines is, according to Waldram, based upon the value of light and the area of sky at each angle. On this is plotted a series of droop lines that represent the geometrical location of points in the hemisphere parallel to the plane of the window. Over this there is another set of droop lines representing geometrical points perpendicular to the plane of the window and over all this is the grid dividing the whole chart into 500 equal squares (each square representing 0.1% of the sky dome). A window, shown outlined in red, is plotted onto the chart illustrating how the droop lines modify the window head as it projects onto the sky dome, and within this window projection an external obstruction is plotted. The area of window unobstructed is calculated using the square grid. The light from two squares (representing 0.2% of the sky dome or 1 Foot Candle where the whole sky provides 500 Foot Candle Illuminance) was deemed sufficient by Waldram.

Traditionally a series of charts was produced for a number of grid points within a room and the values plotted on the grid. The practitioner would then interpret the results and draw a contour line representing the points where 0.2% of the sky dome should be visible based upon this data as in the example at Figure 6. It was then often the case that a further series of checking charts would be produced along the contour line to verify the result where accuracy was important (usually where the loss was deemed to be significant) as shown in Figure 7.

Even for simple situations the charts took several minutes to prepare each and series of charts would run to hours of work with no absolute guarantee of the accuracy of the result. Common errors included:

- Using the wrong Waldram Diagram
- Incorrect assessment of angles
- Incorrect datum levels
- Wrong Grid spacing
- Calculation of Contour Area
- Omitting Distant obstructions

In fact the outcome of any analysis might be subject to an accumulation of errors if the practitioner failed to check his work at each stage.

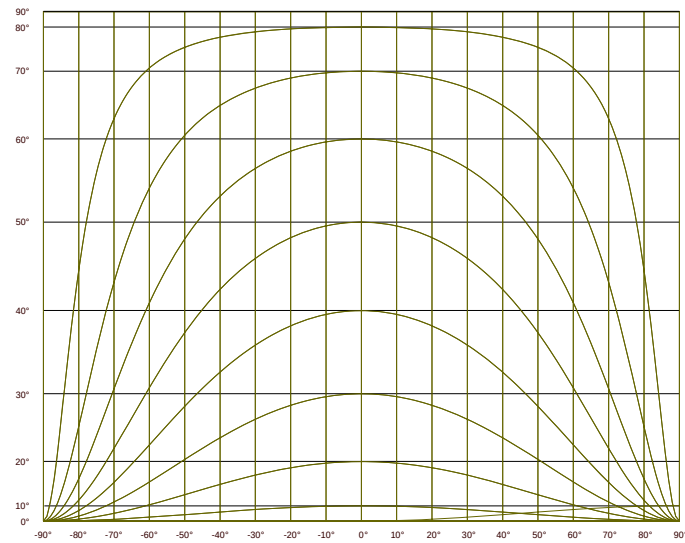


Figure 3: The Basic Waldram Diagram

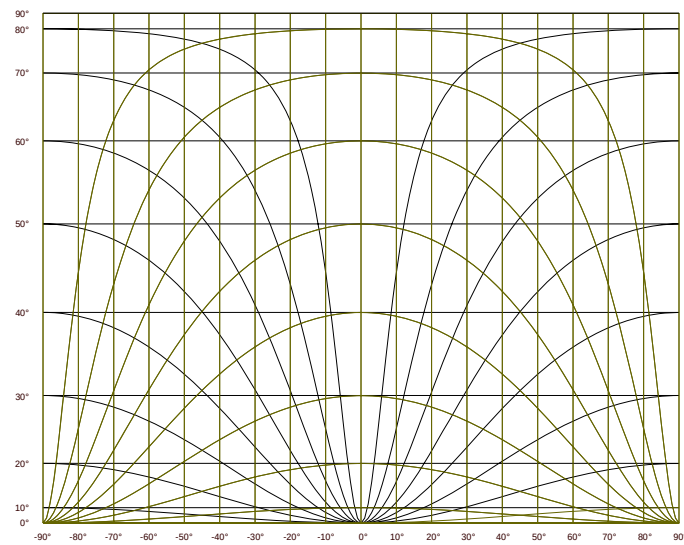


Figure 4: The Basic Waldram Diagram with lines added for obstructions at right angles

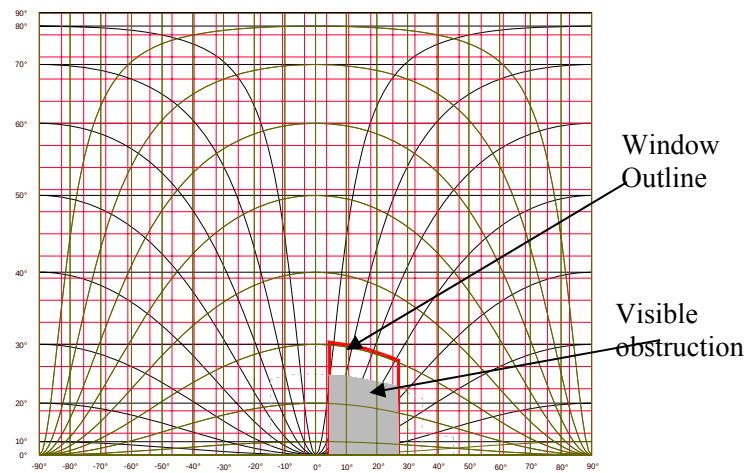


Figure 5: The full Waldram Diagram with Grid added and example of window with obstruction

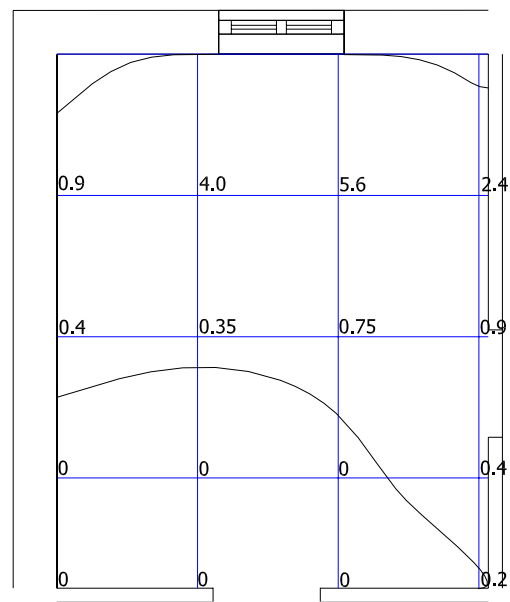


Figure 6: Grid and Typical Contour Drawing

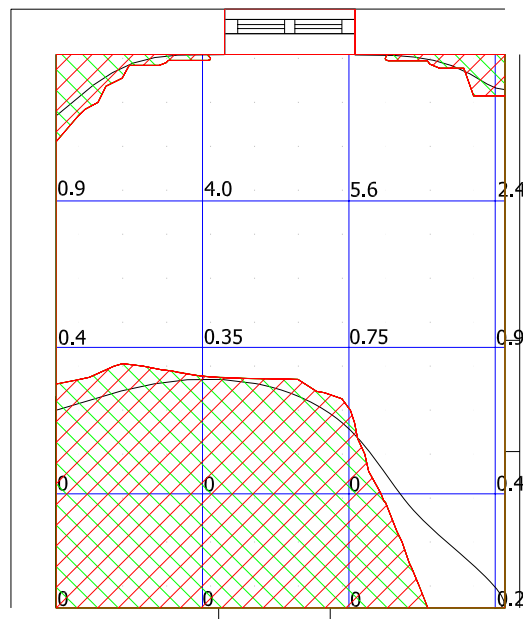


Figure 7: Corrected contour

6. The Development Of A New Diagram In Place Of The Waldram Diagram Used For Rights To Light Cases (The CIEL Diagram)

This investigation has established that the uniform sky model does not provide a satisfactory basis for establishing the amount of light available from various parts of the sky and that the non uniform CIE model is closer to reality. To create a diagram that is the equivalent of the traditional Waldram Diagram it is necessary to modify the value of light from the various altitudes in accordance with Lamberts Formula, which considers the illuminance on a horizontal plane from light at different angles.

To derive a formula that recognises a non-uniform sky and can be used in the calculation of illuminance at any point on the horizontal plane it is necessary to modify the formula used by Waldram and to achieve this it was necessary to seek assistance from the computer software providers who had created the ‘Rights to Light’ programme used in one of the experiments that formed part of this research. In the first instance, a series of questions were posed regarding the potential modifications that could be applied to the original formula and by discussion it was decided that it should be assumed that the non uniform sky should be represented by the CIE sky and then the mathematical analysis would produce the following (Manescalchi 2007):

6.1 Assu mptions

Horizontal Angle = β

Vertical Angle = α

The Differential Surface Area Element (DSAE) for a unit sphere is given by:

$$\cos(\alpha) d\alpha d\beta$$

The CIE Standard Sky transformation can be represented by $(1+2\sin(\alpha))$. The constants 0.33Lz have been dropped for mathematical simplicity.

6.2 *Skydome Modified by CIE values*

If the DSAE is adjusted by the CIE transformation the resulting formula would be:

$$\cos(\alpha) (1+2\sin(\alpha)) d\alpha d\beta.$$

6.3 *Adjusted Sky Values*

If the Waldram Diagram is to represent the luminance of the sky dome onto the x/y plane then the Differential Surface Area Element of the Waldram Diagram must be:

$$\cos(\alpha) (1+2\sin(\alpha)) d\alpha d\beta$$

The horizontal scale is proportional to $\int_0^\beta d\beta = \beta$

i.e., the horizontal scale is uniform. It is just the horizontal angle.

$\cos(\alpha) (1+2\cos(\alpha)) d\alpha$ denotes the rate of change of the area along the vertical axis.

The vertical scale is proportional to

$$\int_0^\alpha \cos(\alpha) (1+2\sin(\alpha)) d\alpha = (\sin(\alpha) - \frac{1}{2} \cos(2\alpha) + \frac{1}{2})$$

6.4 *Lambert Cosine Law + CIE Overcast Sky*

The above formula would have to be modified by Lambert's formula in order to replicate the stated basis of the Waldram diagram

If the DSAE is projected onto the x/y plane (Lambert Cosine Law), this produces:

$$\cos(\alpha) \sin(\alpha) d\alpha d\beta$$

(Note that the angle from the vertical is: $(90 - \alpha)$ and $\sin(\alpha) = \cos(90 - \alpha)$)

If the CIE Standard Sky is added then this produces $(1+2\sin(\alpha)) \cos(\alpha) \sin(\alpha) d\alpha d\beta$

If the Waldram Diagram is to represent the luminance of the skydome onto the x/y plane then the Differential Surface Area Element of the Waldram Diagram must be:

$$(1+2\sin(\alpha)) \cos(\alpha) \sin(\alpha) d\alpha d\beta$$

As before, the horizontal scale is proportional to:

$$\int_0^\beta d\beta = \beta \text{ i.e., the horizontal scale is uniform. It is just the horizontal angle.}$$

$(1+2\sin(\alpha)) \cos(\alpha) \sin(\alpha) d\alpha$ denotes the rate of change of the area along the vertical axis.

The vertical scale is proportional to:

$$\int_0^\alpha (1+2\sin(\alpha)) \cos(\alpha) \sin(\alpha) d\alpha = (-\frac{1}{2} \cos^2(\alpha) + \frac{1}{2} \sin(\alpha) - \frac{1}{6} \sin(3\alpha) + \frac{1}{2})$$

From this a modified Waldram Diagram (or CIEL Diagram) might be constructed using Microsoft Excel where $\text{ATAN}(\text{TAN}(\alpha \cdot \text{PI}/180) \cdot \text{COS}(\beta \cdot \text{PI}/180))$, provides the equivalent angle ' α ' in the CIE Lambert formula above.

Figures 8 and 9 below show the comparison between the traditional Waldram Diagram droop lines and the amended version using the above formula.

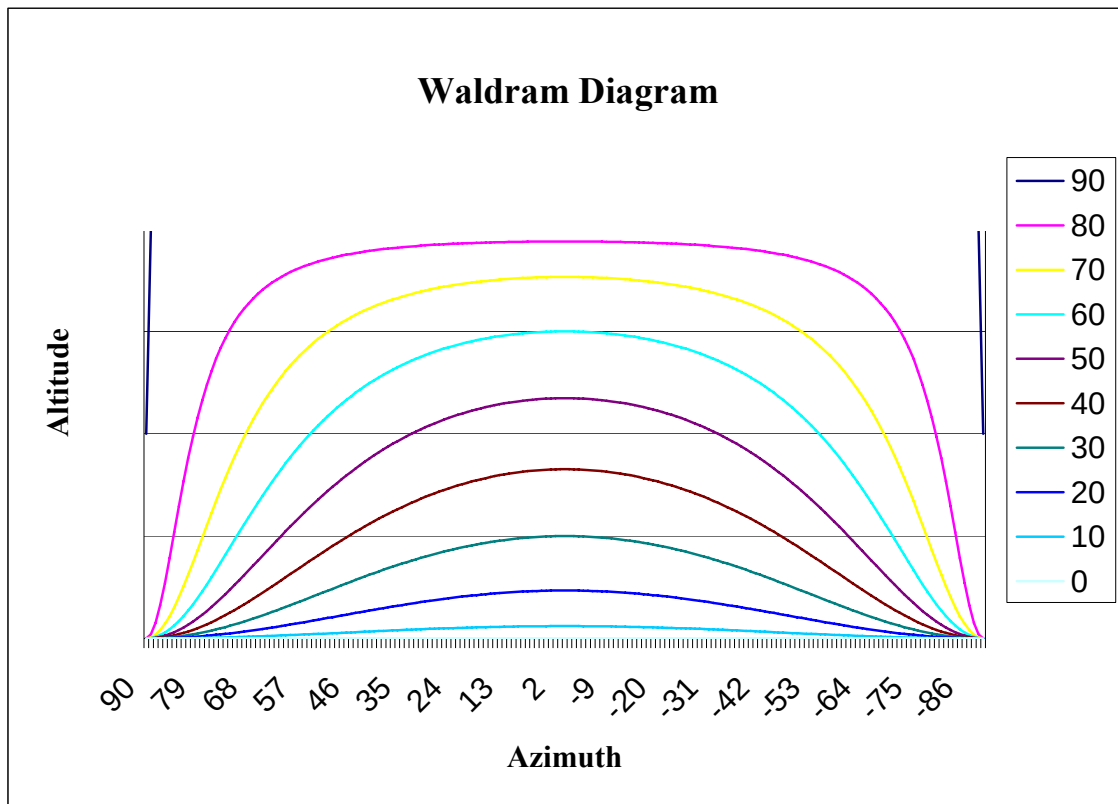


Figure 8: Waldram Diagram

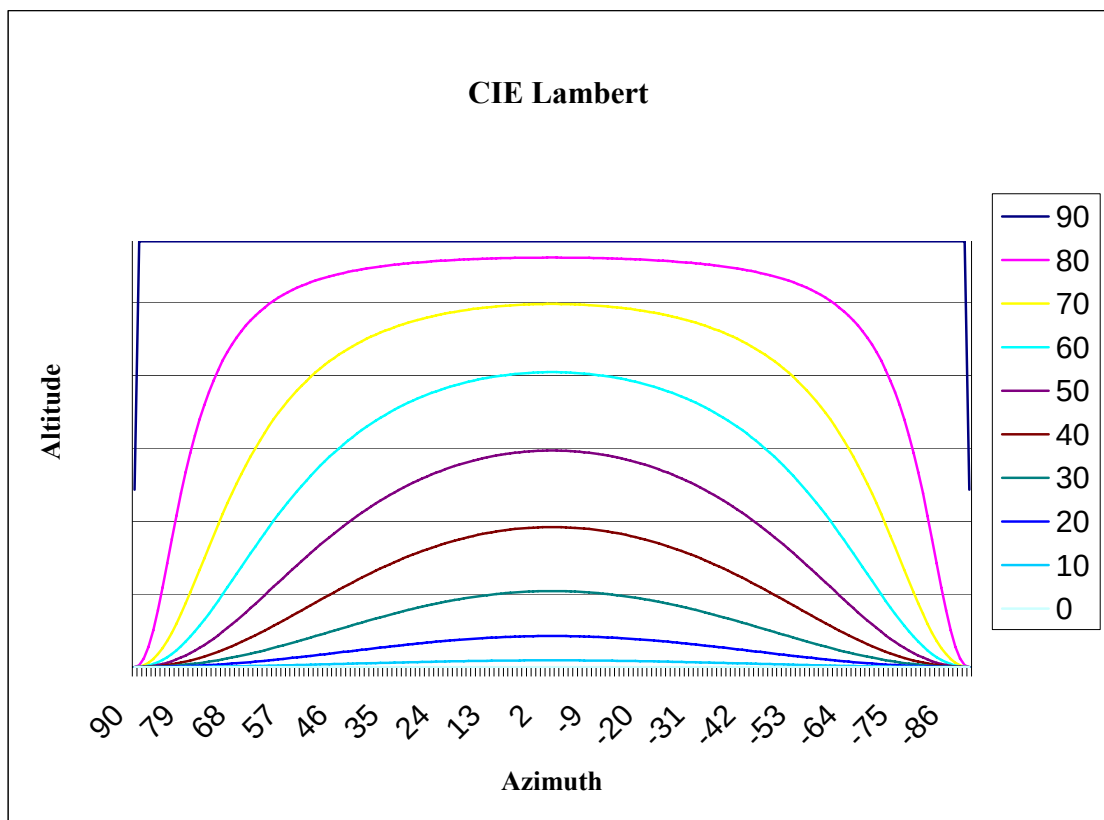


Figure 9: CIEL Diagram

A direct comparison of the two charts shows that spacing of the lines at lower altitudes is less and at higher altitudes greater, using the CIE Lambert Diagram. This means that light from higher altitudes has greater value than at lower altitudes. The significance of this is that in rights to light cases most of the light being received is from the sky above 45 degrees and thus the diagram will predict a greater amount of light entering a room than the Waldram diagram for the same point. Figures 13 and 14 below show how this relates to the results in the test room.

7. Testing of the New Diagram

A test room was provided by the BRE and readings were taken within the room. Figures 8 and 9 below show the outside of the building being tested and the obstruction opposite.

The room was also modelled electronically and physically so that actual results could be compared with those obtained through traditional means, electronically and in the artificial sky dome at Anglia Ruskin University.



Figure 10: Outside of the Building being tested



Figure 11: Obstruction Opposite Window

Figure 12 below shows the view from one point within the room, looking through the window towards the obstruction



Figure 12: View from Within Room

This shows how the obstruction looked through the window and the same projection onto the Waldram Diagram is at Figure 5 above. The amount of sky visible is at the threshold deemed by Waldram to be just sufficient. Note that the Waldram diagram ignores the window frame.

Using the Waldram Diagram results for a series of points within the room produces a contour line as shown in Figures 6 and 7 above.

7.1 Comparison of Results from Two Methods

Both methods, the Waldram Diagram and the CIEL Diagram, were then used to produce a series of contours for different values of illuminance within the room.

Figure 13 shows the results from the Waldram methodology and Figure 14 the results from the CIEL methodology.

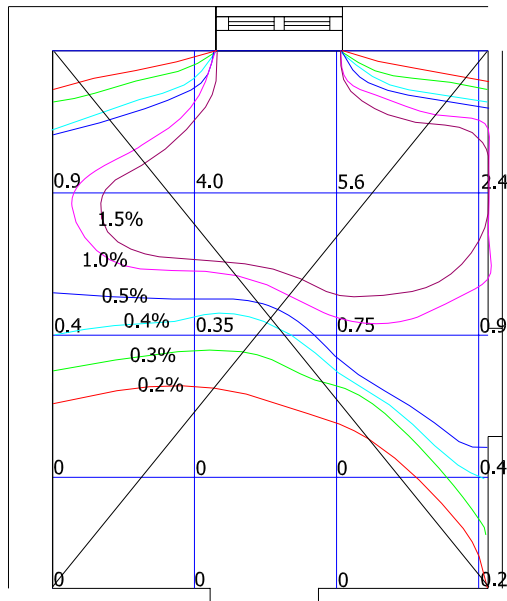


Figure 13: Contours Using Waldram

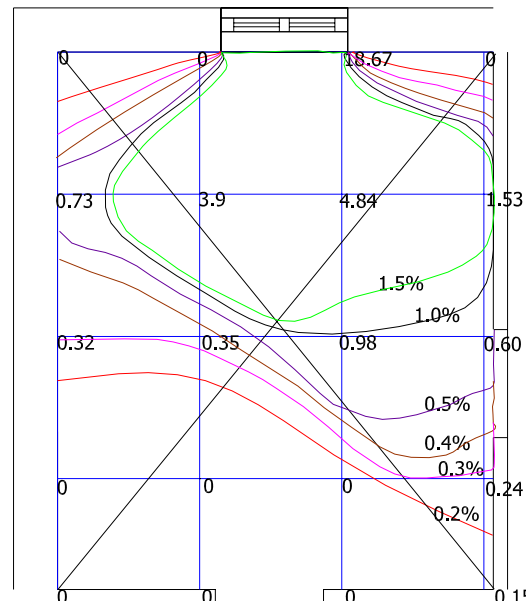


Figure 14: Contours using CIEL

7.2 Comparison with results achieved experimentally

Readings taken on site and in the artificial sky dome using a 'Megatron' meter that allows the user to determine the daylight factors at any point within the room by comparing the value of the light at that point with the unobstructed value of the light from the sky dome. Once it was clear that the model in the sky dome was producing results that were close enough to those achieved on site, the model in the sky dome was then altered so that the internal surfaces were non reflecting and the glazing was removed so that only the sky factor was being measured and the windows were not causing an obstruction..

The outcome of these results was that the value at the centre of the room, when only the sky factor was being measured, which is the closest approximation to the Waldram methodology, achieved a result of 1.5% Daylight Factor. This figure is closest to the predicted result in figure 14 above. By comparison, the Waldram Diagram only predicts a value of 0.4%. These equate, in a 5000 Lux skydome, to 75 Lux and 20 Lux respectively.

8. Consideration of results of research

There were two main areas to this research, firstly to try to establish what illuminance is actually required for normal use if it were assumed that this was to be totally supplied through natural daylight remembering that this excludes specific task lighting requirements, with the aim to demonstrate that the accepted level of 1 foot candle is insufficient. The second was to derive a methodology for determining the actual amount of light that would be received in a room whilst still providing a graphic output from which rights to light practitioners could assess loss and thus advise both their clients and the courts.

It was relatively simple, in a controlled environment, to establish that the majority of participants required rather more than 1 foot candle of illuminance for ordinary use and the fact that over 80% required approximately 2.5 times that was not a great surprise and would not of itself cause surprise amongst practitioners.

The derivation of a new methodology for measurement was rather more difficult in that first the fundamentals of the present measuring system had to be tested and in doing so challenging the assumptions that had been made by Waldram and others for reasons of mathematical simplification more than anything else.

Of most significance was the fact that the traditional methodology, failed to predict the correct value of illuminance within the test room whereas the new CIEL diagram achieved a close approximation.

9. Conclusions

Whilst it has always been known that the Waldram Diagram does not provide a true representation of reality, this research has shown that there is sufficient doubt, that it provides even a close approximation of levels of illumination within a room lit solely by daylight, to warrant a re-examination of the basis of assessment. The proposed CIEL diagram achieves, at least in the property tested, a result that is demonstrably closer to reality whilst still providing a constant basis for assessment i.e. eliminating variables that will remain constant for both sets of calculations, existing and proposed.

When Waldram's assertion, that 1 foot-candle of illuminance is sufficient, is tested under controlled conditions it is apparent, even with a very limited study, that this is not adequate for any normal working conditions without artificial light and that the level required by the majority of people is more likely to be in the region of 25 Lux or approximately 2.5 foot-candles or 0.5% of the sky dome.

Thus, if the example room illustrated above were assessed using the CIEL Diagram and the requirement for the room to be lit over 50% of its floor area to 25 Lux; or more, then it would be on the cusp of being acceptable whereas using the Waldram Diagram and 0.2% of the sky dome the room would be considered more than adequately daylit. It is interesting to note also that even using the Waldram Diagram, if a value of 0.5% is sought, the room would be considered to be only just about sufficiently daylit.

References

- Anstey, B. and Chavassee, M., 1963. *The Right to Light*. London: The Estates Gazette:
- Bickford-Smith, S. and Francis, A., 2000, *Rights of Light, The Modern Law* paragraph. 12.10:
- Building Research Board, 1944. *The lighting of buildings, The Lighting Committee of the Building Research Board of the Department of Scientific & Industrial Research* 12. London: HMSO.
- Chynoweth, P., 2004. *Progressing the rights to light debate Part 1: a review of current practice*. Structural Survey, **22**(3), pp. 131-137.
- Chynoweth, P., 2005. *Progressing the rights to light debate: Part 2: the grumble point revisited*. Structural Survey, **23**(4), pp. 251-264.
- Chynoweth, P., 2009. *Progressing the rights to light debate: Part 3: judicial attitudes to current practice*. Structural Survey, **27**(1), pp. 7-19.
- Chartered Institution of Building Services Engineers, 1999. *Daylighting and Window Design, Lighting Guide LG10*. London: Chartered Institution of Building Services Engineers.
- Commission Internationale de l'Eclairage, 2002. ISO 8995:2002(E) *Lighting of Indoor Workplaces*. Vienna: Commission Internationale de l'Eclairage:
- Daylight Illumination Panel, 1931. *Appendix I, Results of Estimates of Limiting Daylight Factor Contours for Adequate Lighting Made by Members of the Daylight Illumination Panel of the Architectural Sub-Committee of the Illumination Research Committee* (Department of Scientific and Industrial Research). Commission Internationale De L'eclairage, ed. In: CIE, 1931, CIE pp222-223-239.
- Defoe, P.S., 2008. *The Validity of Daylight Calculations in Rights to Light Cases* PrD(BE) Thesis, Anglia Ruskin University
- Defoe, P.S., and Frame I., 2007. *Was Waldram Wrong?* Structural Survey Vol 25 Number 2 p98
- Defoe P.S., 2009. *Waldram was Wrong!* Structural Survey Vol 27 Number 3 p186
- Defoe, P.S., 2008. *Light – How Much is Right?* Building Services journal CIBSE 10.08 p54
- Defoe P.S. and Frame I., 2005. *The Validity of Daylight Calculations in Rights to Light*, ed. In: 5th International Postgraduate Research Conference, Manchester: The University of Salford 948-957.
- Department Of Scientific And Industrial Research, 1937. *Illumination research technical paper*. London: HMSO.
- Hunt, D.R.G., 1979. *Availability of Daylight*. Garston. Watford: Building Research Establishment.

Illuminating Engineering Society, 1972. *Daytime Lighting in Buildings*. Technical Report No 4. London: Illuminating Engineering Society.

International Commission On Illumination, 1932. *Commission Internationale de l'Éclairage, huitième session, Cambridge, Septembre 1931 : recueil des travaux et compte rendu des séances / publié sous la direction du Bureau de la Commission*. 8th Session September 1931. Cambridge: Cambridge University Press.

Manescalchi, G., 2007. Personal correspondence regarding Mathematical Analysis of Waldram Diagram. Held at Elstree Computers Limited.

Nabil, A. And Mardaljevic, J., 2004. Useful daylight illuminance: a new paradigm for assessing daylight in buildings. In: *Lighting Research Technology* (2005), The Chartered Institute of Building Services Engineers, pp. 41-42-59.

Paterson, C.C., 1928. *Penetration of Daylight and Sunlight Into Buildings. Technical Paper No.7*. England: Department of Scientific and Industrial Research London: HMSO.

Paterson, C.C., 1935. *Seasonal Variation of Daylight Illumination. Technical Paper 17*. London: Department of Scientific and Industrial Research: HMSO.

Pitts, M., 2000. The grumble point: is it still worth the candle? *Structural Survey*, **18**(5), pp. 255-258.

Prescription Act. 1832, (c71) England and Wales. England:

Regan v Paul Properties DPF No1 Ltd and Others. EWCA (2006) Cc 1319.

Swarbrick, J., 1938. *Easements of Light* (A synopsis of modern practice and a brief explanation of simplified methods of measuring daylight and assessing compensation). London: Batsford:

Taylor, A.K., 1931. *The Daylight Illumination Required in Offices. Technical Paper No.12*. London: HMSO.

Waldram, P.J., 1925. *The Natural and Artificial Lighting of Buildings*, London: RIBA pp405-406-426.

Waldram, P.J., 1927. *Penetration of Daylight and Sunlight into Buildings. Technical Paper No.7*. London: HMSO.

Waldram, P.J., 1928. *The Estimation of Damage in Ancient Light Disputes*, pp175-176-188.

Waldram, P.J. and Waldram, J.M., 1923. Window Design and the Measurement and Predetermination of Daylight Illumination. *The Illuminating Engineer* (April-May 1923), XVI(45), pp. 86-87-122.

Waldram, P.J., 1909a. A Standard of Daylight Illumination of Interiors. *The Illuminating Engineer*, **3**, pp. 469-470-472.

Waldram, P.J., 1909b. The Measurement of Illumination; Daylight and Artificial: With Special Reference to Ancient Light Disputes. *The Journal of the Society of Architects*, **3**, pp. 131-132-140.

Walsh, J.W.T., 1922. *Elementary Principles of Lighting and Photometry*. London: Methuen.

Walsh, J.W.T., 1961. *The Science of Daylight*. London: Macdonald.

Probabilism, the way out for performance based building regulations

Dr.N.P.M.Scholten PhD

Stichting Expertisecentrum Regelgeving Bouw (Found. Expertcentre Regulations in Building)

n.scholten@bouwregelwerk.org

Abstract

Performance based regulations emerge increasingly. The actual situation in the Netherlands is described. We observe from our consultancy experience permanent obstruction of innovation and prescription of well known solutions because of a lack of understanding by local authorities. Bureaucracy and disasters force opponents and politicians to blame the complicated and scientific nature of modern democratic performance regulations. Political emphasis on reduction of the body of (technical) regulations however is counterproductive. Expert opinion is that a decrease of regulatory-burden is possible by improving regulatory methodology. Meanwhile probabilistic thinking is developing. We cannot live nor build without the risk of failure. The acceptance level of risk should be a political decision. Knowledge and understanding of building regulations should increase and be fostered by improving education. Eurocodes show quite some advance; “tight rope” calculations can be made. Application of probabilistic methods in other areas than structural design is still in its infancy or research stage. New regulations of emissions to ground, air and water have to be developed by reason of a sustainable world, but may not hinder re-use of building products. The content of a building product is not important, but the risk that dangerous emissions will damage the environment. By applying performance based principles, using probabilistic methods adequately, we could really advance. Management by incident by issuing new regulations based on one accident is unholy. When disaster strikes, politicians and the regulators usually create a new host of rules to prevent recurrence. Acceptance of effects based probabilistic judgement should be the issue. Poor regulations, conflicts in practice, huge administrative burden and quite some destruction of capital are observed in building practise. Based on our experience we see blockades for the application of innovative technical and organisational solutions, and the use of unsatisfactory solutions that only can be put to level only at great cost. Owner/user orientation ensures acceptance and proper application. Research and education should aim at tools to manage probabilistics in the building industry and assessment procedures.

Keywords: probabilism, performance, regulations,

1. Dutch situation

The building industry, which provides 9-12 % of the GNP of the Netherlands, does not have any self-guarding capacity. The market is split into a large number of sub-markets, each with its own role and dependent of the other. The building sector has no driving force capable of providing private guidance to ensure that the public is served with a built environment that is safe, healthy, useful, energy efficient and durable. The customer has almost no say in this: the building owner or user has no choice and few possibilities to influence the quality of the product he is provided with. Mainly because there is a scarcity of buildings, particularly houses, so that the customer is forced to accept what is offered. Architects and builders are unable to match their offer to the demand and certainly not to future demand. In such a situation building regulations play a vital role. Ensuring that publicly required minimum standards are met in all aspects is relevant for society.

Fig. 1 Building regulations and the public interest

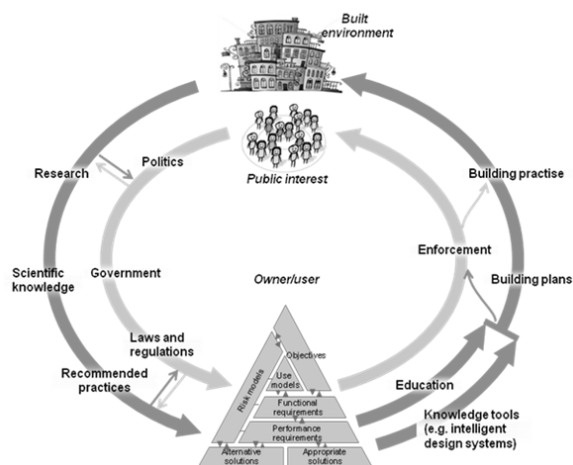


Figure 1 shows schematically the interfaces that building regulations have in the process of safeguarding the public interest. The inner circle describes the public world. The outer circle the private market from research to building owners. Building regulations is part of this world. The satisfaction of the built environment and the public interest should be the driving factors of the whole system. It should be in mind that in Holland the owner is responsible for fulfilling the

regulations. After 1995 fundamental research¹ on building regulations stopped². Technical building regulations have not been part of university and vocational training curricula for more than 20 years. In October 1992 new building legislation in the Netherlands was introduced in the form of the revised Housing Act (Woningwet), the Building Decree (Bouwbesluit) and related technical documents.

Fig.2 Relations between the documents

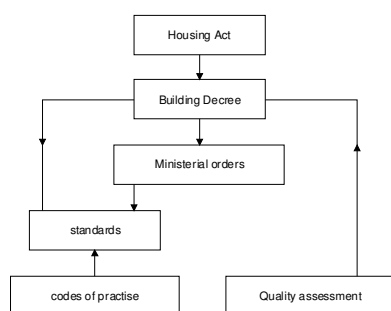


Figure 2 shows the relation between the documents under the new legislation. Next to complying to these technical requirements local authority by-laws must also be satisfied. However these are not technical. From November 2008 the requirements on fire safe use of buildings have been transferred to national level in the new legislation in the Decree on fire safe use of structures (Gebruiksbesluit).

The Building Decree, a general body of administrative regulations based on the Housing Act, is expressed in performance requirements. The performance requirement is based on a functional definition. The definition is thought to express the intention of the performance requirement. The performance requirement consists of a limit value and a determination method. The limit value is the minimum level of performance that has to be attained. The determination method is usually a Dutch Standards Institute standard, so that (elements of) these standards are also part of the Decree.

This means that the standards in the Decree must meet specific criteria and follow the same conceptual model as used in the Decree. The standards have to be performance based and have

¹ Research about scientific content en way of expression of regulations, standards e.g.

² In the PhD-study "De juridische en technische grondslagen van de bouwregelgeving – Woningwet en Bouwbesluit", May 2001, TUD, N.P.M. Scholten, Msc a complete overview of all studies is given.

objectives that are in line with the Decree. The standards' terms and definitions have to be the same as those in the Housing Act and the Building Decree. The boundary conditions in the standards should be clear and the standards should not be in conflict with government policy. In a 1985 policy paper the conditions to be fulfilled were laid down explicitly [1].

A number of large research and standardisation projects have been carried out in the Netherlands to align the standards with the Building Decree. In the period 1985-1992 the "Action plan Building Decree and standards" reformulated more than 80 standards. Between 1992 and 1995 standards were produced to be used in assessing existing building works. In 1996 and 1997 the standards were reformulated again to get complete uniformity between the standards and the Decree. Later on Ministerial orders included only few technical clauses requiring standards. The 2003 Building Decree brought about a change in terminology and modelling and required all standards to be revised in 2001 and 2002. The standards have to be in line with administrative Decree before a Decree can be published. The references to standards have to be correct, otherwise regulations can not be used in practise. So research on reformulating standards starts on a pre-Decree version.

Based on the Building Decree quality assessments (technical approvals, certificates) are thought to be an efficient way to verify that buildings and construction materials comply with the performance-based requirements. The Building Decree provides that formal quality assessments issued by accredited bodies, which are recognised by the Minister of Housing, are acceptable as sufficient proof of compliance with the requirements.

Performance requirements give building contractors and suppliers the freedom to make choices to achieve compliance. Nevertheless there is a need for practical instructions on how to achieve compliance through using current solutions. This need is met by NPR's (Dutch Codes of Practice, Nederlandse praktijkrichtlijnen), which describe the calculated or measured performances of using current solutions. The Building Decree does not refer to these NPR's, but they have been prepared on the basis of the standards referred to in the Decree. Only a few NPR's have been prepared.

In the period 1992 until end January 2010 the Building Decree has changed 27 times. Ministerial orders (in 1992 there were 5 Ministerial Orders and from 2003 only 1) have changed 35 times. A ministerial order is a document, referred to in the Building Decree, that can change qua content in time.

2. Reality

2.1 Introduction

The dissertation [Scholten 2001] [2] proposes a model to determine if existing or still-to-be-developed regulations serve their purpose.

In practice the background to the regulatory requirements and their objectives is poorly documented in the explanatory text of the Building Decree. Documentation can be found in the Expertcentre Regulations in Building library. This is a private initiative. Also the terminology and modelling of building structure in the 2003 Building Decree does not align with the building industry customs. Moreover the 2003 Building Decree is written in legal language that is barely understood³ by professionals and normal people. Forcibly the standards of the Dutch Standards Institute show the same shortcomings.

The determination methods described in the standards need to be target oriented and consequently they are written in scientific language. Furthermore they must be aligned with the underlying principles of the 2003 Building Decree. For example they must be able to accommodate flexible building plans and the principle of equal rights. In practice few people understand and correctly apply the 2003 Building Decree, Ministerial orders and associated standards. That is the experience of all teachers of courses given about parts of the regulations. Also the members of standardisation committees are not competent to formulate the content of their technical expertise in the juridical form of the regulations.

Not all the regulations are yet performance based. The Building decree still contains numerous regulations formulated in functional terms, handling those situations that are either too complex, or where insufficient knowledge exists, to specify quantifiable performance criteria. A building permit requestor or a building owner has to demonstrate that the regulations will be met and the local authority has to agree or disagree. The Court adjudicates in cases of dispute. In practice much discussion and dispute arises due to lack of clarity and precision of the Decree and limited knowledge about de content. Even where parties are in agreement, misinterpretation means that it is uncertain whether their agreement meets the regulations. The reviewing authority can have set its demands too high or has given its agreement wrongly.

³ From post graduate education experience is is clear that the students do not understand the language of the regulations.

Deleted: By

Deleted: s

Deleted: the

2.2 *Equivalent performance*

The administrative burden can increase enormously if tried and tested solutions are passed over in favour of innovative new ones. The 2003 Building Decree and the Decree on fire safe use of structures allow this if proof of equivalent performance is provided. Only if this possibility is driven by common sense, unnecessary costs will be avoided. (Needless discussions in obtaining approvals, unnecessary extra work in building modifications, avoidable costs in (re)constructing buildings for which permission was incorrectly given).

The following are examples from our consultancy:

1. Is a computer controlled ventilation system acceptable that takes into account the number of people in a room at any given time but in which the ventilation levels averaged out over daily and yearly cycles is less than legislated requirements? Is this system acceptable in assessing the energy performance coefficient (epc) of a building? Can a TNO developed computer model be used for this assessment that takes into account outside climate and the activities of the occupants? [3]
2. Can the developments in high efficiency boiler technology be taken into account, by which lower exhaust temperatures are emitted when determining the risk of fire starting in a chimney exhaust pipe? How do we ensure that an incorrect boiler is not actually installed thus creating a fire hazard? [4] and [5]
3. Can “Tritium-Light” emergency exit signs be used instead of the approved traditional ones with their specific colours? Are the standards unknowingly and unjustifiably protecting the manufacturers of traditional signs from competition? [6]
4. What in fact are the criteria for determining adequate emergency routes? What are the criteria for safe operations by the fire brigade? In which cases will it be necessary that the fire brigade comes into action? Can CFD calculations help in demonstrating the circumstances of safe escape or safe extinguishing? Are these calculations comparable to much simpler models such as one and two zone models? We refer then to restricted visibility length due to smoke, temperature and heat radiation, toxicity, etc. [7]
5. How should a structure’s fire resistance limits be determined, beyond which it collapses, if instead of using the standard fire curve the natural fire safety concept is used? How does the positive contribution of a sprinkler system or fire alarm affect the

outcome? What are the implications for the safety which is the objective of the currently accepted performance requirements? [8]

6. When evaluating historic buildings can a risk-based method be used instead of simpler models based on performance requirements? [9]

If probabilism was the highest level of regulations this question could be solved. Condition is however that we have the skills to use this way of thinking on both sides of the table (the market and the competent authorities).

2.3 Exemptions

A source of discussion is the competence that local authorities have in cases of complete or partial reconstruction or modification of existing buildings, to approve lower performances than required for new buildings. The Building Decree 2003, clause 1.11, regulates local authorities' competence and ability to do this.

Taking into account the intentions of the regulations this is only logical because buildings are built for a certain lifetime. Standards are subject to economic constraints: for a building with a short life-span, disinvestments should be avoided. Similarly what is the benefit of a 2.3 m high door in a new extension of an existing building if all the other doors in the building are only 2.1 m high? Is it necessary for this one door to accommodate an ever-taller population? This argument applies to the same door being wide enough to accommodate a wheelchair.

2.4 Education

Building regulations are not a part of the technical education curricula and this creates practical problems. The regulations should be a lot better understood at all levels in the building industry. Owners and users, architects, consultants, builders, fitters, suppliers, authorising bodies and assessors are struggling to apply the regulations correctly is the overall experience of leading experts in the field of building regulations. Leading to unnecessary discussions, unnecessary revised superfluous plans and unnecessary rejections, the cost of which is all born by the end-user.

Certifying bodies and those involved in standards meet the same problems.

The limited extra-curricular and post-doc education on the subject adversely affects also further development in the regulations.

Normally all goes well in straightforward cases. But in more complicated cases, when equivalent performance plays a role or a complex determination method is involved, or worse, when the regulations are not quantified, then everybody gets off track.

A most distressing example is the revision of the regulations in the 29th November 2009 Official Journal on safety requirements for renovations of structures that are not buildings (bridges, tunnels, etc.) [10]. The safety requirements that applied to existing buildings were considered acceptable to apply after a renovation. There was no appreciation of the risks that were thereby considered acceptable. Following the advice of the Expertcentre Regulations in Building an urgent revision to the regulations was made [11] and published.

2.5 Court cases

The situation does not improve when the legal system is used to decide disputes. The General Administrative Law Act [12] seldom leads to a judgement satisfactory based on purely technical considerations. The judge reviews the procedures that have been followed. Has the local authority decision been arrived at reasonably? In few cases does technical expertise enter the discussion to allow a more substantial consideration of the issues. Frequently an understanding of the background to the regulations is needed to judge if unacceptable risks are involved.

The following are examples:

1. Should the fire brigade in a three story artificially ventilated parking garage with a sprinkler installation at any time be able to search the garage and rescue victims close to a fire? An extra investment of € 70.000 for a sectional fire alarm system depends on the answer to this question. If the chance of victims and of a successful search by the fire brigade is sufficiently small then the extra investment may be considered disproportionate. Probabilistics is a tool that can help decide in such cases [13].
2. What should be done in the case of an empty industrial building of 12.000 m² that four squat watchers occupy from 9.00 to 17.00 hours (an artist using 40 m², an office of 40 m² with a computer table and chair, storage of 50 m² for 100 chairs and a furniture maker with 50 m²)? Should the four squat watchers be forced to vacate the building within 24 hours because there is a danger of fire spreading to an office building 5 m distant? In this case a risk-based approach can help find an answer. The administrator should also set this off against the risk of squatters occupying the building and of fire then breaking out [14].

3. What is the chance of the first floor of 66 m² collapsing in a restaurant that can accommodate 75 people in a former hayloft? The floor can withstand 3,5 kN/m² while building regulations require 5 kN/m² - but this is based on a worst-case scenario of a group of dancers in a discotheque. Here also a risk-analysis can end the discussion [15].
4. Should a care hotel located close to a gas pipeline be equipped with a 60-minute fire resistant shield if the probability of a fire occurring with a radiation impact of 22 kW/m² on the end wall is $1 \cdot 10^{-6}$ per year? What is an acceptable level of probability at which we stop investment in further safety measures? Must we be able to withstand all the dangers that could possibly face us, whatever the cost?
5. How should various forms of panic exit devices in a large discotheque be assessed?
6. Should doors be locked which serve under certain circumstances as emergency exits and under other circumstances as security barriers? For example when a building can be used as well partly (only one or a few rooms) or as a whole [16].

Only when the parties to a dispute (local authorities, fire brigades and building owners) consider their differences in terms of risks and probabilities the number of disputes will drastically reduce. This ability is present at only a few experts and they are not working for local authorities where the decisions are made.

2.6 Disasters

Incidents and accidents serve to shape regulations and determine how buildings are appraised and how the regulations are enforced. If probabilistics play a role then “management by accident” will occur much less and regulations can be far less detailed.

In the past 10 years a number of large accidents and disasters have occurred. Volendam (café fire causing 14 deaths and 180 injured), Schiphol (prison fire causing 11 deaths), a near collapse of a market square above a parking garage and inadequate structural load capacity in one of the nearby apartment buildings, partial collapse of a parking garage, collapse of a series of balconies (Fig. 3), the complete burn-down of the architectural faculty building at Delft (Fig. 4),

Fig. 3 Collapse of balconies Maastricht



Fig. 4 Burn-down of faculty building Delft



death of three firemen in a boatyard, collapse of a theatre during construction, health complaints in a new residential district due to a faulty ventilation system. In retrospect we can ask ourselves: how could this happen and how can it be avoided? Are the regulations comprehensive enough? Is it due to inadequate education? Are lessons and experiences from the past coming across to newcomers in the building industry? The lessons we have learnt from the past are not being fed into improved regulations. A limited number of studies are being undertaken to try to learn from collapsed buildings [17] and fires. Until now too little has been documented, which means that the basis for a probabilistic approach to building issues is still very small

3. Government view

3.1 General

The Government understands that revision of the building regulations is needed. Political circumstances hinder development aimed at the optimum solution. The Government is involved only to a limited extent with the technical content of its regulations. This is a result of “lean government” and the loss of expertise in the Ministries. It is also the result of the way in which communication takes place between government and the industry (The formal advisory body receives in practise one week before the advise is due the crucial documents). Feedback from the industry to government is almost non-existent. The Government is mainly concerned with a number of specific political issues: how to reduce the administrative burden for the public? De-regulation is fashionable. This has led to the new Wabo (Law of generic clauses related to regulations having an impact on the environment (Wet algemene bepalingen omgevingsrecht)[18] in which all building related permits have been combined (i.e. 25 separate legal statutes from national, provincial and local government). It is now possible to present one single permit request for all permits together, which the local authority has to quickly decide upon.

However, the technical regulations on the background are unchanged and remain as unconnected to each other as they were before. As a consequence the preparations for a building project become more risky because all the preparatory work needs to be done before the submission of the request for a building permit. This can be considered as a disadvantage. The risk of a large number of legal proceedings stopping the project is however reduced. In the mind of Government and Parliament deregulation means fewer regulations. This may have undesirable consequences for the public in an industry that needs regulations to ensure that minimum levels of safety, health, usability, energy efficiency and durability and sustainability are met. Competitive market forces do not exist and therefore cannot provide what the public needs, also with regard to the future.

Government regulators could have learned more from the lessons of 2003: subsequent to the deregulation of balconies and external storage space in housing, apartments have been built without them. This has produced disadvantaged neighbourhoods where people do not want to live but have to, due to lack of free choice. Parliamentary pressure has resulted in reintroduction of these regulations. Nevertheless the government intends still to discard 25% of the existing regulations to diminish the volume. No research has and will be done to the consequences for the built environment and peoples welfare.

The objective of course should be a reduction of *regulatory burden*. This is not the same as simply scrapping regulations. Formulating the regulations more understandable and making them less complex can achieve effectively a 25% reduction in regulatory burden.

The outcome of the current politically dominated process in the government will most probably be the following – only partly addressing the above mentioned problems:

- combining the legislation on permit requests (Wabo);
- regional pooling of knowledge for local authorities to call on as they wish; in the juridical world the Courts will believe that the regional pools advices stands for material knowledge, but without education the lack of knowledge is not solved. In reality people discuss with each other without the background knowledge of the regulations and that will give bad advice.
- regionalising the fire brigade with the same effects;
- integrating the 2003 Building Decree, local building by-laws, Decree on fire safe use of structures and technical requirements for tunnels and at the same time discarding 25% of the regulations.

The readability and complexity are not addressed, nor the education issue. The regulations will still not be tailored to the needs of their users.

3.2 *European influence*

The Netherlands is a member of the EU and responsibilities are attached to that. National regulations should not hinder free trade and this means that Dutch regulations should be aligned with European standards and quality assessment systems. Multinational companies that pay

little attention to legislative principles largely influence the content of these standards. This can lead to market disruption by formulating unnecessary high performances that are not in line with the goals of public regulation. The ERB equivalent performance declaration about Tritium Lights escape signs demonstrates such related to NEN-EN 1838 [19]. The need to introduce The European Regulation for construction products [20] forms an illustration of the disruption caused by the CPD

Introduction of the Eurocodes in the Netherlands will not help reduce the regulatory burden. On the contrary it will increase because the Eurocodes are ambiguous and a mix of technical, legal and administrative provisions. Consultation needs and discussion with the competent authorities will increase and the cost of all this will be for the owner/consumer. These negative local effects for the Dutch national market are accepted because of the possibility to freely operate in the whole EU in the advisory practise. That was the outcome of discussion between the responsible minister and the Parliament and can be read in Parliaments documentation and the documentation of the mirror standardisation committee 351 001 “TGB-Plenair” and the working group “Implementation Eurocodes”, 351 001 00 01.

Europe is also working on standards for durability and sustainability as part of implementation of the CPD [21]. Environmental regulations are aimed at pollution of air, water and ground. Regulations based on the content of materials frustrate material recycling and so cause unnecessary environmental pollution. Influential European member states do not seem to understand this. Instead of putting society’s interests up front fear of the unknown is creating excessive environmental demands; if something is not included in a product it cannot cause any harm. This way of thinking kills innovation and consequently also the environment. Where do we put all our building rubbish if we cannot recycle it? In the documentation of CEN TC 351 ‘Sustainability of construction works’ documentation about the tough discussions can be found.

4. Three recommendations

Three important Dutch publications have been made concerning rethinking building regulations:

1. the Dekker committee report: [Private if possible, public if it must] [22].
2. the Mans committee report on improving enforcement: [Now is the time] [23]
3. the report of the Expertcentre Regulations in Building [Trust and reliability] [24].

The Dekker report makes a number of recommendations of which the Minister of MHSPE is not convinced. In particular the recommendation to privatise a part of the enforcement process

(for example discontinuing preventive checks on building structures) which would lead to informed solutions. The Dekker committee assumes a beneficial working of market forces that up to now has not proven to be the case. For decades the responsibilities for those in the industry have been clear, but Government supervision is still required to ensure that the public interest is safeguarded. This report will not suddenly change this. A situation where new buildings would have to be closed immediately, because a repressive enforcement system revealed that the building regulations had not been followed, will not be acceptable and will raise political problems. That would be a waste far exceeding the cost of the present system. Authorities will not have the courage to enforce break down completely recent works.

The Mans committee concludes that most local authorities are too small to upkeep their knowledge adequate and up-to-date. Yet, we have to accept local authority autonomy. Pooling knowledge is now actively being pursued, but for the local authorities still voluntary to consult. This is a desirable objective also in the field of building regulations.

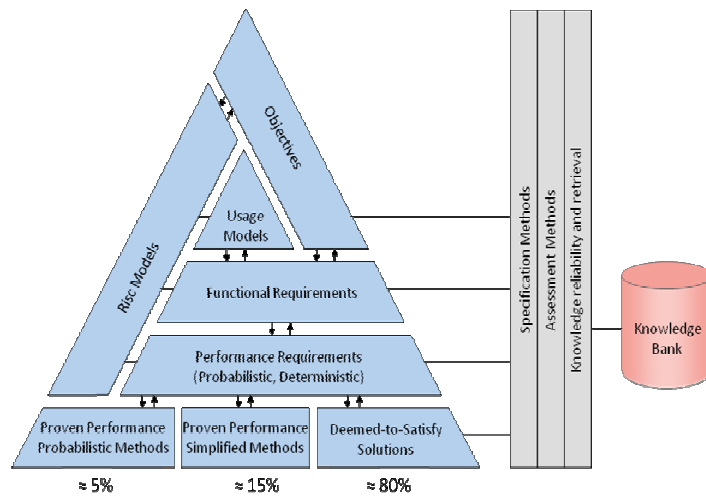
The Expertcentre report puts the user centre stage. Regulations should be tailored to the users' needs, because the user/owner is in the end responsible for fulfilling the Law. This requires a quite radical and comprehensive re-write of the regulations. Knowledge expansion and integration of regulations in technical education curricula is indispensable. The most urgent priority is to document why the regulations exist and what their objectives are.

5. The way out is probabilism

To solve the various dilemmas as sketched above, the actual regulations should, according to the opinion of the ERB-experts, be developed in three levels once the objectives are properly stated. Objectives that satisfy the just needs of owners and users. Detailed into use and risk models the regulatory provisions should be divided in probabilistic, deterministic and deemed to satisfy sets.

The top of Figure 5 is thus representing the objectives in un-quantified terms. These can be translated in functional requirements to be fulfilled.

Fig. 5 : Structure of the future building regulations



The highest level of performances should be given in a probabilistic way by formulating the acceptable risk level in the clauses of the regulation.

Structural safety is already covered in that way by the standard NEN-EN 1990 [25] by reliability-indices. Acceptable models already exist covering a number of other subjects (see summaries of research of TNO for different products and publications of ERB [26, 27]). For others research is necessary to develop appropriate models, the basis for which is risk = probability * effect. The regulations have to specify the acceptable risk of not achieving the objective. The details will depend heavily on statistics. One expects that at approximately 5% of building projects or 5% of the disputes about existing buildings will involve this kind of regulation. For 80% of building projects and existing stock we have to document how plans or existing works can be reviewed in the most straightforward way (deemed to satisfy) in deciding whether or not they meet the regulations. For the remaining 15% of building projects and the remaining part of existing stock, regulations will need to be simplified performance based (the middle between probabilistic and deemed to satisfy). Regulatory solutions are not readily available. So opportunity must be created and resources made available.

6. Research programme

Referring to figure 1, the ideal regulatory framework needs the feedback loop to be closed and a body of knowledge to be built up. The backlog in documenting the background must be quickly resolved into a memory data bank or else too much experience and knowledge will be lost, since the existing regulators are dying out. A new dedicated research programme is needed for this, Providing as well the cradle for future oriented knowledge as well for the breeding of a new generation of regulators.

References

1. MVROM, (1984), Basisfilosofie - technische voorschriften voor het bouwen, Directoraat-Generaal van de Volkshuisvesting van het Ministerie van Volkshuisvesting Ruimtelijke Ordening en Milieubeheer, Zoetermeer
2. Scholten Dr.N.P.M. (2001) [Technical and legal considerations of building regulations – Housing Act and Building Decree] PhD-study, Delft
3. Phaff Ing. J.C., (1992) Beschrijving van het programma VENCON voor woningen. Een model voor ventilatie van gebouwen en verspreiding van binnenluchtverontreinigingen. TNO rapport B-92-1217, Delft
4. Stichting Expertisecentrum Regelgeving Bouw ERB (2008) Verklaring van gelijkwaardigheid voor kunststof rookgasafvoersystemen, ROLUX PP 120, Delft
5. ERB (2009) Verklaring van gelijkwaardigheid voor kunststof rookgasafvoersystemen, CoxDens@PPS en PPS/Alu, Delft, Stichting Expertisecentrum Regelgeving Bouw
6. Varkevisser Ing J.(2009) De menselijke waarnemingsprestatie bij toepassing van “Tritium-Lights” als veiligheidssignalering, Meppel
7. Tonkelaar ir. E.S. den (2009) Rookverspreiding galerij met CFD, DGMR rapport F.2008, Arnhem, 0908.01.R001
8. Herpen ir. R.A.P. and Leur, ir. P.H.E. van de and Vrouwenfelder prof. Ir. A.C.W.M. (2009) Fysisch brandmodel – afstemming risiconiveau aan publiekrechtelijke regelgeving constructieve veiligheid, Zwolle
9. VandeVelde, prof. ir. P. (2005) Fire risk evaluation to European cultural heritage – Quantification of priorities and optimization of fire protection strategies – FIRE-TECH, EU-project
10. Vrouwenfelder prof. ir. A.C.W.M. (2009) Veiligheidseisen bij de verbouw van bestaande bruggen, Delft, TNO
11. MJUS, Regeling tot wijziging van de Regeling Bouwbesluit 2003 inzake de constructieve veiligheid van bouwwerken, geen gebouw zijnde, Stcrt. 2010, 393
12. MJUS, [The General Administrative Law Act], Stb. 1992, 315
13. Uitspraak rechtbank Arnhem in procedurenummer 09/4001 GEMET 78, 23 oktober 2009
14. Uitspraak rechtbank ‘s Gravenhage, AWB 10/161 GEMWT, 13 januari 2010
15. Uitspraak rechtbank Zwolle-Lelystad, Awb 09/767, 26 juni 2009
16. Uitspraak rechtbank Zutphen, LJN: BK 3737, 09/592 en 09/817
17. Herwijnen prof. F. (2009) Leren van instortingen – Waarom bruggen en gebouwen soms instorten en hoe dat is te voorkomen!, Bouwen met staal
18. MINJUS, Wet van 6 november 2008, houdende regels inzake een vergunningstelsel met betrekking tot activiteiten die van invloed zijn op de fysieke leefomgeving en inzake handhaving van regelingen op het gebied van de fysieke leefomgeving (Wet algemene bepalingen omgevingsrecht), Stb. 2008, 496.

19. Verklaring van gelijkwaardigheid van Tritium Lights vluchtrouteaanduidingen, 2009, opgesteld door de stichting Expertisecentrum Regelgeving Bouw, Escape Light bv, GV 90201
20. Council Directive 89/106/EEC of 21 December 1988 on the approximation of laws, regulations and administrative provisions of the Member States relating to construction products, amended by Council Directive 93/68/EEC of 22 July 1993 and REGULATION (EC) No 1882/2003 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 29 September 2003
21. Proposal for a Regulation of the European Parliament and of the Council laying down harmonised conditions for the marketing of the construction products - *Political Agreement*, accepted 25 May 2009, COUNCIL OF THE EUROPEAN UNION, Interinstitutional File: 2008/0098 (COD), 20 May 2010, Brussels
22. Commissie Dekker (2008) [Private if possible, public if it must], Den Haag
23. Commissie Mans (2008) Herziening Handhavingssstelsel VROM-regelgeving (commissie Mans), [Now is the time] , Den Haag
24. ERB (2009) Vertrouwen en Betrouwbaarheid, [Trust and Reliability], Delft , Stichting Expertisecentrum Regelgeving Bouw
25. NEN-EN 1990 (NL), Eurocode - Basis of structural design
26. Publicatiebundel 2004-2008, dr. ir. N.P.M. Scholten, Expertisecentrum Regelgeving Bouw
27. Praktijkboek gelijkwaardigheid 2009, 10 cases en 40 vragen en antwoorden, Expertisecentrum Regelgeving Bouw

Automated compliance checking using building information models

David Greenwood¹, Stephen Lockley¹ Sagar Malsane¹, and Jane Matthews¹

¹School of the Built Environment,
Northumbria University, Newcastle-upon-Tyne, NE1 8ST,
United Kingdom

Email: david.greenwood@unn.ac.uk; stephen.lockley@unn.ac.uk;
sagar.malsane@unn.ac.uk; jane.matthews@unn.ac.uk.

Abstract:

Building designs in the UK are currently checked manually against a frequently changing and increasingly complex set of building regulations. This is a major task for both designers and enforcers, often leading to ambiguity, inconsistency in assessments and delays in the overall construction process. Technical developments in Building Information Modelling (BIM) offer the potential for a new generation of software tools that can automate the checking of compliance with building codes, thus improving the efficiency of building design and procurement. To attain these efficiencies designers must change their working practices and move away from the definition of a building in multiple and disparate documents to a single coherent building model from which the documentation is generated. Theoretically, this building model could contain sufficient information to respond to interrogation at the level of building code compliance, though in practice only a percentage of the required information is normally present. This paper reviews previous research into automated code compliance, identifies the key issues for future development and examines the causes of information paucity for compliance checking in the current generation of BIM tools.

Keywords:

Automated compliance checking, BIM, UK Building Regulations.

1 Introduction

With the development of Building Information Modelling (BIM) software automated compliance checking of building designs using model checking software (MCS) is becoming a realistic prospect (see Choi and Kim, 2008). It is likely that in the near future BIMs will become important digital assets that are not only key instruments in communicating design but also in obtaining approval from statutory bodies (see, for example, Raslan and Davies, 2010). The manual checking of building designs for compliance against national codes is complex and prone to human error with significant cost implications. It is claimed that automated compliance checking would not only prove beneficial to designers but to also building certifiers, consultants, building code authorities, specification writers and builders (Tan *et al.*, 2010).

However, automated compliance checking requires the application of software tools, which are normally generic and international, to codes and regulations, which are specific and local. To date, the strategy adopted in most compliance checking initiatives has been to convert proprietary BIM models into the international standard format (IFC) and then to author bespoke compliance rules that can be executed using this model. The problem with this approach is that the international BIM tools do not populate these IFC models with the all the required data; in particular, the relationships with local codes and regulations. Thus, for reliable compliance checking, additional data must be provided by the design team as a separate activity.

As part of an ongoing funded research project the authors propose a development of this strategy by defining, within the IFC model, a domain extension for England and Wales Building Regulations using the established methodologies of the BuildingSMART Alliance (<http://www.buildingsmart.com/>). The objectives are to work with the National Building Specification organisation to identify concepts, objects and properties that are entrained in the Building Regulations (England and Wales) and explore a range of formal syntaxes for creation of the requisite rules.

2 Technological development

A situation that is analogous to the uptake of BIM is that of the development and early adoption of computer aided design (CAD) in the last decades of the previous century.

The adoption of CAD

During the 1970's and 1980's two-dimensional (2D) computer aided design was developed and deployed by the 'early adopters' amongst construction design practices. By the end of the 1990's 2D CAD was used in the majority of construction design activities and 3D design systems were available, but their uptake was limited. One of the main reasons for the resistance to the use of 3D CAD in construction was the lack of perceived benefits. Essentially, this generation of software tool was drawing-oriented, i.e. the underlying representation in the tools was graphical (in terms of lines, arcs, points, *etc.*): a door, represented in this way would therefore 'behave' like a series of graphical objects, and not a door. For example, changing the opening width of the door meant making a line or arc shorter or longer. Whilst this was acceptable in 2D CAD (being no different to traditional paper-based drawing procedure) in 3D CAD the amount of graphical change required was significantly greater. Thus, although 3D CAD brought great potential benefits (such clash detection and visualisation) the overhead of authoring the models rendered these benefits cost-ineffective. The end result was that most design practices used 3D for *presentational* purposes and disposed of the model once this was complete in favour of traditional 2D drawings.

The initial response of the CAD vendors to resolve this was *parametric object design*: rather than the user having to define the lines and arcs of the door they would be automatically generated from a set of parameters such as height and width. As these parameters were altered so were the resulting graphical representations. This approach effectively accelerated the process of authoring the graphical representation of the building, but essentially the resulting model was still a graphics-oriented model designed to be output as drawings.

A paradigm shift

Concurrently, during the 1990s computer scientists were making a paradigm shift in the way software was designed and authored. This involved moving away from a functional or procedural way of conceiving software to an object-oriented paradigm. Object-oriented (OO) programming required software engineers to construct their systems in terms of the real world objects that were involved in the problem to be solved. This shift in focus was intended to lead to more stable and maintainable software solutions that could be understood by domain experts who were intended to use the systems.

Accordingly, construction industry researchers began to adopt the OO approach to software design and these resulted in several research prototypes (see, for example, Stumpf *et al.*, 1996). By the mid 1990s, the major CAD vendors were adopting this approach however, it became apparent that building designs comprise thousands of different types of real world objects and that it was a task beyond any individual company to create computer models for all of these. There was a need for international standardisation for objects in building models and this led to the beginning of the development of the Industry Foundation Classes (IFCs). It is important to note that this initiative to develop object and product models was not unique to the Construction Industry, but part of an international initiative - the Standards for Product Data Exchange (STEP) - for all sectors of industry (see Pratt, 2001). This pan-industry initiative initially focussed on developing a modelling language (EXPRESS) and file exchange format (ISO 10303 part 21) and these are the cornerstones of the current IFC implementation. These steps enabled the next phase of development, namely, generic resources or libraries that could be shared across all sectors to avoid duplication and accelerate development.

3 Strategies and approaches to system development

Approaches to developing automated building code-checking have been reported in the literature for the last two decades (Eastman *et al.*, 2009).

3.1 Singapore (CORENET)

The BP-Expert system had been available in Singapore from as early as 1995 for checking 2D drawings with a view “to reengineer and streamline the fragmented work processes in the construction industry, so as to achieve quantum improvements in turnaround time, quality and productivity” (Evelyn and Fatt, 2004, p.1). In 2000 it was replaced by e-PlanCheck as part of the Construction and Real Estate NETWORK (CORENET) project (Sing and Zhong, 2001). CORENET was one of the first initiatives in automated code-checking, and was funded by the Singapore Ministry of National Development and carried out by the Construction and Real Estate Network (Choi & Kim, 2008). This aimed to provide an internet based electronic submission system for checking and approving building plans. Building proposals were submitted as a combination of existing 2D drawings with additional information provided in supplementary IFC-based files. The system utilised many of the convergent technologies described in the previous section (OO software design, STEP and BIM); it was considered to be ‘cutting edge’ and conceptually strong, yet there is little evidence of continuing work on the specific initiative.

The aim, as before, was to improve performance, increase coverage and check compliance of building data in an IFC format. However, while the implementation of the IFC by CAD vendors remained focused on geometry many of the requirements for compliance checking were not available. E-PlanCheck addressed this by commissioning an independent platform, FORNAX, to sit on top of the already existing EDM ModelChecker. FORNAX is an object library written in C++. Each object contains all the relevant attributes for the Singapore codes as well as the rules that apply to that object. Each object is designed to be extensible in order to cover the requirements of other countries, and as a result CORENET e-PlanCheck was used as the basis for pilot projects in Norway, New York and Australia (Khemlani, 2005). Despite ongoing attempts to implement performance based checking, reported difficulties with verifying data quality (Solihin, 2004), and its inability to support the checking of design standards throughout the different design stages of the project (Ding *et al.*, 2006) e-PlanCheck in Singapore is still the only system that is currently operational.

3.2 Norway (*Statsbygg*)

The CORENET work was developed and emulated in Norway with the ByggSok system (Haraldsen *et al.*, 2004). This is an e-Government system comprising three modules: an information system, a system for e-submission of building applications and a system for zoning proposals. Driven by the Norwegian Building and Construction industry and supported by Standards Norway and Norwegian buildingSMART it is heavily based on IFC standards. The work is ongoing and currently focussing on the issues of classification, terminology and standardising rule-checking in construction at an international level.

Building upon their e-PlanCheck pilot projects Norwegian developers (Statsbygg) have experimented with multiple systems as part of their efforts to extend the use of IFC to the entire project life cycle in support of their mandate that by 2010 all properties will use IFC based BIM (Sjøgren, 2007). The resulting systems have been piloted on real projects, with data being exchanged through a wide selection of software to suit the various stages / tasks of the project lifecycle. On the HITOS pilot, the code checking efforts have focused predominately on accessible design. Here the building model data are stored and accessed through EDM Model Server in IFC format. The accessibility rules are parameterised, mapped to their associated building objects and executed using Solibri Model Checker's Constraint Set Manager. Solibri communicates directly with building model data in IFC format, but retrieves only the objects it needs – i.e. those mapped to the accessibility rules. The rules implemented to date focus predominantly on geometrical constraints and as such the objects and parameters are supported by the IFC data models produced by current BIM packages. The Statsbygg Solibri system does not support the enhancing of these data models or the export to IFC format, and so cannot currently be used for compliance checking of attributes not supported by the current BIM vendors. The Solibri Constraint Set Manager is implemented in java and ships with a library of built-in parameterised rules which can be configured by adjusting the parameters. New rules, however, must be custom made in collaboration with the Solibri software developers and as such are not easily adapted for other software. Solibri has the benefit of powerful 3D modelling engine which, in combination with the ability to directly read IFC files, allows for clear visual reporting of rule infringements for the

user. Solibri's built-in rule library contains rules for validating a data model prior to rule checking which is useful.

3.3 *Australia (DesignCheck)*

Both the Solibri Model Checker and Express Data Manager were considered as possible platforms for automated code checking in Australia (Ding *et al.*, 2004) again focusing on accessible design regulations. The work was undertaken by Commonwealth Scientific and Industrial Research Organisation (CSIRO) and the University of Sydney and was funded by Australia's Cooperative Research Centre for Construction Innovation. EDM was eventually selected and the resulting automated code checking system – DesignCheck is currently on trial by the construction industry in Australia (Eastman *et al.*, 2009).

Design Check uses object based rules, encoded using EDM. Building data models, in IFC format, are imported into the EDM database and transformed into the Design Check internal model. The Design Check model includes building code specific information not currently implemented by BIM vendors. A mapping schema, written in ExpressX translates the building data model from IFC format into the DesignCheck schema. The strategy is similar to that of e-PlanCheck in Singapore; however, DesignCheck has the advantage of supporting the ability to check for compliance at various stages in the design process, as it has a rule schema for early and detailed design stages as well as for specification. It is therefore targeted at Architects and Designers rather than just Building Control certifiers (Ding *et al.*, 2006). As yet DesignCheck does not have the ability to view 3D models and all reports are text based.

3.4 *The United States (General Services Administration and International Code Council)*

Similar work on code-checking began in the United States around 2000, with the initial emphasis on health, safety and welfare. A major driver of BIM and validation of BIM models in the United States is the US General Services Administration (GSA). The GSA issued BIM-guidelines in late 2006 (GSA, 2006) and in 2007 proposed that all planners seeking funding for their spatial planning projects would need to produce BIM models for validation as an open standard (GSA, 2007).

The aptly named *SmartCodes* is a project driven by the International Code Council, in conjunction with AEC3 and Digital Alchemy. This project has focused largely on addressing the problem of transforming paper-based codes (of which there are thousands) into machine-interpretable rules; generally a lengthy process requiring many iterations between Building Code officials and software developers. In order to streamline this process the SmartCodes project developed a methodology for applying tags to electronic copies of Building Codes using a 'tag dictionary', or ontology (Wix *et al.*, 2008). The rules are then automatically extracted, following a strict mathematical pattern, into an IFC constraints schema. The resulting IFC constraints schema is mapped to the IFC building data model via the tag dictionary. The rules can currently be executed using either Solibri Model Checker, or AEC3 XABIO. The SmartCodes project does not support building code specific information that is not currently implemented by BIM vendors (Eastman *et al.* 2009).

4 Conclusions and implications for future work

The research to date has identified four key requirements for automated code checking

- 1) Computer programmed rules must be easily understood by Regulation authors;
- 2) The lifecycle of the rule base must be independent of software and schema updates;
- 3) All development must be compliant with Open Standards;
- 4) Consideration must be given to the industry processes of model authoring.

These will now be considered in turn, and at each stage the implications for the development of an effective automated checking system for UK building regulations will be considered, alongside what is currently available in the world of BIM systems.

4.1 Easily-understood and accessible rules

Several of the initiatives outlined above have focussed on creating rules and mapping the entities encapsulated in these rules to the international building model schema. This schema is designed to support the needs of an international user and takes little consideration of national semantics (e.g. UK practice and culture). The result is that the authors of the regulations will be required to accept mappings that lie somewhere between the concepts in UK practice and their abstract counterparts in the IFC schema. For example a water closet (WC) is a well-understood description in the UK but in the IFC schema may well be represented as a ‘sanitary flow inlet’. Rules derived from the regulations that are mapped to this IFC schema will be difficult for the authors to understand and in the long term will make maintaining the rule base complex.

Consequently, and following the FORNAX approach (see above) we intend to develop a UK-specific building model schema, embodying concepts that fit with UK custom and practice; mappings will be created between the UK schema and the IFC schema through the domain extension approach, ensuring interoperability and maintainability. It is intended that these mapping will have long term durability. Authors will define rules in terms of the UK schema ensuring comprehensibility and maintainability. Since work on the FORNAX system began there have been significant developments in the IFC schema as well as computing languages that can implement it. These developments make it timely to revisit the FORNAX approach.

4.2 Independence of software and model schema updates

The core information models in the current generation of BIM software tools are designed to support international markets and are intentionally not localised to national jurisdictions by the vendors. It is simply not economically viable for the major CAD vendors to develop multiple local flavours of their product. Rather, their strategy is for localisation to be applied by third parties to comply with specific legal and cultural requirements. Clearly, though, it is not efficient for a design practice operating in both Scotland and England for example, to have to use different BIM tools for local projects. In contrast, building regulations and codes are highly specific to a location and a culture often only applicable to small regions. The UK, for example, has three different regional implementations (shortly to become four with the separation of Welsh regulations from the English). In addition, the development cycles for regulations and software are driven by fundamentally different demands: in the case of software, by market needs and

technology changes; and in the case of regulations, by changes in the law, consultation processes, and extraordinary events. It is essential that any automated regulation-checking solution can be updated independently of the development cycle of the software tools used.

4.3 Compliance with Open Standards

This being the case, the implication that the rule-checking tool must be independent of the software tool points naturally to the use of an open standard for communication. There are several Open Standards for Building models but our research to-date has suggested that the most comprehensive, for the purpose of regulatory control, are the IFCs. Other standards such as CityGML and GBXML are targeted at a specific use and do not model the breadth of concepts required. The IFCs are also widely adopted by the major CAD vendors and are generally accepted as the standard most likely to succeed.

4.4 Consideration of the industry processes of model authoring

So far we have considered the issues of the rule checking schema, but consideration is also needed as to how building models are authored by the design team. There is currently no *single* software application that can fully populate a building model. Indeed it is arguable whether a single comprehensive and coherent building model is a realistic aspiration. At present, the working processes and issues of ownership and liability in the authoring of building design works against the creation of such a single fully-populated model (Greenwood *et al.*, 2010). The approach currently taken is to define views for data exchange between software tools and organisations. Currently the most widely used and agreed view is the ‘Project Coordination View’; others are under development. Essentially, the purpose of the view is to set out formally what information must be present in the building model for it to be considered fit for purpose. The difficulty is that views can only contain entities and properties that are currently defined in the IFC schema. Clearly there are some entities and properties that are specific to UK custom and practice that would never be defined for international usage. Whilst there is a mechanism to work around this problem in terms of data exchange (in the form of IfcPropertySets) there are unlikely to be any widely used software tools that will help all members of the design team author these data.

The overall implication of this is that one of the widely used international CAD tools will be used to author the majority of the building model (elements, materials, geometry etc) and extension programmes will be needed to author data that are not provided.

Part of the current research programme will be directed at investigating optimal approaches to creating these extension programmes. Two strategies are under consideration for identifying data that would be required for automated code-checking but would not normally be present in the model. *Constraint-based* analysis executes the compliance rules on the building model and reports when insufficient information is available to determine compliance. *Schema-based analysis* compares the structure of the building model data with the structure of the UK Building Regulations Schema to determine the required additional data inputs. It is hoped that from this analysis a data entry tool can be generated that is specific to the given building model. Thus the approach would be extensible to any building model, model view or localised schema and allows flexibility in the building model authoring process.

5 Acknowledgements

The authors gratefully acknowledge the co-funding provided for this research by the Royal Institute of British Architects Enterprises Ltd.

6 References

- Choi, J., and Kim, I. (2008) 'An Approach to Share Architectural Drawing Information and Document Information for Automated Code Checking System', *Tsinghua Science & Technology*, **13** (1), pp 171-178.
- Ding, L., Drogemuller, R., Jupp, J., Rosenman, M. and Gero, J. (2004) 'Automated Code Checking' *CRC CI International Conference*, Gold Coast, Australia. Available at: <http://en.scientificcommons.org/50758591> (viewed: 04 June 2010)
- Ding, L., Drogemuller, R., Rosenman, M., Marchant, D and Gero, J. (2006) 'Automating code checking for building designs'. *The Second International Conference - Clients Driving Innovation: Moving Ideas into Practice*, The Cooperative Research Centre (CRC) for Construction Innovations, Gold Coast, Australia, March 2006. Available at: http://web.arch.usyd.edu.au/~john/publications/2006/06Ding_etalCRC06.pdf (viewed: 04 June 2010)
- Eastman, C, Lee, J-m, Jeong, Y-s and Lee, J-k (2009) 'Automatic rule-based checking of building designs' *Automation in Construction*, **18** (8) Pages 1011-1033.
- Evelyn, T. & Fatt, C. (2004) 'Building Smart - A strategy for implementing BIM solution in Singapore'. Available at: http://www.itsc.org.sg/pdf/5_BIM.pdf (viewed: 02 June 2010).
- Greenwood, D., Lewis, S. Lockley, S. (2010) 'Contractual Issues in the Total Use of Building Information Modelling', in Barrett, P., Amaratunga, D., Haigh, R., Keraminiyage, K. & Pathirage, C. (eds) *Proceedings of CIB 2010 World Congress 'Building a Better World'*. Salford, May 2010. The University of Salford.
- Haraldsen, M, Stray, TD, Päivärinta, T, and Sein, MK (2004) 'Developing e-Government Portals: From Life-Events through Genres to Requirements' in Rolland, KHR (ed.), *Proceedings of 11th Norwegian Conference on Information Systems*, Stavanger, pp.44-70.
- Khemlani, L (2005). 'CORENET e-PlanCheck: Singapore's Automated Code Checking System'. AECbytes, October, 2005 [Online]. Available at: <http://www.aecbytes.com/buildingthefuture/2005/CORENETePlanCheck.html> (viewed: 02 June 2010).
- Pratt, MJ (2001) 'Introduction to ISO 10303—the STEP Standard for Product Data Exchange', *Journal of Computing and Information Science in Engineering* **1** (1) 102-4.
- Raslan, R. and Davies, M (2010) 'An analysis of industry capability for the implementation of a software-based compliance approach for the UK Building Regulations 2006', *Building Services Engineering Research and Technology*, **31**:141
- Sing, T. F. & Zhong, Q. (2001) 'Construction and Real Estate NETWORK (CORENET)', *Facilities*, **19**, pp. 419-428.
- Sjøgren, J. (2007). 'BuildingSMART –a smart way for implementation of standards' [Online]. Available at: <http://www.plcs->

- resources.org/papers/s1000d/day3/buildingSMART%20-%20Jons%20Sjogren.pdf (viewed: 29 June 2010)
- Solihin, W. (2004) 'Lessons learned from experience of code-checking implementation in Singapore - Success, Challenges, and Future Outlook' [Online]. Available at: http://www.buildingsmart.com/files/u1/ned_20from_20experience_20of_20code_checking.pdf (viewed: 02 June 2010).
- Stumpf, R., Ganeshan, R., Chin, S., and Liu, L.Y. (1996) 'Object-oriented model for integrating construction product and process information', *Journal of Computing in Civil Engineering*, **10** (3), pp. 204-212.
- Tan, X., Hammad, A., and Fazio, P. (2010) 'Automated Code Compliance Checking for Building Envelope Design', *Journal of Computing in Civil Engineering*, **24** (2), pp. 203-211.
- United States General Services Administration (2006) '02 – GSA BIM Guide for Spatial Program Validation Appendices, GSA BIM Guide Series 02' [Online]. Available at: http://www.gsa.gov/graphics/pbs/GSA_BIM_02_Appendix_v09.pdf (viewed: 08 June 2010).
- United States General Services Administration (2007) 'BIM Guide for Spatial Program Validation, GSA BIM Guide Series 02' [Online]. Available at: http://www.gsa.gov/graphics/pbs/BIM_Guide_Series_02_v096.pdf (viewed: 08 June 2010).
- Wix, J. Nisbet, N. and Liebich, T. (2008). 'Using constraints to validate and check building information models'. *EWork and EBuisness in Architecture, Engineering and Construction: ECPPM 2008*. London, pp. 467- 475.

Uncertainty of Building Regulation for Refurbishment Projects in Malaysia

Azlan-Shah Ali ¹, Adi-Irfan Che-Ani ² and Wan-Zahari Wan-Yusoff ³

¹Department of Building Surveying, Faculty of Built Environment,
University of Malaya, 50603 Kuala Lumpur, Malaysia

Email: asafab@um.edu.my

² Department of Architecture, Faculty of Engineering and Built Environment,
Universiti Kebangsaan Malaysia, 43600 Bangi, Selangor, Malaysia

Email: adiirfan@gmail.com

³ Department of Construction and Real Estate Management, Faculty of Technology
Management, University of Tun Hussein Onn, Batu Pahat, Johor Malaysia

E-mail: zahari@uthm.edu.my

Abstract:

The uncertainty of refurbishment projects is reflected in the difficulty in getting accurate design information during the design process. One of the factors contributing to the uncertainty in refurbishment projects is building legislation. The changing and updating some of the building regulations by government has also affected the approval process in refurbishment projects, especially projects related to conservation. The statutory requirements could cause project delays and cost overruns due to adjustments that need to be made to design in order to comply with the regulations. Therefore, the main objectives of this paper are to present the difficulties that contributed to the uncertainty of building legislation and to show how it affects the overall performance of refurbishment projects. Quantitative research techniques are used, which consists of review of literature and a postal questionnaire survey that involved 234 respondents. From 234 questionnaires sent out, 62 questionnaires were found to be suitable to form a database for analysis. Descriptive and inferential statistics are used in the data analysis. The results conclude that complexity in refurbishment projects in Malaysia is made worse by uncertainty of building legislation. The associative test indicates that performance of refurbishment projects suffered from the uncertainty of building regulations.

Keywords:

Uncertainty, Building-Regulation, Refurbishment, Malaysia.

1 Introduction

Refurbishment work includes upgrading, alteration, extension and renovation to existing buildings to improve facilities and building lifespan. This work does not include routine maintenance works such as daily cleaning works, daily building inspection and monitoring (Ali and Rahmat, 2009; Ali, 2009). In Malaysia, there are several reasons why refurbishment has become a popular sector. The increased number of ageing buildings and rapidly changing technology in Malaysia construction demands building to be altered to accommodate with current building used (Ali et al., 2009). Table 1 shows the increase number in refurbishment projects in Malaysia since 2007.

Table 1: Total Refurbishment Projects 2007- 2009

Source: Malaysia, CIDB (2009)

Types of Refurbishment Works	Number of Projects		
	2007	2008	2009
Upgrading	448 54	8 60	8
Expansion	351 28	2 47	1
Repair	268 24	9 25	7
Renovation	263 28	7 21	5
TOTAL 1	330	1336	1551

Although refurbishment work is an important activity in Malaysia, it is not easy to manage. Ali et al. (2008) highlighted that refurbishment characteristics are complex compared to newly built construction which involved various aspects such as technical, technology, legislative, ecological, social and comfort. This requires proper planning in order to complete the project on time and to meet the client's requirements (Rahmat and Ali, 2010).

In addition, in building regulation, refurbishment design is one of sectors that are affected by the complexities of legislation. Changing and updating building regulations by government have also affected approval process in refurbishment projects, especially projects related to conservation. Highfield (2000) pointed out that some of refurbishment schemes need to comply with current building regulations. Refurbishments for heritage buildings have to face regulations that are more stringent and must comply with the listed building requirements, which limits the extent of alteration work allowed. The statutory requirements could cause project delays and cost overruns due to adjustments that need to be made to design in order to comply with the regulations. Hence, this paper will review the effect of building regulations towards the performance of refurbishment projects in Malaysia.

2 Uncertainty of Statutory Requirements

Design in construction projects needs approval from the appropriate authorities before it can be implemented. Holm (2000) pointed out that construction is one of the sector affected by the complexities of legislation requirements. In refurbishment projects, there

is also a need to comply with statutory requirements (Highfield, 2000). However, only certain types of refurbishment projects, which involve change of use, alteration of facade and historical building are subject more broadly to statutory requirements (CIRIA, 1994). In addition, the requirements for refurbishment of listed building are more stringent and need to be handled with sensitivity by the designers. Fire, thermal and acoustic requirements also usually affect the refurbishment schemes.

Time taken by authorities especially for the issuance of design approval is uncertain and difficult to predict. This could affect the schedule and progress of refurbishment projects. Kincaid (2003) pointed out that the majority of design participants such as architects, engineers, contractors and developers generally agree that planning, building regulation, fire and site approval are the largest obstacle to the progress of refurbishment projects. Similarly, a case study carried out by Mitropoulos and Howell (2002) also found that the main reason for delay in refurbishment projects was the process of getting approval from the local authority. When delay in refurbishment projects occurred, the overall cost for the project would also be affected. McKim et al. (2000) mentioned that one of the factors contributing to cost and schedule overruns is the regulatory requirements. In some cases, the drawings had to be submitted more than once due to amendment that needed to be incorporated. This is sometimes due to the uncertainty of requirements set by the respective authorities. Manavazhi and Xunzhi (2001) stated that one of the reasons that lead to reworking in design is changes in statutory regulation.

The Malaysia Town and Country Planning act, Act 172 (1976) clause 19 under the planning and control section states that for refurbishment work that does not involve change of usage, change of building facade, addition to building height or area and anything that does not conflict with the local plan it is not necessary to obtain approval from the Town Planning Department of the local council. The clause indicates that not all refurbishment work is involved with planning requirement. Only refurbishment projects that do not obtain fall under the abovementioned category would have to abide with the Act. However, the Uniform Building bye Laws (1984) extensively covered for building design work normally coupled with some other requirements set by the local authorities. Each local authority has its own special-requirements that need to be followed by the architects to get approval for a submitted plan. Architects are often unsure of the requirements set, especially when it is the first time they have been involved in the submission of a plan to the respective local authorities. Listed buildings are covered under different planning and building rules and regulations. The relatedness of an existing building to the new rules and regulations determine to what extent a building would allowed to be altered. In some circumstances, to comply with current regulations would require major alterations and would affect the viability of the project. In such situations, it might be wiser to carry out only a refurbishment and maintain the original appearance of the building.

To overcome uncertainty in statutory requirements, Ling (2002) suggested that appropriate attributes of designers are required. Designers who are knowledgeable and enthusiastic about their work could obtain statutory approval speedily. Knowledgeable designers would bring their knowledge in legislation and regulation to acceptable practice.

The review of literature has indicated that the uncertainty of statutory requirements lengthens the time needed to obtain approval. Therefore, it requires knowledge and enthusiasm on the part of the designers to avoid any unnecessary delay for design approval.

3 Research Methodology

This paper takes both a quantitative and qualitative approach to research. Semi-structured interviews were implemented for qualitative part and postal questionnaire surveys were used for quantitative data collection. In order to get a high response rate, the questionnaires were short and simple and did not take much time for respondents to answer. The respondents in this study were designers and architects who are directly involved in getting approval from respective local authorities. A set of questionnaires were sent to the final list of 234 respondents. After filtering the responses, 62 questionnaires were found to be useful for analysis, giving a response rate of around 30 percent. The responses represent 62 different refurbishment projects that the minimum contract value is RM 500,000. The demographic profile of the respondents is shown in Figure 1. The profile shows more than two-thirds of the respondents were principal architects with more than 10 years working experience. This indicates that the data collected is reliable.

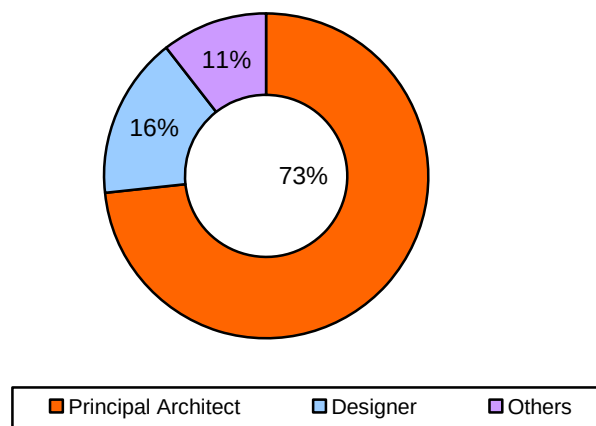


Figure 1: Job Title of the Respondents (n=62)

4 Result and Discussion

The result shown in Table 2 suggests that most of the refurbishment projects did not face much difficulty due to uncertainty of regulations set by the appropriate authorities.

Almost 70 percent of respondents agreed that requirements for refurbishment projects were certain. The architects did not have much problem dealing with the local authorities. Regulations for refurbishment conservation are more stringent. It is probable that the results show few problems regarding regulations because refurbishment projects in the present study did not consist of many conservation projects.

Table 2: The Degree of Certainty of Statutory Requirements

Statutory Requirement	Percentage (n=62)
Very uncertain	2.8
Uncertain	9.7
Neutral	20.8
Certain	25.0
Very certain	41.7
Total	100.0

However, in the semi-structured interviews, ten principal architects in Kuala Lumpur revealed that the majority of local authorities in the country had their special requirements for refurbishment in addition to the general requirements of building byelaws. The special requirements normally were not explicit like the standard byelaws. Normally, the architects had to liaise closely with the building department in the local authority regarding the additional requirements that needed to be fulfilled. As a result, the local councils rejected some of the drawings submitted for approval because they did not comply with the special requirements. It happened to ten architects who were not familiar with the particular local authority. The inconsistency of supplementary requirements of local authorities caused problems for the architects. Moreover, the time taken by the authority to give approval to the design was quite long, and that sometimes caused a delay in the commencement of a project. The interviews also revealed that refurbishment projects that did not involve changes of function, façade of the building, additional floor area or additional height of building were not complicated in fulfilling the requirements. The majority of refurbishment projects in the present study probably did not include these kinds of changes since few problems were faced by the architects. It is likely that the Malaysian Town and Country Planning Act, Act 172 (1976) and Uniform Building Bye Laws (1984) are not extensively applicable in refurbishment projects in this country except for the fire department requirements. Most architects only had to submit their design for fire department approval. The semi-structured interview also highlighted the fact that most of the statutory requirements were related with services information. It could be that the scope of work for most of the refurbishment works involved alteration for fire services system such as sprinkler pipes, smoke spill fan, fire detection system, HVAC and electrical system that needed approval from the fire department.

In some cases, the regulations might be relaxed and there could be an advantage in adapting a building so that it could be classified under current regulations. For instance, in a case of refurbishment projects of the *Takaful Nasional* office-building tower in Kuala Lumpur, it was better to leave it under existing regulations, as only minor refurbishments were required to be carried out. The work was not greatly affected by the current regulations because no alterations were made to the building's façade. However, in the conversion of shop-houses to a commercial shopping complex at *Plaza Warisan* Kuala Lumpur, it was found that regulations concerning car-parking requirements had to be changed. This worked in favour of increasing the overall number of users. A careful analysis of the plot ratio of the existing building to see if it could be maximized under current regulation could be financially rewarding.

4.1 The Effect of Statutory Requirements on Refurbishment Design Performance.

Table 3 shows that the Spearman's rank correlation coefficient detected a significant correlation between the statutory requirements variable and the changes of design variable. Statutory requirements include plan approval from the appropriate bodies, especially the relevant local authority. Problem is that special requirements for local authorities are not uniform, are inconsistent and are frequently changed. This prolongs the process of plan approval.

Table 3: The Correlation Matrix between Project Variables and Design Performance

Project variable	Changes of design during the construction stage	Provisional sum to contract value	Time variance	Cost variance
Statutory requirements	-.257*	-.150 -.	.030	-.094

* Correlation at 5% significance level

** Correlation at 1% significance level

The inconsistency of special requirements between different local authorities influences the amount of changes in refurbishment design. Manavazhi and Xunzhi (2001) discovered a similar problem: they found revisions in design were a result of inconsistency in statutory regulations.

The result implies the need for designers to have good relationships with personnel in the local authority in order to avoid breakdowns in communication, which can prevent the required information flow. Statutory requirements normally involve approval from the local authority. This needs initiative on the part of the architects to coordinate with the authority's personnel to find out any special requirements that need to be fulfilled. More initiative is needed from the designers to ascertain any new requirements implemented by the relevant authorities. The processing of plans for approval could be expedited if all the authorities' requirements were fulfilled at the outset.

Semi-structured interviews show that handling statutory requirements is not an easy task. It needs positive architects' attributes to manage the situation since statutory requirements mostly involve dealing with approval from government agencies. Approval of the submitted plan depends on information obtained from the other key participants. Requirements such as structural and sewerage system are the province of the C&S engineer; fire requirements, power supply and HVAC are the province of the M&E engineer etc. As a result, architects need more meetings and contacts for architects to collect information required for plan approval. Inability of the architects to coordinate with other designers could affect the progress of the refurbishment project by delaying design approval, as mentioned by Kincaid (2003), Mitropoulos and Howell (2002). Second, the result implies that it is important to maintain a good relationship with the authority's personnel. It is easier for architects to check the status of a submitted plan, to get information on any amendment or special requirements and frequently they are able to expedite the plan approval process.

5 Conclusion

In conclusion, the results show that uncertainty of building regulations for refurbishment projects influence the design performance. This indicates that the performance of refurbishment projects suffered from uncertainty depending upon the nature of projects. However, the semi-structured interviews show that the problem could be minimized by having good relationship between designer and authority personnel, in addition to positive initiatives from the designer's.

6 References

- Ali, A.S. & Ismail Rahmat (2009). Coordination methods in managing the design process of refurbishment projects. *Journal of Building Appraisal* Palgrave, UK. pp 87-98
- Ali, A.S. (2009). Cost decision-making in building maintenance practices in Malaysia. *Journal of Facilities Management* Vol. 7 (4), Emerald, UK pp 298-306
- Ali, A.S., Kamaruzzaman, S.N. & Salleh, H. (2009). The Characteristics of Refurbishment Projects in Malaysia. *Journal of Facilities*, Emerald, UK. Vol. 27 No. 1/2 pp 56-65
- Ali, A.S., Rahmat, I. & Hassan, H. (2008). Involvement of key design participants in refurbishment design process, *Journal of Facilities*. Emerald, UK. Vol. 26 No. 9/10 pp 389-400
- CIDB, (2009). *Construction quarterly statistical bulletin- third quarter 2009*. CIDB, Malaysia.
- CIRIA, Construction Industry Research & Association, (1994). *A guide to management of building refurbishment, CIRIA report no 133*. Construction Industry Research and Association, UK.
- Highfield, D. (2000). *Refurbishment and upgrading of buildings*. E&FN Spon, London and New York
- Holm, M.G. (2000). Service management in housing refurbishment: a theoretical approach. *Journal of Construction Management and Economics*, Vol. 18, 525-533.
- Kincaid D. (2003) *Adapting buildings for changing uses, guidelines for change of use refurbishment*. Spon Press London.
- Ling, Y.Y. (2002). Model for predicting performance of architects and engineers. *Journal of Construction Engineering and Management*, Vol. 128, No. 5, 446-455.
- Malaysia, Town Planning Act 1972 (1976). International Law Book Services.
- Malaysia, Uniform Building Bye Laws (1984). International Law Book Services.
- Manavazhi, M.R. & Xunzhi, Z. (2001). Productivity oriented analysis of design revisions. *Journal of Construction Management and Economics*, Vol. 19, pp 379-391.
- McKim, R. Tarek, H. & Attalla, M. (2000). Project Performance Control in Reconstruction Project. *Journal of Construction Engineering and Management*, Vol. 126, No. 2, pp 137-141.

- Mitropoulos, & Howell, (2002). Renovation projects: Design process problems and improvement mechanisms. *Journal of Management in Engineering*, Vol. 18, No. 4, pp 179-185.
- Rahmat, I. & Ali, A.S. (2010). The involvement of the key participants in the production of project plans and the planning performance of refurbishment projects. *Journal of Building Appraisal*, Palgrave, UK pp 273-288

A Study of Building Contractors' Compliance with National Building Regulations in Cape Town

Abimbola Windapo
University of Cape Town
Abimbola.Windapo@uct.ac.za

Keith Cattell
University of Cape Town
Keith.Cattell@uct.ac.za

1. Abstract

In South Africa a number of building regulations and other laws exist that control or influence the operating environment of the construction industry. Compliance with these regulations and laws is a statutory requirement in the construction of buildings. This study investigated: the level of awareness amongst Cape Town contractors of building regulations; the degree of compliance by contractors with building regulations; the perceived quality and availability of regulation documentation to contractors; and whether there are significant differences between the levels of compliance by the different levels of building contractors studied. An exploratory survey was undertaken of site managers/agents working for small, medium and large-scale building contracting firms on buildings under construction in Cape Town. Descriptive statistical analysis revealed that a significant number of the site managers surveyed held no formal qualification and that the larger registered contractors tended to comply more with the building regulations than the unregistered small-scale firms.

Keywords: Building Regulations, Compliance, Knowledge and Quality Standards

2. Introduction and background

Building regulations stipulate how a building should be constructed, altered or extended to ensure safe and healthy accommodation and the conservation of energy (Brett, 1997; Noam, 1983; Burbyet al., 2000). Watermeyer and Milford (2004) describe building regulations as legal instruments intended to ensure that buildings perform in such a way as to provide essentially equivalent, socially acceptable levels of health, safety, welfare and amenities for their occupants and for the communities in which they are located. Building regulations have evolved through the ages and are now becoming more complex and exacting (Baicheet *al.*, 2006). Van de Heijden *et al.* (2007) and Baicheet *al.* (2006) note that many players in the building sector are critical of the complexity of building regulations.

The fact that a variety of buildings play a substantial role in environmental health, human welfare and economic stability, has made it essential to have official rules, regulatory controls and standards to govern their design, construction and operation; and to cover such diverse areas as structural stability, fire safety, heating, lighting, ventilation, plumbing, sanitary facilities, indoor air quality and

sustainability (Watermeyer and Milford, 2004; and Circo, 2008). In the context of the latter point, it should be noted, for example, that the housing sector in South Africa is responsible for just less than 25% of the total energy demand (Watermeyer and Milford, 2004).

Most countries have laws to establish technical standards and requirements, which provide detailed descriptions of the minimum standards to ensure, among other things, health, safety and energy efficiency. In addition they typically publish official documents to regulate such aspects as standards, approved solutions and administrative procedures (Visscher and Meijer, 2007). The South African National Building Regulations and Standards Act 103 of 1977 (as amended) publishes the building regulations to be used countrywide, while the Construction Standards Department within the South African Bureau of Standards (SABS) develops and maintains national standards in respect of both products and practices for the Civil Engineering and Built Environment sectors (CIDB, 2007; Landman, 2002 and Freeman, 1990).

According to CIDB (2007), while South Africa has a well-developed set of national standards that enable manufacturers and contractors to provide consumers with high quality products, concerns have been raised about increasing non-compliance of materials and products with national standards (CIDB, 2007). Many projects are poorly specified; artisans and foremen are not accredited in terms of their performance in achieving necessary standards; many of the current specifications are perceived to present a barrier to entry of small-scale entrepreneurs and exclude their participation in particular markets; a limited amount of clients are reportedly not requiring materials to comply with South African Bureau of Standards (SABS); and there is a lack of capacity amongst building inspectors to evaluate compliance requirements. As a result, owners of buildings and infrastructure are reportedly not always satisfied with the end product CIDB (2007).

Non-compliance with enacted building regulations and by-laws has also been reported and observed in Cape Town. Examples include: residents of a Cape Town suburb protested against a developer, the Community Housing Company's (CTCHC) poor workmanship, where it was alleged that the houses built by the developer were damp and mouldy and tended to flood in winter (Booyesen, 2010); building developments had proceeded without approved building plans, contravening the National Building Regulations (Petersen, 2010); and illegal construction was allegedly added to a major building development (Fischer, 2010).

However, reliable evidence indicating the nature and extent of non-compliance with building regulations in South Africa is undocumented. This together with the complaints by end-users and the lack of satisfaction in the building product alluded to by CIDB (2007) formed the basis for this research.

2. Research aim and questions

Imrie (2004) notes that the building control system has rarely been the subject of enquiry and debate and that there are a few, if any scholarly writings about the subject and its substance, or content about its regulatory functions other than in the most descriptive of senses. The research therefore aimed to investigate the extent of non-compliance with building regulations in Cape Town, through a study of building contractors in the city. The following research questions were addressed:

- a. Are building contractors aware of the National Building Regulations and Standards Act 103 of 1977 (as amended), the municipal building by-laws and other legislation that regulate building construction? If they are aware of these statutory building laws, which ones do they know of?
- b. Do building contractors perceive the quality and availability of building regulation documentation to be adequate?
- c. What is the level of comprehension of the building regulation documentation amongst building contractors?
- d. What is the level of compliance of building contractors with building regulations and by-laws?
- e. Are there differences between the degree of compliance with building regulations and by-laws by building contractors according to their level of registration with the Construction Industry Development Board (CIDB)?

The importance of the study lies in its intention to inform the development of appropriate quality policies and management procedures for the enforcement of building regulations. If appropriate regulations and policies are developed, all parties involved in compliance and enforcement of building regulations will be able to do so more efficiently.

4. Building Regulations and Registration of Contractors

There is no reason to doubt that wherever building regulations exist, there is non-compliance and that efforts to manage this problem include the registration of contractors. However, little readily available literature could be sourced to show the extent of this, so it should not be inferred from the references used in this section that the problem is confined to the countries about which they report.

The State is a potential resource or threat to every industry in the society, the construction industry being no exception (Stigler, 1971). As a regulator, the state establishes technical standards and social norms, permits construction activities and establishes liabilities (Winch and Campagnac, 1995). *Anget al.* (2005) observed that State regulations are the outcome of technical, legal and scientific efforts, strongly influenced by political, governmental and social forces, and evaluated against the consequences of the building process, the building cost and the costs-in-use. In most countries, the enactment of legislation in the building and housing sectors stemmed from the need by government to improve public health and derive from previous experience of natural disasters (Van de Heijden *et al.*, 2007; Meijer *et al.*, 2002; Winch and Campagnac, 1995; Meacham *et al.*, 2005 and Spence, 2004).

The desirability of having uniform building regulations in South Africa, to replace the variety of local building regulations and by-laws has long been recognised (Freeman, 1990). The idea was first mooted in 1947 by the SABS, which proposed the formation of a set of building regulations to be mandatory throughout South Africa. An important change in policy came about with the promulgation in 1977 of the National Building Regulations and Standards Act (Act 103 of 1977), which provided legislative authority for the compulsory introduction of National Building Regulations which, when promulgated, superseded all existing municipal building regulations and building by-laws (Freeman, 1990). In 1985 the National Building Regulations, mandatory throughout South Africa, replaced the building regulations and by-laws of individual local authorities. In 1988, these regulations were withdrawn and replaced by an entirely new set of National Building Regulations.

Whilst the 1985 Regulations made extensive use of ‘functional’ regulations, supplemented by so-called ‘deemed-to-satisfy requirements’, which were set out in Schedules to the Regulations, the 1988 Regulations took this approach much further. Except for regulations of an administrative nature, the 1988 Regulations are now almost entirely ‘functional’ and the ‘deemed-to satisfy requirements’ have been removed. These requirements are now embodied in a separate Code of Practice document, entitled SANS 10400 “The Application of the National Building Regulations”. This is a non-statutory document, which contains technical information needed for the practical application of the Regulations.

Apart from the Building Regulations there are other legislative measures that impact directly and indirectly on the operating environment of the construction industry. These include: Local Authority By-laws; Standards (local and ISO); Codes of Practice; the Occupational Health and Safety Act (No. 85 of 1993); and the Physical Planning Act, 1967 (Act No. 88 of 1967). In addition, Housing Associations exist that enforce rules, regulations and controls such as: ensuring compliance to aesthetic controls and architectural style; building restrictions; and land use control (Landman, 2002).

Most countries, including South Africa, require certain processes to be followed to show compliance with these regulations. Such processes include requiring the client or builder to submit building plans, health and safety plans and surface water drainage plans to the Local Authority for approval. The Local Authority has technical requirements and procedures for checking building plans against the requirements and issues building permits in this respect (Meijer *et al.*, 2002). Like countries in the European Union (Vischer and Meijer, 2007; Loet *et al.*, 2002) South Africa requires building permits to be obtained before construction work can begin on site.

In 1997 Hong Kong's Buildings Ordinance provided for the establishment of a Contractors Registration Committee. Among its main functions, the committee examines the qualifications of applicants and, if it considers it necessary, inquires whether an applicant has the relevant experience (Lo *et al.*, 2002). South Africa has also instituted a similar system in which contractors are required to register with the Construction Industry Development Board (cidb), which grades firms according to a scale from 1 to 9 for public tenders. In addition, the National Home Builders Registration Council (NHBRC), established in terms of the Housing Consumer Protection Measures Act, 95 of 1998, protects the interest of consumers and regulates the home building industry. The NHBRC is mandated via the Act to assist homebuilders, through both training and inspections, to achieve and maintain satisfactory home building technical standards. Section 10 of the Act requires all contractors who build houses, whether government-subsided or for the private sector, to register with the NHBRC and comply with the requirements of the NHBRC Home Building Manual (Van Wyk, 2004).

5. Knowledge/Awareness of the Building Regulations

It is likely that studies about the knowledge and awareness by contractors of building regulations, and the effect of this on compliance, have been carried out in many countries. However, this section refers only to findings of studies that are documented and readily available.

The Australian Joint Select Committee on the Quality of Buildings (JSCQB) was established in 2002 to investigate and report on the quality of buildings (Parliament of New South Wales, 2010), in response to the decline among builders in the knowledge of the building code and the consequent effect of this, i.e. poor quality buildings.

Baicheet *al.*'s (2006) study into the compliance with building regulations in England and Wales revealed that big construction firms tend to have technical and design sections that obviously know the building regulations, while small firms often tend to rely on drawings given to them and have a poor understanding of the building regulations. Baicheet *al.* (2006) noted that the growing

complexity of the building regulations is making it increasingly difficult for small firms to understand them, and hence, errors on plans and drawings are eventually replicated in buildings. Imrie's (2004) study into how knowledgeable builders were of the provisions of Part M (Access to and Use of Buildings) of the UK Building Regulations contradicted this view. Imrie (2004) noted that, although Part M is easy to achieve because it does not ask for a great deal, conflict situations arise because builders are unaware of the requirements of the approved documents. Moreover, some building control officers claimed that, although builders are given copies of design guides on access, they rarely look at them.

Baicheet *al.* (2006) also found that most site managers were not trained for this role. However, the most professional site managers include many who have no qualifications at all, but have achieved this status solely on their experience, which suggests that holding a qualification may not affect whether an individual is aware of, or knows the building regulations.

6. Indicators of Compliance

Strong enforcement in general, and the frequency of inspections in particular, are critical elements in obtaining compliance with building regulations (Burby *et al.*, 2000). The key to obtaining compliance is the level of effort building departments expend on enforcement (surveillance, plan checking, field inspection and technical assistance).

Compliance is tested in terms of the principal goal of any building regulation, which is to establish a baseline standard with respect to a particular aspect of design or construction (Bell and Lowe, 2000). In the 'performance approach', proposed solutions can be evaluated by testing or calculation, or by a combination of testing and calculation (Glaeser and Gyourko, 2003). Bordasset *al.* (2001) are, however, critical of this approach. They comment that there has been much interest in performance indicators in many fields, but, while recognising the importance of such monitoring and benchmarking, the indicators can become ends in themselves, rather than aids to understanding and assessing the contribution to performance in different contexts.

In most countries, some requirements are given by law and in building regulations, building standards etc., but others are provided in contract documents, site meeting records, and other project documentation, as well as other details that are difficult to specify (Josephson and Hammarlund, 1999).

According to Visscher and Meijer (2007), writing in the context of European building regulations, a final inspection certificate or Certificate of Occupancy is required prior to occupation in most

countries to declare that the structure has been built in accordance with the regulations. Bengé (1999) describes that compliance with the mandatory clauses of the New Zealand Building Code is tested through a performance-based process. However, an alternative route to proving compliance with the mandatory clauses is provided, through the ‘Verification Method’ (calculations or test methods) and the ‘Acceptable Solution’ (cookbook method), which two additional tiers are collectively called the ‘Approved Documents’.

In South Africa, the National Building Regulations and Standards Act (Act 103 Of 1977) also requires a Certificate of Occupancy to be issued for conventionally constructed buildings and lists certain additional regulations, *etc.* with which contractors must comply.

7. Non-Compliance with regulations

The results of a survey of houses under construction in England and Wales by Baicheet *al.* (2006) revealed the following typical areas of non-compliance by construction companies:

- shortcomings in the construction of cavity walls – insufficient number of and poorly installed wall ties, excessive debris in cavity trays and wall ties, and faulty building-in of insulation layer
- poor integration of lintels and steel beams and junctions between walls, windows and external doors
- incorrect installation of wall and roof restraining straps
- omission of damp proof course at thresholds
- incorrect (too high) level of landscaping and paving around completed buildings;
- damaged underground drainage pipe work
- poorly installed roof insulation

Baicheet *al.* (2006) also questioned occupants of houses that had completed within the two years prior to the survey. The following faults and deficiencies were reported:

- faults and cracks in walls (most common)
- damp patches in walls
- water leaks from pipes, and blocked drains
- unstable or uneven floors
- draughty doors and windows
- malfunctioning heating systems
- faulty electrical systems

Baicheet *al.* (2006) included building control inspectors in the survey and concluded that these faults were generally the result of poor workmanship and a consequence of non-compliance with building regulations during the construction of the houses. The building control inspectors reported that the greatest area of non-compliance problems occurred at the entrances and approaches to the dwellings. This, as noted above, is covered by Part M of the UK Building Regulations, which, according to the inspectors is clearly prescriptive and hence leaves little scope for differences of interpretation. Building inspectors cited lack of knowledge of building regulations by tradesmen as the reason for problems of non-compliance and stressed the importance of competence in the running of a building site to minimise non-compliance. The inspectors also reported that the majority of site managers were unqualified.

8. Determinants of the Quality and Availability of Building Regulation Documentation

Anget *al.* (2005) report that users continue to complain about the difficulty of determining the applicable regulation in a given situation. They note that regulatory information remains a problem in the Netherlands as a result of the complexity of performance-based regulations and the lack of time available for practitioners to study and internalise them. Imrie (2004) describes a similar situation in the UK. In his review of the UK Building Regulations he notes that, although the interpretation and application of Part M are guided by an ‘approved document’, building inspectors view it as very limited and woolly.

Gannet *al.* (1998) note that the approved documents form the regulatory framework within which the building sector operates. These approved documents are widely used because practitioners prefer the simplicity of a published solution representing satisfactory current practice and because the ‘functional requirements’ are found to be often too vague for builders to create independent criteria without the risk of rejection by local authorities (Gann *et al.* 1998). Baicheet *al.*’s (2006) study reported building inspectors in England and Wales as agreeing that, generally, none of the sections of the UK Building Regulations was too impractical or difficult to put into practice. Inspectors reportedly had no particular difficulties in enforcing the regulations and believed them to be specifically written and clearly codified.

Findings from an evaluation of experience with the 2003 building regulations published by the Dutch Architects Association (V2BO Consult, 2003, cited by Anget *al.*, 2005:114) and research carried out by Listokin and Hattis (2004) reveal the problems with the quality and availability of building regulation documentation to include the following:

- “accessibility of the expanding body of regulations”

- “an increasing number of specialized staff is needed to assist the designers in the field of regulations”
- “causing certain design decisions to be taken earlier than the client can accommodate”
- “new terminology such as ‘type of functional use’, ‘class of degree of occupation’ and ‘common elements’, create interpretation problems”
- “create insecurity and doubt about whether well-proven historical solutions are still acceptable”
- “requires questionable improvements”
- “restricts cost – saving materials and technology”
- “impedes scale and efficient production”
- “limits competition among developers and professionals working in a given location and thus increase cost i.e. ‘cartel effect’”
- “poorly written and disorganized – takes more time to comprehend and use the regulations”

9. Research Approach

The research was an exploratory survey of randomly selected on-going construction projects in Cape Town, South Africa. The main criteria for selection was that the project was in the first stage of construction, i.e. site establishment and foundation excavation.

Eleven sites were identified after a two-week search in the Cape Town area. Registration and employee details of these firms are presented in Table 1, below:

Table 1 Details of firms managing sites studied

No. of Firms	Registration	No. of employees
3	cidb Grade 4,5, or 6	20-100
2	cidb Grade 9	> 400
4	unregistered	unknown
2	NHBRC	unknown

The projects were observed serially for a period of three months. Semi-structured interviews were conducted with the site managers, recording their: experiences; knowledge of building regulations; perceptions about the quality and complexity thereof; understanding thereof; and perceptions about where improvements were considered necessary.

The site inspection and the recording of areas of non-compliance with the NBR was done using Baichee *et al.*'s (2006) survey protocol, comprising a list of aspects of building regulations to be checked for non-compliance. Rigorous inspections, based on a checklist of NBR requirements, were

undertaken. Non-compliance with the regulations was considered serious if the deviations were difficult or impossible to rectify at later stages in the work without serious expense and disruption, or even rebuilding. Recording of non-compliance was mainly photographic, with some descriptive sentences. Inspections were limited to a visual assessment of the work and workmanship relating to compliance with the NBR.

10. Research Findings and Discussion

The findings of the study are presented and discussed under the following sub-headings:

10.1 Site Managers/Agents Qualifications and Years of Experience

Table 2, below, presents the site managers'/agents' qualifications and years of experience.

Table 2 Distribution of Respondents by Site Managers/Agents Qualifications and Years of Experience

Qualification	Years of Experience				Total
	Less than 5	5 – 10	10 – 25	More than 25	
None	2		1	2	5
Qualified Tradesman (Plumber, Carpenter, Cabinet maker, etc.)				2	2
N1 Foreman Qualification				1	1
Community Development Certificate		1			1
Degree in Construction Management / Civil Engineering			2		2
Total	2	1	2	5	11

Table 2 shows that five of the site managers/agents used on the sites had no formal qualifications. Two of these unqualified respondents had more than 25 years experience, one had between 10 and 25 years experience, and the remaining two had less than five years experience. Two of the respondents were qualified tradesmen with more than 25 years experience in the industry, one had an N1 Foreman qualification and more than 25 years experience and one had a Community Development Certificate and between 5 and 10 years experience. Only two of the site managers/agents were degree holders. They were employees of the two construction companies registered at Grade 9 level on the cidb register.

10.2 Knowledge of National Building Regulations and Standards Act (Act 103 of 1977) and other legislation regulating construction

Two site managers/agents surveyed admitted that they were unaware of the existence of any legislation regulating building construction. Both these respondents had no formal construction qualifications and both had less than 5 years of industry experience.

All other respondents were aware of the building regulations and other legislation controlling building construction. The regulations, statutes, etc. that most of the respondents claimed to have knowledge of includes:

- the National Building Regulations and Standards Act (Act 103 of 1977)
- the Occupational Health and Safety Act (No. 85 of 1993)
- the NHBRC Home Building Manual
- City of Cape Town's "Model Land Use Planning, Development Conditions and Requirements"
- Estate enforced building regulations/guidelines for those constructing buildings within estates.

10.3 Perception of the Quality and Availability of Building Regulation Documentation

Only seven of the respondents surveyed disclosed their perceptions about the quality and availability of building regulation documentation. One respondent regarded the quality of regulation documentation to be 'high', four perceived it as 'good' and two felt it was 'poor'. As to the availability of such documentation, five respondents considered it to be 'easily available', one reported that it was 'available' and one considered it 'difficult to access'.

One of the respondents who regarded the quality of building regulation documentation as 'good', believed it to be 'too strict' and believed that it increases building costs. Another respondent perceived the legal sections, prohibitions, exemptions, appeals and the approval process as one of the most complex sections of the National Building Regulations and that felt that improvements would be required in these sections.

10.4 Level of Compliance with National Building Regulations and other Laws and the areas of Highest Compliance and Contravention

Table 2, below, presents the level of compliance of the building contractors to statutory building laws and the areas of highest compliance/contravention.

Table 3 Level of Compliance with National Building Regulations and other Laws and the areas of Highest Compliance and Contravention

Level of Compliance	Areas of Non-Compliance						Total
	None	Materials Storage	Health & Safety	DPM	Fencing & Hoarding	Cavity Wall Construction	
Very Good	2						2
Good	3						3
Non-compliance in 1 instance		1	1	1		1	4
Ditto, in 2 instances		1				1	2
Ditto, in 3 instances		1			1	1	3
Total	5	3	1	1	1	3	14

Key:

Very Good = All areas relating to building regulations and other statutory laws are being strictly followed and in some instances have gone beyond the minimum requirements.

Good = Non-compliance with building regulations could not be identified by visual inspection.

DPM = Damp proof membrane

It can be inferred from Table 2 that only five of the contractors were seen to have complied fully and even gone beyond the minimum requirements stipulated in the regulatory documents. Four of the contractors were observed not to have complied in the area of stacking and storage of materials, health and safety requirements and in wall construction; one contractor in two non-compliance instances of materials storage and cavity wall construction and another in three non-compliance instances of materials storage, cavity wall construction and fencing and hoarding requirements.

10.5 Level of Compliance by Grade of Registration with cidb

Table 4 cross-tabulates level of compliance with Grade of cidb registration.

Table 4 Level of Compliance by Grade of Registration with cidb

Level of Compliance	CIDB Register									Not Registered	Total
	1	2	3	4	5	6	7	8	9		
Very Good									1	1	2
Good					1				1	1	3
Non-compliance in 1 instance								1		3	4
Ditto in 2 instances				1							1
Ditto in 3 instances										1	1
Total				1	1			1	2	6	11

It can be inferred from Table 4, above, that three of the five firms registered at cidb Grades 5 and 9 were rated as 'very good' or 'good' in their level of compliance with the building regulations. However, four out of six of the unregistered contractors had instances of non-compliance with the building regulations. With such a small sample, no firm inference can be drawn, but it could indicate that registered large construction firms tend to comply more with the building regulations than smaller registered or unregistered firms.

12. Conclusions

The findings of the study are consistent with what was found in the literature regarding compliance, or no-compliance, with building regulations. The evidence, albeit from a small sample, suggests that there is a skills and knowledge gap in the South African construction industry and it is apparent that most of the key managerial positions are held by people who, though uneducated, have extensive experience. Those with the most experience are close to retirement age, which is a threat likely to widen the knowledge gap when their knowledge is lost to the industry.

It is also evident that the level of non-compliance tends to be more significant amongst the unqualified and less experienced firms not registered with the cidb. The findings of the study suggest that training of site managers in such firms is necessary. They also raise questions for further research into the capacity of the local authorities to enforce building regulations, and into the appropriateness of standards.

Acknowledgement

An Emerging Researcher Grant from the University of Cape Town Research Council, South Africa, has supported this study.

References

- Ang, G., Groosman, M., Scholten, N. (2005) Dutch performance-based approach to building regulations and public procurement. *Building Research and Information* (March – April) 33 (2), 107 -119.
- Baiche, B., Walliman, N., and Ogden, R. (2006) Compliance with Building Regulations in England and Wales. *Structural Survey*, 24(4), 279 – 299.
- Benge, C., (1999) "Dealing with alternative construction methods through performance based building codes", in Proceedings of CIB TG29 *Construction Industry Development In The New Millenium*, October 27-29, Singapore.
- Booyesen, H. (2010) City sells land to suspect builders: previous complaints about housing. *Claremont/Rondebosch Peoples Post*, Tuesday 9 February, 3.
- Bordass, B., Leaman, A. and Ruyssevelt, P. (2001) Assessing building performance in use 5: Conclusions and Implications. *Building Research and Information*, 29(2), 144 – 157.
- Brett, P. (1997) *An Illustrated Dictionary of Building: A Reference Guide for Practitioners and Students*; 2nd Ed; Butterworth-Heinemann, Oxford.
- Burby, J. R., May, P. J., Malizia, E. E., and Levine, J. (2000) Building Code Enforcement Burdens and Central City Decline. *Journal of the American Planning Association*, 66 (2), Spring, 2000, 143-161.
- Circo, C. J. (2008) Using Mandates and Incentives to Promote Sustainable Construction and Green Building projects in the Private Sector: A call for more State Land Use Initiatives; 2010 (web-based publication), 112 Penn State Law Review/ACREL pp 4; Accessed on 7th May, 2010.
- CIDB, (2007) *The Building and Construction Materials Sector, Challenges and Opportunities*. Construction Industry Development Board, Brooklyn Square, Pretoria.
- Fischer, T. (2010) City blocks hospital building plan. *Claremont/Rondebosch Peoples Post*, Tuesday 9 March, 1.
- Freeman, C. J. (1990) *The National Building Regulations: An Explanatory Handbook*; 2nd Ed; Juta, Kenwyn.
- Gann, D. M., Wang, Y. and Hawkins, R. (1998) Do regulations encourage innovation? – The case of energy efficiency in housing. *Building Research and Information*, 26(4), 280–296.

- Josephson, P. E. and Hammarlund, Y. (1999) The causes and costs of defects in construction. *Automation in Construction*, 8, 681–687.
- Landman, K., (2002) *Gated Communities in South Africa Building Bridges or Barriers?* Paper presented at an International Conference on Private Urban Governance, Mainz, Germany, June 6 – 9.
- Listokin, D. and Hattis, D. (2004) *Building Codes and Housing*, 2010 (web-based publication)
- Lo, S. M., Lam, K. C., Yuen, K. K., and Fang, Z. (2002) An Investigation of Building officials' Perception for the Use of Performance – Based Fire Engineering Approach in Building Design; *Fire technology*, 38, 271 – 286.
- Meacham, B., Bowen, R., Traw, J. and Moore, A. (2005) Performance – Based Building Regulation: Current Situation and Future Needs, *Building Research and Information*, (March – April) 33(2), 91 – 106.
- Meijer, F. M., Visscher, H. J., and Sheridan, L. (2002) *Building Regulations in Europe. Part 1A Comparison of the System of Control in Eight European Countries*, Housing and Urban Policy Studies 23; Delft University Press, ISBN: 90-407-2373-7.
- Noam, E. M. (1983) The Interaction of Building Codes and Housing Prices, *AREUEA Journal* vol. 10 pp 395 – 404.
- Parliament of New South Wales (2010), Joint Select Committee on the Quality of Buildings.; Accessed at: <http://www.parliament.nsw.gov.au/qualityofbuildings> on 15th August, 2010.
- Petersen, T. (2010) *Burglaries On The Rise in University Estate*: Southern Suburb Tattler, Thursday, February 11, pp. 7
- Spence, R. (2004) Risk and Regulation: Can Improved Government Action Reduce the Impacts of Natural Disasters? *Building Research and Information*, (September – October) 32(5), 391 – 402
- Stigler, G. J. (1971) The Theory of Economic Regulation; *The Bell Journal of Economics and Management Science*, vol.2, No 1, pp 3 – 21.
- Van Wyk, L. (2004) Review of the South African Construction industry Part 1: Economic, Regulatory and public Sector Capacity Influences on the Construction industry, 2004 (web-based publication), Information Resource Produced by CSIR Building and Construction Technology, March 2004 Issue; Accessed at: www.buildnet.co.za/akani/2004/mar/02.html on 29th October, 2009.
- Visscher, H., and Meijer, F. (2007) *Dynamics of Building Regulations in Europe*, being text of a paper presented at W09 – Legal Aspect of Housing, Land and Planning – ENHR 2007 International Conference on Sustainable Urban Areas, Rotterdam.
- Watermeyer, R. B and Milford, R. V. (2003) *The Use of Performance Based Building Codes to Attain Sustainable Housing Objectives: The South African Approach*; Paper presented at Global Policy Summit on the Role of Performance-Based Building Regulations in Addressing Societal Expectations, International Policy and Local Needs, National Academy of Sciences, Washington DC, USA, Nov 3-5; Accessed at: www.csir.co.za/akani on 27th October, 2009.
- Winch, G., and Campagnac, E. (1995) The Organization of Building Projects: an Anglo/French Comparison; *Construction Management and Economics* 13, 3 – 14.

The Duty to Co-operate in Construction Contracts – an International Review

Evaldas Klimas¹

¹Vilnius University, Law Faculty,
Sauletekio al. 9-I, 10222 Vilnius
Lithuania

Email: Evaldas.Klimas@rln.lt

Abstract:

Contractually non-described issues are inevitable when implementing relational contracts. Mutual understanding and co-operation are vital for solving such issues. Disputes regarding construction contracts are usually caused by a lack of co-operation between the parties involved – often because they do not understand or underestimate the obligation to co-operate. The purpose of this paper is to emphasize the importance of the duty of co-operation when implementing the rights and obligations in construction contracts. The paper elucidates the substance of the duty to co-operate in construction contracts and analyses the importance of notices, meetings, assistance, and good management on site. The market situation created by the worldwide credit crunch makes scrutiny of such issues all the more important.

The analysis is based on the following widely accepted international documents: UNIDROIT Principles of International Commercial Contracts (2004); The Principles of European Contract Law (2003), prepared by The Commission of European Contract law; the UNCITRAL Legal Guide on Drawing Up International Contracts for the Construction of Industrial Works (1987); with references also to the FIDIC Red Book of 1999, several jurisdictions of Common law and Continental law systems. Analysis of the above-mentioned sources leads to the following inferences. The duty to co-operate is related to the principles of good faith and fair dealing, which permeate the law of contract. In order to achieve the optimal result for the parties involved, a reasonable level of co-operation must be maintained; the parties should be well aware of certain measures to be taken during the construction process. Close attention to proper implementation of the duty to co-operate leads to better mutual understanding between parties and may prevent tension, conflicts and disputes from arising while implementing the construction contract.

Keywords:

Duty to co-operate, construction contracts.

1 Introduction

A contract is not merely a meeting point for conflicting interests but must also, to a certain extent, be viewed as a common project in which each party must co-operate. Duty to co-operate is established in such instruments of the unification of private law (which shall be examined) as the UNIDROIT Principles of International Commercial Contracts (2004) (hereinafter – the UNIDROIT Principles), The Principles of European Contract Law (2003) (hereinafter – the PECL), the UNCITRAL Legal Guide on Drawing Up International Contracts for the Construction of Industrial Works (1987), and the Draft Common Frame of Reference (hereinafter – the DCFR), prepared by the Study Group on a European Civil Code and Research Group on EC Private Law (*Acquis Group*). The duty to co-operate may be also found in FIDIC model documents, civil codes in Continental law system countries and case law in Common law system countries. Therefore, duty to co-operate may be considered as a general principle of law.

This paper analyses the above-mentioned documents drafted by international organisations/research teams. As these documents are drafted on the basis of international comparative analysis, the paper does not analyse legislation of particular jurisdictions, with the exception of several cases. The Author uses means of analysis and comparative methods in this paper. The paper's primary purpose is to elucidate the substance of the duty to co-operate, which is a substantive principle, and also to answer the question: what is the impact of proper implementation of the duty to co-operate in preventing conflicts and disputes, and making the construction process more fluent and coordinated? This is a very pertinent topic in long-term construction contracts, where proper implementation of the contract depends heavily on a wide range of participants: architects, engineers, and technical supervisors etc.

2 The principle of duty to co-operate

In many Civil law systems, it is considered that the duty to co-operate derives from the principle of good faith. The duty to co-operate is also recognised as a general principle of *lex mercatoria* (Vogenauer and Kleinheisterkamp 2009, p. 543). The duty to co-operate is implied in the key instruments of the unification of private law. These instruments are analyzed below in attempt to elucidate the substance of the analysed principle.

2.1 The principle of duty to co-operate in the UNIDROIT Principles 2004

The UNIDROIT Principles clarify that when concluding a contract, contracting parties receive not only rights but also obligations towards each other. In certain cases, one contracting party should perform or refrain from certain actions in order to allow the other party to implement its duties – i.e. the parties are under a duty to co-operate.

The duty to co-operate is stated in the fifth chapter (Content and third-party rights), regulating the basic implied terms of contract law. Under Article No. 5.1.3 (Co-operation between the parties) of UNIDROIT Principles (2004 edition, p. 102), each party must co-operate with the other party when such co-operation may reasonably be expected for the performance of that party's obligations. Thus, the obligation to co-operate is settled next to the very basic implied principles: good faith and fair dealing,

reasonableness etc. (article No. 5.1.2.). The placing of the duty to co-operate in the fifth chapter, and the fact that it is distinguished as a separate principle, show the great importance attached to the duty and implies that it is considered a general principle of contract law.

The cited provision clearly states that the duty to co-operate is not an ordinary obligation of one party to another party. The duty to co-operate cannot be considered as one party's absolute right to demand from another party the performance of certain actions. The provision draws a line between the right to demand co-operation and the obligation to co-operate. As the duty to co-operate is not an ordinary obligation, but a principle, the line is flexible and can be best expressed by one word – reasonable.

UNIDROIT Principles commentary (1994 edition, p. 103) states that “the duty of co-operation must of course be confined within certain limits [the provision refers to reasonable expectations], so as not to upset the allocation of duties in the performance of the contract.” Although the principal concern of the provision is the duty not to hinder the other party's performance, there may also be circumstances which call for more active co-operation. Thus, the obligation of one party to co-operate with the other is defined with the abstract aspect of reasonable expectations, which depend on the factual circumstances. It is very difficult to determine the substance of the duty to co-operate, as necessarily the duty significantly depends on the factual situation. Indeed, the circumstances of the contract implementation may require the parties to be either passive or to be active in the performance of their contractual obligations.

The latest commentary on the UNIDROIT Principles (Vogenauer and Kleinheisterkamp 2009, p. 543, 544) distinguishes two dimensions of the duty to co-operate: (i) each party is under a duty to remain passive if a particular action might hinder the performance of the other party. Parties must refrain from obstructing the other party's efforts to perform; (ii) each party is under a duty to engage in actions if such actions are required to enable or facilitate the other party's performance. In the first case, if for example the owner should provide the site and if the contractor so requires, the owner should refrain from performing any activity on the site if such activity could obstruct the proper implementation of the contractor obligations. In the second case, for example, the contractor may request the owner to provide the contractor with necessary solutions or responses to the contractor's questions or help obtain instructions from an architect.

The UNIDROIT Principles do not define the scope of the duty to co-operate in detail. They merely provide abstract criteria of “reasonable expectations”, which help to identify the level of necessity to co-operate and actions which should be performed. The Principles specify, however, that the duty to co-operate may be performed in both active or in passive ways, depending on the facts in a particular case. Such abstract determination of the duty to co-operate is thought-provoking and requires the parties to act on their best endeavour, while seeking proper implementation of the co-operation.

2.2 The principle of duty to co-operate in the PECL

The duty to co-operate is also found in the PECL article No. 1:202 (Duty to Co-operate; PECL 2003). The provision states that each party owes to the other a duty to co-operate in order to give full effect to the contract. Apparently, the PECL wording differs from

that of the UNIDROIT Principles. The PECL, as well as the UNIDROIT Principles, require the parties to co-operate as “reasonably expected.” However, the PECL stipulates additional criteria, which help to define whether the party duly implements the obligation to co-operate – i.e. the requirement to act in order “to give full effect to the contract.” On first sight, it might be thought that one party may request unreasonably much from its counterparty. However, each party has its own obligations under the contract. Moreover, the party may not be deemed in breach of the duty to co-operate, if such party does not perform what it was not obliged to perform (Lando and Beale 1995, p. 60). Therefore, even though parties are under a duty to co-operate, a reasonable balance between the parties must be maintained, while paying attention to their obligations.

While the wording of the duty to co-operate differs in the UNIDROIT Principles and PECL, the nature of this principle is the same. Both parties should act in their best ability while implementing the contract. In order to assess whether the duty to co-operate was properly implemented, it is necessary to analyse the balance of the party obligations, the factual situation and the final result – i.e. whether it was or was not achieved due to the lack of co-operation of one of the parties.

Under the PECL, non-co-operation of one party is considered as non-performance (with all the consequences thereof). Paragraph 4 of Article 1:301 (Meaning of Terms) provides that ‘non-performance’ denotes any failure to perform an obligation under the contract, whether or not justified. The “failure to perform” includes delayed performance, defective performance and failure to co-operate in order to give full effect to the contract. Moreover, under the UNIDROIT Principles, non-co-operation is considered as non-performance as well (Vogenauer and Kleinheisterkamp 2009, p. 547). This implies that proper implementation of the duty to co-operate should be treated not only as a theoretical principle, but also as an obligation. This obligation is especially important in contracts obliging parties to work together; for example, construction contracts where a contractor may implement the idea of the owner (or his architect) only if worksheets are clear.

2.3 The principle of duty to co-operate in UNCITRAL Legal Guide

The UNCITRAL Legal Guide (1987, p. 48) suggests that successful implementation of a contract depends on co-operation between the parties. The Guide clarifies that, while it may be impossible to list all instances when parties should co-operate in the contract, it is desirable that the contract generally obliges each party to both co-operate with another party to the extent necessary for the performance of the other party’s obligations, and also avoid conduct interfering with that performance (UNCITRAL Legal Guide 1987, p. 48). In practice, it is impossible to foresee all the conditions in one single long-term construction contract. Questions during the construction process are inevitable. However, such questions can be addressed by mutual consultations – i.e. proper co-operation. Early solving of questions can prevent disputes.

2.4 The principle of duty to co-operate in the DCFR

While there is no doubt that all parties should co-operate while implementing construction contracts, the scope of the principle is not clear-cut. The DCFR, issued in 2009, attempts to clarify the scope of the duty co-operate. The document was prepared

by the legal academics Study Group on a European Civil Code and Research Group on EC Private Law (Acquis Group) and reflects a compromise between different legal systems, as it was based on legal regulation and best practice of different EU member states. The provisions of the DCFR go a little further regarding the duty to co-operate than the UNCITRAL Principles and the PECL. As the DCFR is reminiscent of a classical code, the duty to co-operate is determined in several chapters.

In the DCFR, the duty to co-operate is set as a general principle of the law of obligations. DCFR III. – 1:104 (Co-operation) states the general principle of co-operation. It provides that the debtor and the creditor are obliged to co-operate with each other when and to the extent that can reasonably be expected for the performance of the debtor (Study Group on a European Civil Code and Research Group on EC Private Law (Acquis Group) 2009, p. 230). The wording is closer to the wording of the UNIDROIT Principles than of the PECL. The duty to co-operate is determined by the criteria of reasonable expectations.

The DCFR classifies construction contracts as services contracts. The specific obligations stemming from the duty to co-operate in the construction process are found in the following provisions: (i) Article No. IV. C. – 2:103 (Obligation to co-operate, stated in Chapter 1 (General provisions) of Part C – (Services); and (ii) Article No. IV. C. – 3:102 (Client's obligation to co-operate), stated in Chapter 3 (Construction) of Part C – (Services). The provisions lay down each party's obligations for proper implementation of the duty to co-operate. Article No. IV. C. – 2:103 states that (1) the obligation of co-operation requires in particular:

- (a) the client to answer reasonable requests by the service provider for information insofar as this may reasonably be considered necessary to enable the service provider to perform the obligations under the contract;
- (b) the client to give directions regarding the performance of the service insofar as this may reasonably be considered necessary to enable the service provider to perform the obligations under the contract;
- (c) the client, insofar as the client is to obtain permits or licences, to obtain these at such time as may reasonably be considered necessary to enable the service provider to perform the obligations under the contract;
- (d) the service provider to give the client a reasonable opportunity to determine whether the service provider is performing the obligations under the contract; and
- (e) the parties to co-ordinate their respective efforts insofar as this may reasonably be considered necessary to perform their respective obligations under the contract (Study Group on a European Civil Code and Research Group on EC Private Law (Acquis Group) 2009, p. 304, 305).

In addition, obligations specific to the construction industry are laid down in Article No. IV. C. – 3:102, which provides that the duty to co-operate requires the client to:

(a) provide access to the site where the construction has to take place insofar as this may reasonably be considered necessary to enable the constructor to perform the obligations under the contract; and

(b) provide the components, materials and tools, insofar as they must be provided by the client, at such time as may reasonably be considered necessary to enable the constructor to perform the obligations under the contract (Study Group on a European Civil Code and Research Group on EC Private Law (Acquis Group) 2009, p. 311).

On first sight it appears as if the DCFR refers only to active actions of the parties for the proper implementation of the duty to co-operate: i.e. give answers, provide directions, obtain permits, coordinate, provide access, and provide goods. It is not an exhaustive list of obligations comprising the duty to co-operate. In practice, where specific obligations are not defined, the general principle of duty to co-operate applies.

Nevertheless, in the opinion of the author of this paper, under the DCFR the duty to co-operate may require a party to stay passive in certain cases. For example, the contractor may have to provide necessary information to the client and to refrain from interfering, to allow the client to make independent observations, and notify anticipated non-conformity or directions to the contractor thereafter. All these obligations reflect common construction practice and are (or may) be respectively acknowledged in Civil law system countries or in Common law systems. However, exact obligations may vary and should be reviewed accurately.

If a construction contract does not specify the actions essential for implementing the duty to co-operate, the necessity to perform certain actions can be determined according to the reasonableness test. This test has already been analysed and is embedded not only in the DCFR but also in the UNCITRAL Principles. Finally, it should be noted that the above listed obligations are only part of all the obligations constituting the duty to co-operate. In other cases, where the exact duty is not defined, the main principle of duty to co-operate applies with proper implementation of the reasonableness test.

2.5 The principle of duty to co-operate in construction contract forms

2.5.1 The principle of duty to co-operate in the FIDIC

Some of the best-known construction contract forms are the FIDIC construction contracts. These contracts are widely used because of the implemented balance between the parties. Of course, some authors may disagree with this statement, but this is a question of separate discussion. One of the basic FIDIC contract forms is the Construction Work Contract Form of the first edition 1999 – the so-called Red Book 1999, where parties' rights and obligations are defined in twenty chapters.

The Red Book 1999 provides for the duty to co-operate both implicitly and explicitly: for example, an employer's obligation to give the contractor right of access to, and possession of, all parts of the site (clause 2.1); or a contractor's obligation to give notice to the engineer as soon as practicable if the contractor encounters adverse physical conditions, which he considers to have been unforeseeable (clause 4.12).

The duty to co-operate is also explicitly established as a separate contractors' obligation in the FIDIC Red Book 1999. Sub-clause No. 4.6 (Co-operation) states that the contractor shall, as specified in the contract or as instructed by the engineer, allow appropriate opportunities for carrying out work to: (a) the employer's personnel, (b) any other contractors employed by the employer, and (c) the personnel of any legally constituted public authorities, who may be employed in the execution on or near the site of any work not included in the contract (FIDIC 1999). The same sub-clause clarifies that any such instructions shall constitute a variation of and to the extent that it causes the contractor to incur unforeseeable cost. Services for the personnel and other contractors may include the use of contractor's equipment, temporary works or access arrangements which are the responsibility of the contractor (FIDIC 1999). As we can see, the clause of co-operation is not analogous to the above analysed clauses of the duty to co-operate in UNIDROIT Principles or PECL. The FIDIC Red Book 1999 does not provide for the duty to co-operate as a general obligation of both parties. As Glover states, despite being headed "co-operation", this sub-clause does not seem to refer to co-operation in the partnering sense of the word, or arguably in any real sense at all (2006, p. 92).

The above cited FIDIC Red Book 1999 sub-clause of co-operation indicates several issues which can be considered as substantive elements of the duty to co-operate. Of course, this does not mean that the parties do not have the duty to co-operate while implementing the contract. The duty to co-operate is a general principle of contract law, which is widely acknowledged in international practice and is a codified norm in national laws or an implied legal concept arising from the case law. Therefore, the parties have a duty to co-operate even under the FIDIC Red Book 1999 though it simply lists a limited number of obligations, which are a part of the general principle – duty to co-operate.

2.5.2 The principle of duty to co-operate in other standard construction contract forms

The duty to co-operate is also an important principle while implementing obligations under other well-known standard contract forms, such as the JCT and the NEC. The duty to co-operate is not directly expressed in the JCT standard terms, but this does not mean that the duty to co-operate can be ignored. It is an implied term under English law that the employer will do all that is reasonably necessary on his part to bring about completion of the contract (Furmston 2006, p. 164). Implementation of a building contract embodying the JCT general conditions does require close co-operation between the contractor and the architect (Furmston 2006, p. 171). The importance of the duty to co-operate is also emphasised in the Early Contractor Involvement (ECI). Mcinnis (2003, p. 289) states that co-operation is key in both relational contracting and under the NEC (New Engineering Contract).

The importance of the duty to co-operate is acknowledged in the NEC as well as in long-term contracts. This was influenced by relational contract theory, the basics of which were elaborated by MacNeil¹. Mcinnis (2003, p. 130) presents a short distillation of MacNeil's work: the success of the contractual relationship depends less upon what

¹ For more see MacNeil, I. R. (1980), *The New Social Contract*, Yale University Press, New Haven.

has been agreed than upon how the parties will agree to handle events in the future. Relational contract theory is only one of many modern contract law theories trying to overcome shortcomings of traditional contract law. Under classical contract law, there ought to be no room for an additional unexpressed obligation to co-operate (Quick 2007, p. 7). However, common law courts have insinuated the obligation into contracts by a variety of devices including the general rules of interpretation and implication of terms (Quick 2007, p. 7), and more will be discussed in the next section of this paper.

Clause 10.1 of the NEC3 states that an employer, contractor, project manager and supervisor must act as stated in the contract and “in a spirit of mutual trust and co-operation.” The clause provides for two obligations: i) to act as stated in the contract; and ii) to act in a spirit of mutual trust and co-operation (Eggelston 2006, p. 85). In other words, the duty to co-operate goes beyond the provisions of the contract and obliges the parties to act on the basis of mutual trust and co-operation. According to some commentators, contractors, rather than employers, are likely to be the main beneficiaries of any uncertainty, as they have more opportunities to capitalise on uncertainty as to whether individual provisions are to be applied in a spirit of mutual trust and co-operation and they are better placed to prove damages in any claim based on failure to do so (Eggelston 2006, p. 85). However, the author of this paper considers that such concern is unsubstantiated. Uncertainties in relational contracts may be addressed and issues arising from such uncertainties may be solved or avoided all together if both parties co-operate, i.e. act fairly and put their best efforts to implement the contract.

Moreover, clause 16 of NEC3 sets out early warning procedures. The clause obliges the contractor to warn the project manager on certain matters. Both the contractor and project manager are then required to attend an early warning meeting. The main purpose of such a meeting is to discuss possible problem solutions. The clause provides for a partnering based approach – to resolve an issue before it causes a dispute. In that respect, co-operation between the parties generally and early warning meetings in particular allow the parties to discuss and resolve the matter in the most efficient manner.

Consequently, the parties have a duty to co-operate, regardless of which above mentioned contract form was chosen for the project implementation, but national law is not less important as chosen contract form.

2.6 The principle of duty to co-operate and national law

Different jurisdictions follow different approaches regarding good faith and the duty to co-operate (these differences are especially apparent when comparing continental and common law jurisdictions), but the comparison of these approaches requires separate research.) *Travaux preparatoire* of the UNCITRAL Principles, the PECL and the DCFR contain a rather comprehensive analysis and comparison of different legal systems. For the purposes of this paper, the author focuses his analysis on international documents as they reflect the common view of different legal systems on the principle of duty to co-operate. However, for the sake of completeness, the author nevertheless refers to several examples of national legislation on the duty to co-operate.

2.6.1 *Duty to co-operate under German law*

In German civil law, duty to co-operate is derived from the principle of good faith. Under article No. 242 (Performance in good faith) of German Civil Code (the BGB)² an obligor has a duty to perform certain actions as required by the principle of good faith. The concept of good faith is one of the core legal principles in German civil law (while common law jurisdictions follow the opposite approach)³. Moreover (as co-operation duty is essential in the contracts to produce work), article No. 642 (Collaboration by the costumer) of the BGB provides that if a customer must perform certain actions in order to achieve the result envisaged in the contract, the contractor may demand reasonable compensation if the customer, by failing to perform the act, is in default of acceptance. As we see, the general principle that the customer of the contractor has a duty to co-operate is firmly established in one of major jurisdictions in Europe – German civil law.

2.6.2 *Duty to co-operate under Lithuanian law*

Co-operation duty is determined in the codes of small continental law jurisdictions, especially where new codes were enacted recently; for example, article No. 6.691 (Co-operation of parties to a contract of construction independent work) of the Civil Code of the Republic of Lithuania⁴ provides that in implementing a construction contract, parties to the contract are obliged to co-operate. In the event of obstacles hindering implementation of the contract, each of the parties is obliged to take all available reasonable measures to eliminate such obstacles. Accordingly, the party that fails to perform this duty is deprived of the right to claim damages caused due to the relevant obstacles. Both parties of the construction contract have a general duty to co-operate that is understood as the duty to remain passive and the duty to engage in actions. Notably, such wording of the provision gives national courts ultimate freedom of interpretation as to which party will be considered responsible for failure to co-operate on a case-by-case basis. Therefore, not only the rules of the codes but also judicial practice should be considered when identifying the concept of duty to co-operate under exact jurisdiction.

2.6.3 *Duty to co-operate under common law*

Common law jurisdictions have already firmly established duty to co-operate in long-term contracts. For example, under English law, in leading case *London Borough of Merton LBC v Stanley Hugh Leach Ltd*⁵ two general employers' obligations were established:

- (a) The employer will not hinder or prevent the contractor from carrying out all its obligations in accordance with the terms of the contract, and from executing the works in a general and orderly manner;

² German Civil Code (BGB), http://www.gesetze-im-internet.de/englisch_bgb/englisch_bgb.html#BGBengl_000G1, viewed 03/06/2010.

³ For more see: WHITTAKER, S. and ZIMMERMANN, R., 2000. Good faith in European contract law: surveying the legal landscape. In: ZIMMERMANN, R. and WHITTAKER, S., ed. *Good faith in European contract law*. Cambridge: Cambridge University Press, 7- 62.

⁴ The Civil Code of the Republic of Lithuania (with amendments) (Valstybės žinios, 2000, No. 74), http://www3.lrs.lt/pls/inter3/dokpaieska.showdoc_l?p_id=370997, viewed 03/06/2010.

⁵ (1985) 32 BLR 51.

(b) The employer will take all steps reasonably necessary to enable the contractor to discharge all its obligations and to execute the works in a regular and orderly manner.

As Murdoch and Hughes (2008, p. 172) observe, these two general obligations are in reality “the positive and negative aspects of the same thing.” Together they make in the employer’s duty to co-operate with the contractor in all aspects of the construction work (Murdoch and Hughes 2008, p. 172). Duty to co-operate is usually referred to as employer’s obligation, because the contractor is usually the weaker party in construction contracts, and in most cases the co-operation of the employer is required for the proper implementation of the contract.

The author notes that the concept of duty to co-operate may vary from country to country, and attention should be paid to national legal regulations when determining the duty to co-operate.

3 Implementing the duty to co-operate in the construction process

Nowadays the construction process is increasingly complex as huge and expensive projects (e.g. road construction, bridges, airports, sports stadiums, music arenas etc.) are planned and implemented. Such projects involve many participants: architects, engineers, supervisors, consultants, contractors, subcontractors, etc. The implementation of such projects usually lasts a long time. During the implementation of long-term projects, unforeseen circumstances may often arise. This leads the parties to perceive long-term contracts as common projects. According to Campbell and Harris (1993, p. 167), efficient long-term contractual behaviour must be understood as consciously co-operative. Parties perceive long-term contracts as analogous to partnership.

If the construction process is indeed perceived as a common project, and the construction contract as analogous to partnership, the duty to co-operate is thus a pertinent element of that process. However, in practice, the importance of the principle of the duty to co-operate in contract law is underestimated. Mcinnis (2003, p. 293) analysed the duty to co-operate in depth; he concluded that the duty to co-operate is not an additional obligation under classical contract law. However, Mcinnis argues, the courts have imposed the duty to co-operate on the parties. In other words, the duty to co-operate was acknowledged to exist in itself.

3.1 Duty to co-operate prevents disputes

Usually conflicts arise because the parties misunderstand each other. Occasionally even minor contradictions may develop into a complicated dispute. Proper implementation of the duty to co-operate may at the very least prevent many initial contradictions.

First, it is very important that each party is able to express its needs and expectations. As Jansen (1998, p. 147) observes, the employer is under an obligation to inform the contractor about his ideas, needs and expectations in order to facilitate the contractor’s performance of his obligations under the contract. In a traditional construction process, the communication of expectations is a condition *sine qua non* to the contractor’s ability to commence performance of his obligations under the contract (Jansen 1998, p. 148). Such communication of expectations can be reflected in various ways, e.g. as Jansen

(1998, p. 154) suggests, the description of the work to be carried out by submitting bills of quantities is a characteristic feature of the general conditions of building contracts used in English legal culture. Other countries also have similar measures under which the expectations can be reflected clearly.

Second, disclosure of information and providing timely directions is equally important. Even if the needs and expectations of the parties are expressed, withholding of some information may lead to significant damage. As Jansen (1998, p. 170) states, the employer is under an obligation to disclose information by means of directions; and this general obligation derives from the general obligation to provide the contractor with necessary information under various legal systems such as France, Germany, and England. Moreover, proper implementation of this obligation may reinforce the bonds between the parties. Usually, the engineer's and surveyor's roles are of high importance while implementing the obligation to give instructions in reasonable time.

Third, exchange of information between the parties should be ensured. The best way to do so may be by holding meetings. Meetings as a dispute prevention tool was identified in the UNCITRAL Legal Guide (1987, p. 111), which suggests that a contract may provide that representatives of the parties are to meet periodically at specified intervals on the site to promote co-operation between them and to resolve outstanding issues concerning the construction on site. Such meetings may help to resolve routine problems or misunderstandings, and thus dispel the need to invoke contract provisions on dispute settlement.

Finally, good conflict management may prevent parties from disputes. Disagreements that arise in meetings can be solved only if the parties possess conflict management skills. In practice, usually professional engineers and surveyors are called upon to settle disagreements between the parties. Wilmot-Smith (2006, p. 54) observed, "contractors had professionals on their staff that were often as good as or better than the professional engineers and surveyors who were employed by the firm of architects or engineers which was engaged by the employer". It should be noticed that sometimes personnel of the contractor are held to be better at contract conflict management than many professionals, but engineers as mediators or adjudicators are usually called upon only when serious questions arise and where the management of the conflict is complicated. This shows that contractor staff can't handle all situations and professional dispute settlers should be involved in the process.

In conclusion, the emergence of conflicts may be prevented by proper implementation of the following elements of the duty to co-operate: (i) clear statement of parties' needs and expectations, (ii) disclosure of information and provision of timely directions, (iii) exchange of information (especially during meetings) and (iv) good conflict management.

3.2 Duty to co-operate conditioned by outside factors

Even if parties do their best, it is very difficult to predict future difficulties, which are beyond the control of the parties. Proper implementation of the duty to co-operate may help to solve the problems during the implementation of the contract.

One such example can be exempting impediments. If one party invokes an exempting impediment, it should give written notice of the impediment to the other party without undue delay. This notification could facilitate the taking of measures by the other party to mitigate the loss caused or which is likely to be caused by failure of performance (UNCITRAL Legal Guide 1987, p. 239). Upon receipt of notice on impediment, as the duty to co-operate requires, the parties should meet and consider the measures to be taken in order to prevent or limit the effects of the impediment. The parties certainly also have further duties, such as the duty to prevent and mitigate any loss which may be caused by the impediment.

Such events as breakdowns of economic systems or other upheavals or exceptional circumstances can also considerably influence the conditions under which the parties had assessed their risks and estimated costs before concluding the contract. Such unforeseen circumstances can distort the balance of performance of parties' contractual obligations. The question then arises of which party should bear the risk of such unforeseen changes and to what extent. It is generally acknowledged that the question is to be addressed by applying two fundamental principles – the sanctity of contract (*pacta sunt servanda*) and the equal fundamental principle of good faith (Brunner 2009, p. 391). Without going deeply into the civil law system doctrine of *clausula rebus sic stantibus*, it should be emphasised that hardship cases are solved similarly in different countries. In order to prevent a dispute, parties should co-operate and solve problems, and evaluate which of them should bear the risks. In other words, again, the parties should co-operate – give notices, communicate and negotiate in attempt to sustain their contractual relationship. The UNIDROIT Principles and the PECL establish that in the case of hardship, the disadvantaged party is entitled to request restarting negotiations⁶.

Variation is one more issue which is worth mentioning. As Pfeiffer (2007, p. 166) states, if the subject of contract is complicated then the contract needs to be sufficiently flexible to accommodate later change, and in the construction industry this has always been the situation. Of course, there should not be too many variations, because this can complicate the execution of the contract. Even in the case of variations, if the parties wish to implement the contract, they should co-operate: exchange information, communicate their expectations.

4 Conclusions

The above analysis yields the following main conclusions:

1. The duty to co-operate is a general principal of contract law, which is closely connected with the principle of good faith and fair dealing and plays a major role while implementing construction contracts, and permeates the law of contract;
2. It is very difficult to determine the substance of the duty to co-operate, as the necessity to apply the duty significantly depends on the factual situation and national legal regulations. Nevertheless, two dimensions can be easily distinguished: (i) duty to remain passive and (ii) duty to engage in actions;

⁶ For more see: Brunner 2009, p. 393 - 399.

3. The framework of the duty to co-operate is determined by the criteria of reasonable expectations. Proper implementation of the duty to co-operate can be estimated by analysing the balance of the obligations of the parties, factual situation and final result, whether it was reached or was not reached because of the lack of co-operation of one of the parties;

4. The substance of the duty to co-operate while implementing the construction contract comes from common construction practice. The duty to co-operate comprises such obligations as: (i) to give answers; (ii) to provide directions; (iii) to obtain permits; (iv) to coordinate; (v) to provide access; (vi) to provide goods as agreed; (vii) to provide opportunity for determination (viii) clearly state needs and expectations; (ix) to disclose information and provide timely directions; (x) to provide exchange of information during the meetings; (xi) to take care of good management.

5. The parties should co-operate even if one party invokes an exempting impediment, unforeseen circumstances occur or one party requests variations. Proper implementation of the duty to co-operate may determine the smooth implementation of the construction process, prevent the onset of conflicts or help to solve disputes in their initial stages.

5 References

- Brunner, C. (2009), *Force Majeure and Hardship under General Contract Principles: exemption for Non-Performance in International Arbitration*, Kluwer Law International, Austin.
- Campbell, D., Harris, D. (1993), 'Flexibility in Long-term Contractual Relationships: The Role of Co-operation', *Journal of Law Society*, Volume 20 (issue no. 2), pp. 166-191.
- Commission on European Contract Law (2003), *The Principles of European Contract Law*, <http://frontpage.cbs.dk/law/commission_on_european_contract_law/PECL%20engelsk/engelsk_partI_og_II.htm> accessed 13 January 2010.
- Eggliston, B. (2006), *The NEC3 engineering and construction contract: a commentary. 2nd edition*, Blackwell Publishing, Oxford.
- Furmston, M. (2006), *Powell-Smith & Furmston's Building Contract Casebook. 4th edition*, Blackwell Publishing, Oxford.
- Glover, J., Hughes, S., Thomas, C. (2006), *Understanding the New FIDIC Red Book. A Clause-By-Clause Commentary*, Sweet & Maxwell, London.
- International Federation of Consulting Engineers (FIDIC) (1999), *Conditions of Contract for Construction (First Ed. 1999)*, FIDIC, Switzerland.
- International Institute for the Unification of Private Law (1994), *UNIDROIT Principles of International Commercial Contracts*, <http://www.unidroit.org/english/principles/contracts/principles1994/1994fulltext-english.pdf>, viewed 26/05/2008.
- International Institute for the Unification of Private Law (2004), *UNIDROIT Principles of International Commercial Contracts 2004*, [http://www. Unidroit.org/English/principles/contracts/principles2004/blackletter2004.pdf](http://www.Unidroit.org/English/principles/contracts/principles2004/blackletter2004.pdf), viewed 26/05/2008.
- Jansen, C. E. C. (1998), *Towards a European building contract law. Defects liability: a comparative legal analysis of English, German, French, Dutch and Belgian law*, Kluwer, Deventer.

- Lando, O. and Beale H. (1995), *Principles of European Contract Law Part I: Performance, Non-Performance and Remedies*, Martinus Nijhoff Publishers, Dordrecht.
- Mcinnis, A. (2003), 'The new engineering contracts: relational contracting, good faith and co-operation – part 1', *The International Construction Law Review*, Volume 20, pp. 128 – 153.
- Mcinnis, A. (2003), 'The new engineering contracts: relational contracting, good faith and co-operation – part 2', *The International Construction Law Review*, Volume 20, pp. 289 – 325.
- Murdoch, J.R. and Hughes, W. (2008), *Construction Contracts: Law and Management*, 4th edition, Taylor & Francis, London.
- Pfeiffer, T. (2007), 'New Mechanisms for Concluding Contracts', In: Schulze, R., *New Features in Contract Law*, Sellier. European Law Publishers, Munich, pp 161 – 168.
- Quick, R. (2007), 'Queensland's ECI contract', *The International Construction Law Review*, Volume 24, pp. 4 – 35.
- Study Group on a European Civil Code and Research Group on EC Private Law (Acquis Group) (2009), *Principles, Definitions and Model Rules of European Private Law. Draft Common Frame of Reference (DCFR)*, Sellier. European Law Publishers GmbH, Munich.
- United Nations Commission on International Trade Law (UNCITRAL) (1987), *Legal Guide on Drawing Up International Contracts for the Construction of Industrial Works*, United Nations, New York.
- Wilmot-Smith, R. (2006), *Construction Contracts: Law and practice*, Oxford University Press, New York.
- Vogenauer, S. and Kleinheisterkamp, J. (2009), *Commentary on the UNIDROIT Principles of International Commercial Contracts (PICC)*, Oxford University Press, Oxford.

Legal Implications of New Equipment Anti-idling Requirements

Thomas M. Korman¹ and Lonny G. Simonian¹

¹Construction Management Department,
California Polytechnic State University, San Luis Obispo, CA
USA

Email: tkorman@calpoly.edu; lsimonia@calpoly.edu

Abstract:

Diesel engines have been documented to emit a complex mixture of air pollutants, composed of gaseous and solid material. The visible emissions in diesel exhaust are known as particulate matter (PM), particles with a diameter of less than 2.5 micrometers, air quality problems. In 1998, California identified diesel exhaust PM as a toxic air contaminant based on its potential to cause cancer, pre-mature death, asthma, and other health problems. Based on the year 2000 emissions in California, diesel PM contributes each year to 2,000 premature deaths, thousands of hospital admissions, asthma attacks, and other respiratory symptoms, and lost workdays. Overall, diesel engine emissions are responsible for the majority of California's known cancer risk from outdoor air pollutants. In addition, diesel soot causes visibility reduction and is a potent global warmer. The California Air Resources Board (CARB) recently adopted a regulation to reduce diesel PM and oxides of nitrogen (NOx) emissions from in-use (existing) off-road heavy-duty diesel vehicles in California. Such vehicles are used in construction, mining, and industrial operations. This legislation prevents the idling of heavy-duty vehicles used on construction sites, such as concrete trucks, scrapers, graders, dump trucks, backhoes, loaders, cranes, concrete pump trucks, etc. Since the inception it has been affecting the daily construction operations, construction cost estimating, and the management for construction equipment fleet management. This paper reviews the current regulations regarding anti-idling and its effects on construction operations, estimating, and fleet management and the legal implications

Keywords:

Air Quality, Construction Equipment Operations, Engine Idling, Equipment Maintenance, Inspection.

1 Introduction

Diesel engines emit a complex mixture of air pollutants, composed of gaseous and solid material. The visible emissions in diesel exhaust are known as particulate matter (PM). In 1998, California identified diesel exhaust PM as a toxic air contaminant based on its potential to cause cancer, pre-mature death, asthma, and other health problems. Diesel engines also contribute to California's fine PM,

particles with a diameter of less than 2.5 micrometers. Those most vulnerable are children whose lungs are still developing and the elderly who may have other serious health problems. Based on year 2000 emissions in California, diesel PM contributes each year to premature deaths, thousands of hospital admissions, asthma attacks, and other respiratory symptoms, and lost workdays. Overall, diesel engine emissions have been determined to be responsible for the majority of California's known cancer risk from outdoor air pollutants.

2 Summary of the Adverse Impacts of Diesel Particulate Matter

Diesel engines emit a complex mix of pollutants, the most visible of which are very small carbon particles or "soot". As described above, this is known as diesel PM and contains over 40 cancer-causing substances, most of which are readily adsorbed as soot particles. Several international and national health agencies have concluded that diesel exhaust has the potential to contribute to cancer and other health effects. These agencies include the National Institute of Occupational Safety and Health, (NIOSH) the International Agency for Research on Cancer, the World Health Organization (WHO), the National Toxicology Program and the U.S. Environmental Protection Agency (EPA). Overall, diesel engine emissions are responsible for the majority of California's estimated cancer risk attributable to air pollution. In addition, diesel PM is a significant fraction of California's particulate pollution problem.

Estimates of the magnitude of the contribution of diesel PM to air pollution related deaths are based on health studies involving ambient PM pollution in general. The method used to quantify diesel PM's effect likely underestimates its contribution because this method assumes that diesel PM, per unit of concentration, has the same health impact as general ambient particulate matter. However, the combination of several factors strongly suggests that diesel emissions pose a significantly greater health threat than non-combustion particle pollution. These factors include the following:

- Diesel PM is more often emitted close to people so high exposures often occur
- Diesel PM more readily deposits in the lungs and can be absorbed in the body
- Diesel PM contains compounds known to damage DNA and cause cancer

In addition, there are other important reasons to reduce diesel PM pollution:

- The value of the health benefits are expected to vastly exceed the control costs
- Diesel PM causes visibility reduction, and
- Diesel PM is a potent contributor to global warming

Table 1. Summary of the Health and Environmental Effects of Diesel PM Exposure for California

(Source: California Air Resources Board, 2008)

Effects of Diesel PM Exposure for California	Occurrences Per Year
Premature Deaths	2000
Lung Cancer	250

3 Background and History of the Legislation and Regulations

California Health and Safety Code, Division 26 (Air Resources) authorizes the CARB to adopt regulations to protect public health and the environment through the reduction of toxic air contaminants (TAC's) and other air pollutants with adverse health effects. Following the identification of diesel PM as a toxic air contaminant in 1998, the CARB adopted the Diesel Risk Reduction Plan (DRRP) in October 2000 to reduce diesel PM emissions and resultant health risk to "near zero" by 2020.

The airborne toxic control measure for diesel particulate matter from off-road heavy-duty diesel vehicles was one in a series of rules designed to reduce diesel PM from most diesel-fuelled heavy-duty vehicles in California. Prior to enacting the legislation restricting idling for heavy-duty trucks used on construction sites, similar legislation was enacted for school buses, solid waste collection vehicles, and on-road diesel-fuelled residential and commercial vehicles. The following section summarize these legislative policies.

On July 16, 2003, the CARB approved an airborne toxic control measure (ATCM) that limits school bus idling at or near schools to only when necessary for safety or operational concerns. This ATCM to limit idling is intended to reduce diesel exhaust particulate matter and other TAC's and air pollutants from heavy-duty motor vehicle exhaust. The school bus ATCM was the ARB's first priority because children riding in, and playing near, school buses and other commercial motor vehicles are disproportionately exposed to pollutants from these sources. Generally, children are more vulnerable than adults to air pollutants because they have higher inhalation rates, narrower airways, and less mature immune systems.

The ATCM requires the driver of a school bus, transit bus, or other commercial motor vehicle to manually turn off the vehicle engine upon arriving at a school and to restart no more than thirty (30) seconds before departing. The driver of a school bus or vehicle is subject to the same requirement when operating within 100 feet (30 meters) of a school and is prohibited from idling more than five (5) minutes at each stop beyond schools - such as parking or maintenance facilities, school bus stops, or school activity destinations. The driver of a transit bus or other commercial motor vehicle is prohibited from idling more than five (5) minutes at each stop within 100 feet (30 meters) of a school. Idling necessary for health, safety, or operational concerns is exempt from these restrictions. In addition, the ATCM requires the motor carrier of an affected bus or vehicle to: ensure that

drivers are informed of the idling requirements, track complaints and enforcement actions, and keep records of these driver education and tracking activities.

In September 2003, California's solid waste collection vehicle rule was passed to reduce the harmful health impacts of exhaust from diesel-fueled waste collection trucks. The goal of the solid waste collection vehicle regulation (SWCV) is to reduce cancer-causing particulate matter and smog-forming nitrogen oxide emissions from these trucks by requiring owners to use CARB verified control technology that best reduces emissions, following a phased-in schedule from 2004 through 2010.

On July 26, 2007, the CARB adopted a regulation to reduce diesel particulate matter and oxides of nitrogen (NOx) emissions from in-use (existing) off-road heavy-duty diesel vehicles in California. Such vehicles are used in construction, mining, and industrial operations. This legislation prevents the idling of heavy-duty vehicles used on construction sites, which includes concrete trucks, scrapers, graders, dump trucks, back hoes, loaders, cranes, concrete pump trucks etc. The anti-idling regulations restrict idling to no more than five (5) minutes

4 Inspection and enforcement of the new regulations

Currently, enforcement is performed by the CARB. The CARB is divided into two geographic sections, splitting the responsibility for the monitoring and inspection between northern and southern California. In the future, the ARB intends to shift inspection responsibilities from the State to local government agencies, counties and local municipalities. Prior to the September date, field audits were only conducted when complaints of excessive idling were filed by the public with the CARB. Responsibility for inspection and enforcement will eventually fall upon local government agencies similar to the National Pollutant Discharge Elimination System (NPDES) permit, where the State Water Resources Control Board shifted inspection and enforcement responsibilities to local government agencies for Storm Water Pollution Prevention Plans (SWPPPs).

5 Effect on construction equipment operations

As with any new regulation, the full effects on operations are difficult to foresee. According to the current legislation idling is acceptable under the following situations:

- delays in traffic congestion
- inspecting, testing, servicing, repairing or diagnosing vehicles
- verification that the vehicle is in safe operating condition
- warming engine to operating temperature
- operating a power take-off device (e.g. concrete mixing, lifting, etc.)
- immobility caused by adverse weather conditions or mechanical failure
- queuing for construction operations (required to be beyond 100 feet from any residential area)
- engine meets the optional NOx idling emission standard and vehicle is located beyond 100 feet (30 meters) from any residential area.

Construction firms interviewed for this research began instituting idling awareness programs by posting notices and by monitoring potential changes in operations. Much like the idea of having competent persons for safety, i.e. trench safety, firms are specifically training key foreman and superintendents to watch for and change idling patterns. Most firms surveyed acknowledged that they were aware of the regulations, but also agreed that they may not fully understand the potential affects on construction operations and the impact or the fines that they could be cited with.

Interviews with several general engineering and grading contractors determined that the increase in production cycle times is a major concern as construction equipment is shutdown and restarted repetitively during construction work shifts. Production cycle times for grading and earthmoving operations are expected to increase, as equipment is shutdown and restarted when loads are not ready for transfer. During the load cycle, an increased wait time may become a common factor of the overall production cycle. Delays may also potentially occur as construction equipment operators have a tendency to relax when the equipment is turned off.

In addition to the potential for increased cycle times, research revealed that changes in project specifications are likely to incorporate similar air quality control measures. Construction equipment engine sizes will be limited on projects, such that construction contractors will be required to use the minimum practical size to perform specific construction operations. Additionally, the number of construction equipment that can operate simultaneously may be limited through construction clauses reading “contractor shall employ efficient management practices to ensure that the smallest practical number is operating at any one time.”

Regulations currently limit idling of heavy-duty diesel trucks during loading and unloading; however, it is foreseeable that project specifications may begin to also include clauses requiring that auxiliary power units must be used whenever possible. In addition, project specifications are likely to include provisions that will limit the number of construction worker trips to and from the project site, by requiring carpooling and by providing for lunch onsite. Specification clauses and measures such as these may be required to be shown on grading and building plans.

6 Effect on construction cost estimating

The effect on construction cost estimating will occur in three ways: potential fines imposed by the CARB, increased production cycles, and equipment operating cost. Field audits for off-road heavy-duty diesel vehicles began September 15, 2008. Fines for first time offences are currently set at \$300 per violation and subsequent offences can be \$1,000 to \$10,000 per violation, depending on the severity of the idling. On October 9, 2008 the California EPA sent out a news release titled “Commercial diesel vehicle owners warned of idling restrictions – Grace Period Over”. The impact on larger operations could be substantial when a branch or division of a company in one location is fined first for the initial \$300 fine, and then another branch or division of the same company is cited in another location and

fined \$1,000 for the second offence. Fines can quickly accelerate upwards to become one very large fine.

On large earthmoving operations, which require grading and hauling, the increase in production cycle times will be quantifiable and therefore be able to be incorporated into the overall construction estimate. On smaller operations, quantifying the effect will be more difficult.

With regards to operating cost, the survey of construction contractors who operate diesel equipment in California all noted that they were not changing their hourly equipment costs. Some noted that they were predicting a savings on fuel costs and overall equipment hours decreasing. Their major concern revolved around wear on equipment components caused by the oil draining to the lower pan, and at start-up due to operators not waiting the necessary time to allow oil to move throughout the engine.

7 Effect on construction equipment fleet management

All equipment greater than 25 HP is affected. According to the U.S. EPA the non-road inventory for construction and mining equipment is 148,000 units, and non-road inventory for industrial equipment is 15,000 units. An equipment list of affected equipment is maintained and updated annually by the EPA. It is based on the MacKay construction equipment universe study (MacKay 2003), which analyzed the exhaust from different types and models of equipment. Standard construction equipment listed includes concrete trucks, scrapers, graders, dump trucks, backhoes, loaders, cranes, concrete pump trucks etc.

Fleet managers will be required to develop and establish a written idling policy, which should be provided to all vehicle operators in the fleet. Current legislation requires a fleet manager to report their fleet information as it exists on March 1, 2009 to the CARB. Reporting is free and there is no cost for fleets to register vehicles with the CARB. All vehicles subject to the regulation must be labeled with an Equipment Identification Number (EIN) assigned by the ARB after the initial reporting period. The label is to remain on the vehicle for the remainder of its life. Fleet managers are required to label vehicles within thirty (30) days of receiving the EIN and fleet managers are responsible for purchasing or creating labels in accordance with the CARB specifications, as labels are not issued by the CARB.

Vehicles owned by a rental or leasing company must be included. If rented or leased for less than one (1) year, the vehicle must be included in the owner's fleet. If rented or leased for greater than one (1) year, the vehicle may be excluded from the owner's fleet, and included in the lessee's fleet only if written into the lease agreement.

Alternatives to developing and establishing a written idling policy include incorporating exhaust retrofit technologies on existing fleet equipment. Systems that reduce specific pollutants can control one or more pollutants (such as PM or

NOx), or both, and are installed on an in-use vehicle. Systems include, but are not limited to:

- Diesel particulate filters (DPFs)
- Flow through filters (FTFs)
- Diesel Oxidation Catalysts (DOCs)
- Selective Catalytic Reduction (SCR)

Additionally, listed below are some available idle reduction technologies:

- Battery-Powered Auxiliary Power Systems
- Fuel-Fired Heaters
- Diesel-fueled Auxiliary Power Systems

8 Legal implications and legal challenges to the new regulations

In 2010, the American General Contractors of America (AGC) formally challenged the new regulations with a study that determined that there was no scientific reason for CA Emission Rule. The AGC report states that emissions from California's construction and other off-road diesel equipment are less than 28% of what state officials estimated. As a result, they stated that the CARB had no scientific justification for enacting the new rule requiring construction contractors to spend billions of dollars on their existing equipment. The AGC general counsel stated, "There is no scientific reason for the CARB to pursue its punishing off-road rule." The new emissions study was based primarily on data assembled by board officials, including how many pieces of off-road diesel equipment are in use and how much they run.

The CARB original report stated that there 192,000 pieces of equipment. The AGC claims there are only 157,000 pieces of this equipment in California, 7.5% of this equipment is low use. Other data came from the California State Board of Equalization, the U.S. Department of Energy and records detailing the actual hours worked by equipment operators. Using this data, the AGC stated that state officials overestimated emissions from the state's off-road diesel fleet. The AGC stated that emissions parallel the consumption of diesel fuel, and in 2009, off-road construction equipment burned only 164 million gallons of fuel. If the board's original estimates were correct, that number consumed would have totalled 581 million gallons. Therefore, without the new regulations, future emissions of nitrogen oxides will be 64% to 74% below the state's emissions goals through 2025.

The AGC report states that future emissions of PM will be more than 30% below the state's emissions goals through 2019 and has argued that board officials eliminate their rule's distinctions between fleets of different sizes, and to treat all fleets as if they were small. Additionally, the AGC noted that the cleanest construction equipment would not become available until 2015. The AGC stated that "Giving the industry a little more time won't put the state's air quality targets at risk, but it will give thousands of construction workers a chance to resume their careers and rebuild their lives." New construction employment figures for metropolitan areas in California underscore how badly the recession has hurt the

state's construction industry. The AGC noted that 26 out of 28 metro areas in the state lost construction jobs between March 2009 and 2010, while the other two remained stable. It is estimated that metropolitan Sacramento lost one out of every five-construction jobs (8,300 jobs, 19%) over the past year. El Centro California lost the highest percentage of jobs (29%, 500 jobs), while the Los Angeles area lost the most jobs (22,700 jobs, 18%). Napa County (27%, 800 jobs); San Luis Obispo County (26%, 1,500 jobs); and Chico (24%, 600 jobs) also experienced high rates of job losses. And Hanford and Yuba City escaped any construction job losses, reporting no change in employment levels between March 2009 and 2010.

9 Conclusions

Challenges to the new regulations have continued with challenges to the fines levied, impacts the rule has had on construction schedules that have required equipment operations to reschedule around school days, discrepancies between project specifications and legal ramifications generated, and change orders resulting from impacts on contractors means and methods. It is the author's opinion that there will be continued pressure on manufactures to reduce emission and provide reduced cost retrofits to existing equipment. The role of the construction manager and construction inspector will broaden to include monitoring and oversight of construction equipment for violations prior to their occurrences. Finally, the authors believe that others states may soon follow suit with similar legislation regarding construction activities.

10 References

- ARB (2008). Mobile Source (On-road) URL: <http://browser.grik.net/www.arb.ca.gov/msprog/msprog.htm>, viewed: 31/12/2007.
- ARB (2008). Mobile Source (Off-road) URL: <http://browser.grik.net/www.arb.ca.gov/msprog/msprog.htm>, viewed: 31/12/2007.
- Gaslan, D.P., (2008). Interview with Darryl P. Gaslan, Manager, California Air Resources Board, August 2008.
- Granite (2008). Telephone interview with Granite Construction, Wastonville, CA, October 2008.
- MacKay (2003). Landers, Kirk. America's fleet: the millennium census (MacKay and Co.'s 1999 Universe of Construction Equipment and Machinery study) Construction Equipment, June 1, 1999
- Ross, C. (2008). Interview with Charles Ross, Investigator/APS, California Air Resources Board, August 2008.

When clients do not pay: a critical analysis of the legal remedies available to the South African building and civil engineering contractors and consultants

Prof. Marthinus J Maritz¹, Adv Dirk C Robertson²

¹Head of Department, Department of Construction Economics,
Faculty of Engineering, Built Environment and Information Technology,
University of Pretoria, Pretoria, 0002, RSA

²Greene Consulting Engineers, Centurion, Pretoria, 0171, RSA

Emails: tinus.maritz@up.ac.za; dirk@greenegroup.co.za

Abstract:

Smooth cash flow ensures the effective delivery of projects and is fundamental to develop and sustain a healthy, professional and competitive construction industry. The adverse effect of late or non-payment of contractors and consultants are well known to all in the construction industry. Late and non-payment problems have forced countries like the United Kingdom, Singapore, New Zealand and certain states of Australia to introduce legislation to regulate the payment of contractors and consultants in terms of a building or a construction contract. From South African specific surveys conducted by the Construction Industry Development Board (CIDB) and Consulting Engineers South Africa (CESA) it appears that South African building and construction contractors and consultants have similar problems, if not more so, than their international counterparts when it comes to payment of work completed or services rendered. In light of the above this paper investigates the legal remedies available to the South African building and civil engineering contractors and consultants to enforce their right of payment for work completed/services performed, to establish how effective the enforcement is, and to suggest what possible solutions there are in order to improve payment practices in the South African building and construction industry.

Keywords:

Construction contracts, construction industry, legal remedies and payment.

1. Introduction

Good cash flow management is fundamental to develop and sustain a healthy, professional and competitive construction industry. Smooth cash flow ensures the effective delivery of projects. The prospect of prompt payment for completed works offers a strong incentive to consultants and contractors to deliver quality service.¹

The adverse effects of non payments and/or late payment by Employers on Contractors and Consultant are well known to all in the construction industry. Several related studies have

¹ Construction Industry Review Committee, Construct for Excellence: Report of the Construction Industry Review Committee, para.5.77, Hong Kong, 2001.

been conducted in developed countries which addressed the problems related to payment issues in the construction industry. Examples are the Latham Report (Latham, 1994) and the Egan Report (Egan, 1998). Late and non-payment problems have forced countries like the United Kingdom, Singapore, New Zealand and certain states in Australia, e.g. New South Wales to legislate their construction specific statutory payment security regime. These legislations purposefully enact provisions to address issues on prompt payment in the construction industry to eliminate poor payment practices and smooth the contractor's cash flow.

From a South African perspective it appears from a CIDB survey conducted by Marx² and the Consulting Engineers South Africa (CESA) that Contractors and Consultants have similar problems to their international counterparts when it comes to payment of work completed/services provided.

In the light of the above stated CIDB and CESA findings and the growing international trend to implement construction specific legislation in order to, *inter alia*, ensure/facilitate prompt payment practices in the construction industry, it has been decided to address, for this paper, the following problem statement:

“What are the legal remedies available to the South African Building and Civil Engineering Contractor and Consultant to enforce their right of payment for work completed/services performed and how effective are they in enforcing said right of payment.”

The structure of this paper is as follows:

- Section 2 summarises and addresses recent surveys and research relevant to payment practices and trends in the South African construction industry.
- Section 3 provides an overview of the common-law position of the South African building and civil engineering contractor.
- Section 4 provides an overview of the statutory position of the building and civil engineering contractor.
- Section 5 summarises and addresses the remedies available to contractors in South Africa to enforce payment in terms of the CIDB endorsed standard building and construction contract.
- Section 6 investigates other possible remedies in order to enforce payments that were not addressed in Section 3, 4 or 5.
- Sections 7 and 8 discuss the structure and findings of a survey that was conducted among South African contractors and consultants.

The research for this paper was delimited as follows:

² Construction Industry Development Board (CIDB): Results of the 2008 survey of the CIDB construction industry indicators and comparison with the 2005 & 2007 survey, South Africa, 2009.

- The study is limited to selected South African building and civil engineering contractors and consultants. Although extensive use was made of international literature for the literature survey, only South African contractors and consultants were interviewed.
- The legal remedies in terms of the following CIDB endorsed forms of contract for construction and building work were researched:
 - o JBCC Series 2000: Principal Building Agreement, edition 5.0, 2007, published by the Joint Building Contracts Committee, South Africa [PBA].
 - o The General Conditions of Contract for Construction Works, 1st edition, 2004, published by the South African Institution of Civil Engineering. [GCC 2004]
 - o New Engineering Contract 3: The Engineering and Construction Contract, June 2005, published by the Institution of Civil Engineers, United Kingdom. [NEC3 ECC]
 - o Conditions of Contract for Construction for Building and Engineering Works designed by the Employer, 1st edition, 1999, published by the International Federation of Consulting Engineers, Switzerland. [FIDIC Red]
- The legal remedies in terms of the following CIDB endorsed forms of contract for the provision of professional services were researched:
 - o CIDB Standard Professional Services Contract [CIDB PSC]; or
 - o New Engineering Contract 3: The Professional Service, June 2005, published by the Institution of Civil Engineers, United Kingdom. [NEC3 PSC].

2. Literature survey

2.1 Marx Report (2009)³

Construction Industry Indicators (CII's) have been developed by the Department of Public Works and the CIDB with assistance from the CSIR to play a useful role in developing a sustainable industry and to be adopted as a tool for improving performance in the South African construction industry.

The CIDB CII's measured above have been captured since 2003, and are currently being captured in partnership with the Department of Quantity Surveying and Construction Management of the University of the Free State. A full report was published in March 2009 on the results of the 2008 survey for projects completed in 2007.

Regarding payment delays experienced by contractors for the years 2004 to 2007, the following was reported:

There was a decrease from 24% to 9% in the number of all projects where payments were made timely within 14 days, if the 2004 results are compared with the 2007 results.⁴ In 2007

³ Marx HJ, 2009, *Results of the 2008 survey of the CIDB Construction Industry Indicators and Comparison with the 2005 & 2007 survey results*, University of the Free State

the private sector clients were the worst early payers, with payments made within 30 days on only 35% of their projects. The best performing client categories with 59% and 56% of project payments made within a month were the public private partnerships and provincial departments respectively.⁵ The percentage of projects with payments that took more than 30 days increased from 2004 to 2007 from 43% to 56%. In 2007 the contractors for 20% of all public corporation projects and 21% of all private sector and provincial department projects were only paid after 60 days. There is an encouraging reduction in the percentage of payments done later than 120 days from 13% to 3% if the 2004 and 2007 projects are compared. It is of great concern that only 44% of all contractors in 2007 were paid on time within 30 days.⁶ Contractors refrain from standing up to their contractual right to be paid on time for fear of losing job opportunities in the future.

With regard to consultants, the following was reported:

The consultants' fees were paid within 30 days for only 45% to 51% of all projects completed between 2004 and 2007.⁷ In 2007 the provincial and national departments were the slowest payers of fees with fees only paid after more than 60 days on 30% and 22% of all their projects respectively. This is followed by the regional/district councils and public private partnership client categories where the consultants were only paid after 3 months on 14% of all their projects. On 14% of all public private partnership projects the consultants were only paid 4 months after submission of fee accounts.⁸ The tendency for late payment of consultants has grown if the 2007 results are compared with the 2006 results.⁹

2.2 MBSA, 2009, "Payment Quality Survey" Draft Report¹⁰

In response to the Marx report, the Master Builders South Africa (MBSA) has performed a survey amongst its members to ascertain the prevalence of delayed or non-payment as well as the possible causes of the delayed or non-payment.

A questionnaire was distributed amongst its members.

From the draft report issued by the MBSA the findings from the survey is summarised as follows:

- Of all the projects surveyed, those in the Free State reported the most frequent delays in payments (93%) followed by the Northern Cape (74%). The province reporting the least frequent delays was Gauteng (24%). There were also three provinces in which some projects were listed as "never experiencing delays". These were Gauteng (43%); KwaZulu Natal (29%) and Mpumalanga (9%).
- Nationally, across all clients, (all projects included in the survey) 54% of projects were paid within 30 days; 26% between 31 and 60 days; 13% between 61 and 90 days; 3% between 91 and 120 days and 4% more than 120 days

⁴ Table 15 and 17, Marx HJ, *supra*

⁵ Table 15 and 17, Marx HJ, *supra*

⁶ Table 15 and 17, Marx HJ, *supra*

⁷ Table 37, 38 and 39, Marx HJ, *supra*

⁸ Table 39, Marx HJ, *supra*

⁹ Table 37, and 39, Marx HJ, *supra*

¹⁰ MBSA, 2009, "Payment Quality Survey" Draft Report

- The National Government and Private Sector seems to have the best payment record with 85% and 79% of payments respectively made within 30 days. The remaining 15% of National Government projects are paid between 31 and 60 days, while some payments in the private sector are delayed up to 120 plus days.
- Payments for local and provincial government projects mostly occur between 31 and 60 days. (45% and 44% respectively) with only 38% and 23% of payments respectively made within 30 days from date of invoice. 15% of the value of local government projects is only paid between 90 and 120 days after date of invoice. While 2% of the value of provincial government projects are only paid after 120 days.
- Sub-contractors seem to be affected by delayed payments as only 50% of payments are received within 30 days from date of invoice.

2.3 CESA Biannual Economic and Capacity Survey (June 2009)¹¹

An Economic and Capacity Survey is conducted by the Consulting Engineers South Africa (CESA) every 6 months. The purpose of this survey is to report on the prevailing conditions in the consulting engineering industry. The survey addresses aspects like financial indicators, human resources, capacity utilisation and competition in tendering and pricing. Questionnaires were distributed to all member firms of the Consulting Engineers South Africa.

According to the latest survey (June 2009), consulting engineers reported a percentage fee income outstanding for 90 days or more of 9.5%. The comparative figure for June 2007 was 10.3%, 11.3% for December 2007, 11.1% for June 2008 and 12% for December 2008.¹² If one looks at the employers, the situation is as follows:

For June 2009, 7.3% of fee claims submitted to central government is outstanding for 90 days or more. For provincial government the figure is 3.8%, local government 13.2%, state owned enterprises 1.4%, private sector 11.9% and foreign employers 13%¹³.

2.4 Maritz, MJ, 2007, “An investigation into the adjudication of disputes in the South African Construction Industry”¹⁴

The purpose of this paper was to provide an overview of the development of adjudication as an alternative dispute resolution process in South Africa and its effectiveness in solving disputes in the local construction industry. The following findings are relevant to this research:

- 63% and 26% of the respondents respectively agree and strongly agree that “*there exists a chronic problem of delayed and non-payment in the South African construction industry affecting the entire delivery chain.*”¹⁵
- 50% and 13% of the respondents respectively agree and strongly agree that “*allowing all disputed matters to come before adjudication would also reduce payment disputes.*”¹⁶

¹¹ Table 15: CESA, 2009, “CESA Biannual Economic and Capacity Survey”

¹² Table 15 *supra*

¹³ Table 15 *supra*

¹⁴ Maritz, M.J. 2007. *An investigation into the adjudication of disputes in the South African construction industry*. International Conference: RICS COBRA 2007, Atlanta, USA.

¹⁵ Table 1: Maritz, M.J. 2007

- 39% and 30% of the respondents respectively agree and strongly agree that “*South Africa should introduce a Construction Industry Payment and Adjudication Act similar to those in the UK, Australia, New Zealand and Singapore.*”¹⁷

2.5 Maiketso NC, Maritz MJ, 2009, “What are the requirements for the South African construction industry to fully utilise adjudication?”¹⁸

The purpose of this research was to investigate what the requirements are for the South African construction industry to fully utilise and benefit from adjudication. The researcher, *inter alia*, reviewed the contractual, institutional and legislative framework for adjudication in South Africa. The following findings are relevant to this paper:

- 75% of the respondents agreed that “*South Africa needs a “Payment and Adjudication Act” similar to that in the UK*”. This finding correlates with the Maritz paper as discussed above.¹⁹
- 60% of the respondents agreed that “*such legislation should address minimum payment terms, 90% agreed with statutory adjudication, and 95% agreed with remedy in case of non-payment and other countries.*”²⁰

3. Common-law position of building and civil engineering contractors

The general principals of the South African law apply to building and construction contracts. In the case of standard construction contracts and where contracts with identical or similar wording have been interpreted by the Courts, the Courts will consider previous decisions in its judgements.²¹

In general the following principals apply where a contractor claim for payment for work done in terms of a *locatio conductio operis*²². The contractor needs to allege and prove:

- *The terms of the contract relied upon:* If the terms involve a negative the onus is on the contractor to prove the negative. For example if the employer alleges that the contract comprises of x + y and the contractor alleges it comprises of x only, the contractor must prove that y is not a term of the contract.
- *The work that had to be performed:* It is usually an implied term of the contract that the contractor will use materials that are suitable for the purpose of the works.²³ If the employer supplies the material to the contractor for the works, it may be that the employer accepts the risks of inferior material. It is another implied term of the contract that the contractor will perform the work in a workmanlike fashion. The level

¹⁶ Table 2, Maritz MJ, *supra*

¹⁷ Table 3, Maritz MJ, *supra*

¹⁸ Table 2: Maiketso NC and Maritz, M.J., 2009. “What are the requirements for the South African construction industry to fully utilise adjudication?” International Conference: RICS COBRA 2009, Cape Town

¹⁹ Table 2: Maiketso NC and Maritz, M.J, *supra*

²⁰ Table 2: Maiketso NC and Maritz, M.J *supra*

²¹ Smith v Mouton 1977 (3) SA 2 at 12

²² Harms, L T C. 1998. *Amler’s precedents of pleadings*. 5th ed. Durban: Butterworths, 265

²³ Colin v De Guisti 1975 (4) SA 223

of skill and diligence to be employed is that possessed and exercised by other members of the trade to which the contractor belongs.²⁴

- *The remuneration applicable:* The Contractor must allege and prove 1) that the remuneration was, in terms of the contract payable, and 2) the amount of the remuneration payable. If the contract is silent with regard to remuneration, remuneration will be payable and should be fair and reasonable. Where the contract was silent on whether remuneration will be payable or not, remuneration will be payable. In such a case, it is an implied term that the remuneration will be reasonable. If in dispute, the onus will be on the contractor to prove that nothing was said concerning remuneration.
- *Performance:* The contractor must allege and prove that he has done all that was required to be done in terms of contract²⁵

4. Statutory position of building and civil engineering contractors

South Africa does not have 'construction specific' legislation to address the need for prompt payment of building and civil engineering contractors. The Public Finance Management Act, 1999 (PFMA) determines that all contractual obligations (and accounts) must be settled within 30 days from its receipt. [Section 38(1)(f) read with Part 4, Regulation 8.2.3 of the Regulations.] These provisions are mandatory and an accounting officer of the guilty official may be found guilty of an offence in terms of the PFMA.

In several other countries acts, addressing this need for prompt payment in the construction industry were endorsed. Acts and the respective countries and states, which enacted them to address the problem of late and non-payment, are:

- Housing Grants, Construction and Regeneration Act 1996 – United Kingdom;
- Building and Construction Industry Security of Payment Act 1999 (NSW) – New South Wales, Australia;
- Building and Construction Industry Security of Payment Act 2002 (VIC) – Victoria, Australia.
- Building and Construction Industry Payment Act 2004 (QLD) – Queensland, Australia;
- Construction Contracts Act 2002 – New Zealand;
- Building and Construction Security of Payment Act 2004 – Singapore.

²⁴ Randaree NNO v WH Dixon & Associates 1983 (2) SA (1)

²⁵ BK Tooling (Edms) Bpk v Scope Precision Engineering (Edms) Bpk 1979 (1) SA 391

5. Remedies to enforce payment in terms of the CIDB endorsed standard building and construction contracts

5.1 Right to interim and final payment certificates

A contractor's obligation to complete the work is generally indivisible. The mere completion of a specific subdivision of the work does not entitle a contractor for payment of the work done. In the absence of contractual provisions that does allow for interim payments, a claim for partially completed work done would be met with the *exceptio non adimpleti contractus*.²⁶ Only upon completion of the work as a whole would the contractor be entitled to payment.

As a rule the average contractor does not have/command the necessary resources to complete a construction contract before requiring payment for the work completed. In order to provide the contractor with the necessary cash flow to complete the work, most construction contracts provides for the issue of interim payment certificates. In such a certificate the employer's representative records his reasonable, but only approximate, assessment of the total of work executed and materials supplied up to a given date

This certificate entitles the contractor to payment of the amount certified within a set number of days. Failing payment, the contractor may sue the employer on the strength of the certificate, and the strength of the certificate alone. The claim would be one based on the express terms of the contract. It is not an enrichment claim even though the amount may be certified as a "reasonable estimate of the total of the work and materials"²⁷.

From the comparison made between the four CIDB endorsement construction contracts, the following general observations were made:

- Payment certificates are certified by independent persons.
- The frequency of interim payment certificates are defined in the contract.
- It is clear when the interim payment certificates should be issued by the Employer.
- All four CIDB endorsed construction contracts allows for payment of materials on site and only the GCC 2004 does not expressly allow for payment of material off site.
- It is clear when the interim payment certificates should be paid by the Employer.
- It is clear when the final payment certificates should be issued by the Employer.
- It is clear when the final payment certificates should be paid by the Employer.

²⁶ Qwa Qwa Regeringsdiens v Martin Harris & Seuns OVS 2000 (3) SA 339

²⁷ Simmons v Bantoesake Administrasieraad (Vaaldriehoekgebied) 1979 1 SA 940 (T)

5.2 Right to interest on late payments

If the employer fails to pay money due under the contract the contractor may elect to charge interest on the amount due. The easiest way to recover interest would be in the case where the contract has express provisions that provides for the payment of interest in specific circumstances at a quantified rate²⁸.

From a comparison made between the four CIDB endorsed construction contracts, the following general observations can be made relevant to interest on late payment:

- All four the CIDB endorsed construction contracts contain provisions that provides for the payment of default interest (“finance charges” in the case of the FIDIC Redbook).
- The circumstances when default interest may be charged are defined.
- The time from when the interest runs is defined.
- The rate of interest chargeable is defined in the contracts.

5.3 Payment guarantee

Relevant to the South African construction industry, a payment guarantee could be defined as a contractual undertaking by a third party, the guarantor, towards the contractor, that the guarantor will pay to the contractor the amount of works done under the construction contract, up to the guaranteed amount or a percentage of the price of the works done, in case the employer defaults in its payment obligations.

Of the four CIDB endorsed contract documents, only the FIDIC Redbook and the JBCC PBA 2007 contracts expressly provides for the use of payment guarantees. See clause 3.1, JBCC PBA 2007 and the example clause on page 17 of the guidance notes of the FIDIC Redbook. Both contracts have *pro forma* payment guarantee forms that could be used by the parties.

It was a specific objective of the research to determine the frequency of use of payment guarantees in South Africa.

5.4 Right to terminate

The Contractors have to complete the complete construction contract before he will be entitled for payment for the work done. In the case where the work is only partially complete, the contractor’s claim for payment of the partially completed work could be met with a counter claim from the employer based on *exceptio non adimpleti contractus*.²⁹ Following from this common law position, a contractor can not abandon site if the employer fails to pay the contractor for partially completed work.

All four the CIDB endorsed contract documents contain provisions that allow for the suspension of work and/ or the cancellation of the contract in the case of failure by the employer to pay interim payment certificates.

²⁸ Another cause of action for a claim for outstanding interest would be that the defendant was placed in mora on the date from which the interest is claimed. See *Standard Bank of SA Ltd v Lotze* 1950 (2) SA 698 (C)

²⁹ *Hamman v Nortje* 1914 AD 293 at 296.

The following aspects should be considered in the case where a contractor wants to leave site/terminate the contract as a result of the employer's failure to make payment for work completed.

- When the party wishes to enforce a cancellation clause, the conditions for its implementation have to be strictly complied with.³⁰
- In the case where it is required by the contractor to give the employer notice of his intention to cancel the contract as a result of the employer's failure to make the required payment, the notice to be given should be an express, extra-judicial announcement and such notice cannot be implied or given by notice of motion.³¹
- In the absence of a contractual cancellation clause, a contractor will not be able to cancel a contract if an employer fails to make an interim payment. The rationale for this is as follows:
 - In the case of an interim certificate, the contractor has not completed the work in total, and until he has not completed the work the contractor has not performed in terms of the contract.
 - If the contractor abandon site, as a result of the non-payment by the employer, the contractor will be in material breach of his obligations to deliver the completed work to the employer.
 - If the contractor cancels the contract, his cancellation may be held to be a repudiation of the contract, in other words, an indication that he no longer intends to be bound by the terms of the contract, and this would afford the employer the right either to accept such repudiation, bringing the contract to an end, or to refuse to accept the repudiation in which case the contract remains alive and both parties are obliged to continue to honour their obligations to each other. In either event, the employer would be entitled to such damages as he could show he has sustained as a consequence of the repudiation.

6. Other remedies to enforce payment

6.1 Evidence from the Employer regarding financial arrangements for the project

The FIDIC Red book provides for evidence to be provided by the Employer to the contractor whereby, inter alia, the Employer prove that it has access to/has the funds necessary to pay the contract price. Clause 2.4 Employer's Financial Arrangements reads as follows:

³⁰ Die Wet NO v Uys No en andere 1998 (4)

³¹ Shroobree NO Simon 1999 (2) SA 488 (SE) See also clause 55.1 and 56.1 of GCC 2004 and clause 36.3 and 38.2 of JBCC.

“The Employer shall submit, within 28 days after receiving any request from the Contractor, reasonable evidence that financial arrangements have been made and are being maintained which will enable the Employer to pay the Contract Price (as estimated at that time) in accordance with Clause 14 [Contract Price and Payment]. If the Employer intends to make any material change to his financial arrangements, the Employer shall give notice to the Contractor with detailed particulars.”

The mechanism for the provision of evidence by the Employer is technically not a remedy to enforce prompt payment by the Employer, but it can certainly be regarded as a mechanism that will assist the contractor to identify, upfront, any possible risks pertaining to capability of the Employer to pay for work completed by the contractor.

Similar provisions could not be found in the JBCC PBA 2007, NEC ECC and GCC 2004 documents.

6.2 Order of court

It is common practice in South Africa to make an arbitration award an order of the court. An arbitration award can be made an order of the court of competent jurisdiction by any party³². An award that has been made an order of the court can be enforced in the same way as any judgement or order to the same effect. After an award has been made an order of the court, the party enforcing its rights can, for example, issue a writ of execution to be executed by the sheriff of the court.

A contractor or consultant armed with an order of the court, resulting from a successful arbitration award or any other procedure, can enforce the order by applying for the following:

- A finding and order of contempt of court by and committal of the defaulting employer, or
- A writ of execution followed by an attachment of assets of the defaulting employer, and sale thereof.

The procedure whereby a successful litigant can enforce the judgement is crucial in the legal process. The idea that a court must be able to give an effective judgement is the major foundational concept to determine jurisdiction³³. Effectiveness is inextricably linked to execution, for if the court through its officers has no control over either the defendant or the defendant's property, any judgement granted in favour of a plaintiff would amount to little more than a declaration of a theoretical benefit.

Up to June 2008 it was not possible to obtain an order of execution against the property of the State. Section 3 of the State Liability Act, 1957 (in its current form) provides that:

“No execution, attachment or like process shall be issued against the Defendant or Respondent in any action or proceedings or against the property of the State...”

³² Section 31 of the Arbitration Act 42 of 1965

³³ Eilon v Eilon 1965 1 SA 703 (A) 725.

On 2 June 2008 and in the matter of *Dingaan Hendrik Nyathi v The MEC, Department of Health, Gauteng and Others* the Constitutional Court declared section 3(2) of the State Liability Act, which limits the State's liability, unconstitutional. The effect of this judgement is that the State is likely to have to comply with judgements sounding in money on demand in the same way as any ordinary debtor.

Failure of the State Attorney to honour a judgement debt has caused the Constitutional Court not only to declare unconstitutional certain legislation designed to limit the state's liability, it ordered the Department of Justice to report back to it with a plan to ensure the speedy settlement of all unpaid court orders. It was estimated by the court that approximately 200 court orders was unpaid to date of the judgement. The Court further expressed strong views against officials who have become a law unto themselves and openly violate peoples' rights.

The declaration of invalidity was suspended until 3 June 2009 to allow the State sufficient time to pass legislation that will provide for the effective enforcement of court orders against the State.

In August 2009 the Minister for Justice and Constitutional Development applied for an extension of the period of suspension of the order of constitutional invalidity. The extension of the period of suspension was extended in order to allow the State additional time to pass legislation that will provide for the effective enforcement of court orders against the State. As part of an interim order, the court had laid down the requirements for execution steps against the state. In the case of *The Department of Transport & Roads, North West Provincial Government and 5 others v Kafulani Civils North (Pty) Ltd and 4 others*, case number 70680/09, Bosman AJ confirmed and summarised said requirements as follows:

- a. "To summarise, after a final order has not been satisfied within 30 days of the date of judgement, the judgement creditor may serve the Court in terms of the Rules on the relevant National or Provincial Treasury, the State Attorney, the Accounting Officer of the relevant National and Provincial Department as well as the Executive Authority of the department concerned.
- b. The order served should be accompanied by a certificate by the Registrar certifying that no appeal, review or rescission proceedings are pending in respect thereof.
- c. The relevant treasury shall forthwith within 14 days of service of the order, cause the debt to be settled, or make acceptable arrangements with the judgement creditor for settlement of the debt.
- d. If (c) above has not been complied with, the judgement creditor may apply for a writ of execution in terms of Rule 45 of the Rules of Court against the movable property owned by the State and used by the relevant department.
- e. The Sheriff shall, pursuant to the writ, attach but not remove the identifiable moveable property.
- f. In the absence of any application contemplated in paragraph (g) (hereunder), the Sheriff may, after the expiration of 30 days from the date of attachment, remove and sell the attached movable property in execution.

- g. During the periods referred to in (f), any party having direct and material interest may apply to the Court which granted the order, for a stay on grounds that the execution of the assets is not in the interest of justice.”

In June 2010, the period when this paper was written, no changes to current legislation or new legislation has been effected to allow for the issue of a writ of execution against the State.

7. Questionnaire Survey

In addition to the literature study for this paper, a questionnaire survey was conducted amongst randomly selected consultants and contractors in the South African construction industry. The questionnaire was designed to be brief, concise and straightforward to encourage a high response rate from the potential respondents. Two different sets of questionnaire forms were used in the survey; one for the consultants and the other for the contractors.

Both the questionnaire for the consultants and the contractors comprised of five distinct sections:

- Section 1 established the background information of the respondent.
- Section 2 established the levels of use and knowledge of the respective CIDB endorsed contracts. For the consultant's questionnaire the respondents were requested to rate their knowledge and use of the CIDB PSC 2005 and the PROCSA 2009. For the contractors questionnaire respondents were requested to rate their knowledge and use of the JBCC PBA 2007, GCC 2004, NEC3 ECC and FIDIC Redbook
- Section 3 was used to rate the sufficiency of the remedies in terms of the CIDB endorsed contracts. For the consultant's questionnaire the respondents were requested to rate the perceived sufficiency of certain contractual clauses to ensure prompt payment of professional fee accounts. Clauses pertaining to interim monthly accounts, interest on late payments, written proof of funding available for ongoing requirements, payments guarantees, suspension of services and termination of services were rated. For the contractors questionnaire clauses pertaining to issue of interim payment certificates, payment of interest on late payments, payment guarantees, suspension and / or termination of work were rated
- Section 4 rated the attitude and perceptions of the respondent regarding the effectiveness of litigation in securing payment for professional services and construction work duly executed.
- Section 5 proposed possible solutions on how to improve current payment practices in the South African construction industry. For both the consultant's and the contractor's questionnaire the respondent's opinion regarding the introduction of statutory prompt payment provisions were measured. In addition the respondent was also requested to indicate what the prompt payment process should provide for a minimum.

Considering that there would be a wide range of expected or possible responses, questions that were open-ended were avoided. For most of the questions a 5-point Likert scale ranging

from 'very low sufficiency' to 'very high sufficiency' or 'strongly approve' to 'strongly disagree' were used. The questionnaire was accompanied by a covering letter which explained the reasons for and background of the research.

8. Trends indicated by the survey

Table 1 to 8 is a summary of the trends indicated by the survey.

Table 1: The use of interest on late payment/finance charges provisions and provision of payment guarantees (Consultant)

Findings pertaining to the use of interest on late payment/finance charges provisions and provision of payment guarantees						
			Never	Rarely	Often	Always
A1	How often do you/your company charge interest on late payment of professional fee accounts.	No	9	9	5	2
		%	36	36	20	8
A2	How often do you/does your company insist on the provision of payment guarantees from the client?	No	6	6	1	
		%	43	43	7	7

Table 2: The use of interest on late payment / finance charges provisions and provision of payment guarantees (Contractor)

Findings pertaining to the use of interest on late payment / finance charges provisions and provision of payment guarantees						
			Never	Rarely	Often	Always
A1	How often do you/your company charge interest on late payment of professional fee accounts.	No	28	28	13	7
		%	37	37	17	9
A2	How often do you/does your company insist on the provision of payment guarantees from the client?	No	27	27	18	0
		%	38	38	25	0

Table 3: The attitude and perceptions regarding the effectiveness of litigation in South Africa
(Consultants)

Findings with regard to the attitude and perceptions regarding the effectiveness of litigation in South Africa							
			Strongly disagree	Disagree	Neutral	Agree	Strongly agree
A1	Litigation takes a long time and a successful verdict may often come too late to prevent financial harm to your company.	No	0	0	0	8	24
		%	0	0	0	25	75
A2	Because of the high non-recoverable costs of litigation, a successful verdict may often be a paper victory. (A worthless judgement).	No	2	2	0	7	23
		%	6	6	0	21	68
A3	State departments and municipalities often ignore an order of court and therefore a successful verdict together with an order of court may often be a paper victory. (A worthless judgement).	No	0	0	9	7	16
		%	0	0	28	22	50
A4	Once you / your company have/has instituted litigation against a party (including private companies, state departments and municipalities), chances are slim that you will get further work from that party in future.	No	0	0	1	6	25
		%	0	0	3	19	78

Table 4: The attitude and perceptions regarding the effectiveness of litigation in South Africa
(Contractors)

Findings with regard to the attitude and perceptions regarding the effectiveness of litigation in South Africa							
			Strongly disagree	Disagree	Neutral	Agree	Strongly agree
A1	Litigation takes a long time and a successful verdict may often come too late to prevent financial harm to your company.	No	0	0	1	3	38
		%	0	0	2	25	73
A2	Because of the high non-recoverable costs of litigation, a successful verdict may often be a paper victory. (A worthless judgement).	No	0	0	3	5	34
		%	0	0	6	29	65
A3	State departments and municipalities often ignore an order of court and therefore a successful verdict together with an order of court may often be a paper victory. (A worthless judgement).	No	0	0	6	30	16
		%	0	0	12	58	31
A4	Once you/your company have/has instituted litigation against a party (including private companies, state departments and municipalities), chances are slim that you will get further work from that party in future.	No	0	0	4	25	23
		%	0	0	8	48	44

Table 5: Possible solutions to improve current payment practices in the South African construction industry (Consultants)

Findings on possible solutions to improve current payment practices in the South African construction industry							
			Strongly disagree	Disagree	Neutral	Agree	Strongly agree
A1	Statutory prompt payment provisions will improve late payment practices in the South African construction industry	No	22	13		17	0
		%	66	38		50	0
A2	A commission should be established to investigate errant payments.	No	55	21			8
		%	14	14	5	46	22
A3	Councils/professional bodies for professional consultants in the South African construction industry should be enabled to suspend the licences / membership of defaulting main consultants (main consultants that do not promptly pay sub-consultants).	No	44	910			8
		%	11	11	26	29	23

Table 6: Possible solutions to improve current payment practices in the South African construction industry (Contractors)

Findings on possible solutions to improve current payment practices in the South African construction industry						
		Strongly disagree	Disagree	Neutral	Agree	Strongly agree
A1	Statutory prompt payment provisions will improve late payment practices in the South African construction industry	No	000	14		36
		%	000	28		72
A2	A commission should be established to investigate errant payments.	No	000	8		43
		%	000	16		84
A3	The CIDB should be enabled to suspend the registration of defaulting main contractors (main contractors that do not promptly pay sub-contractors)	No	000	8		43
		%	000	16		84

Table 7: Possible provisions of a statutory prompt payment provisions
(Consultants)

Findings with regard to the question: “To what extent should statutory prompt payment provisions provide for the following”			To a very small extent	To a small extent	To an average extent	To a large extent	To a very large extent
A1	Statutory adjudication or a similar dispute resolution mechanism to ensure swift dispute resolution of payment disputes.	No	2	22	4		7
		%	7	77	52		26
A2	A right to regular payment.	No	0	03	7		22
		%	0	09	22		69
A3	A right to a defined time frame for payment.	No	0	02	0		20
		%	0	06	31		63
A4	A right to interest on late payments.	No	1	12	1		17
		%	3	36	34		53
A5	A restriction of the right to set-off or withhold sums due.	No	1	14	8		13
		%4		4	15	30	48
A6	Provision for a mechanism that will ensure that a client cannot withhold payment from a consultant unless he has given an effective notice of his intention to withhold such payment.	No	0	02	8		22
		%	0	06	25		69
A7	A right to suspend services coupled with the right to reimbursement and additional time as a result of the suspension.	No	1	10	8		23
		%	3	30	24		70
A8	Prohibition of “pay when paid” clauses.	No	2	23	7		15
		%7		7	10	24	52

Table 8: Possible provisions of a statutory prompt payment provisions
(Contractors)

Findings with regard to the question: “To what extent should statutory prompt payment provisions provide for the following”							
			To a very small extent	To a small extent	To an average extent	To a large extent	To a very large extent
A1	Statutory adjudication or a similar dispute resolution mechanism to ensure swift dispute resolution of payment disputes.	No	1	10	23		28
		%	2	20	43		53
A2	A right to regular payment.	No	0	02	22		28
		%	0	04	42		54
A3	A right to a defined time frame for payment.	No	0	02	19		31
		%	0	04	37		60
A4	A right to interest on late payments.	No	1	10	33		18
		%	2	20	62		34
A5	The provision of escrow accounts, or similar trust accounts, to the benefit of the contractor and for retention money retained from the contractor.	No	1	14	28		19
		%	2	28	53		36
A6	A restriction of the right to set-off or withhold sums due.	No	0	02	21		29
		%	0	04	40		56
A7	Provision for a mechanism that will ensure that a client cannot withhold payment from a consultant unless he has given an effective notice of his intention to withhold such payment.	No	0	04	15		33
		%	0	08	29		63
A8	Statutory provision for a contractor’s lien.	No	1	15	13		33
		%	2	29	25		62
A9	A right to allow for stage payments for material in advance of their arrival on the construction site.	No	1	14	14		33
		%	2	28	26		62
A10	A right to suspend services coupled with the right to reimbursement and additional time as a result of the suspension.	No	0	06	14		32
		%0		0	12	27	62
A11	Prohibition of “pay when paid” clauses.	No	0	04	14		34
		%	0	08	27		65

9. Summary and conclusions

From the literature survey the following is concluded:

- There exists a chronic problem of delayed and non-payment in the South African construction industry.
- From 2004 to 2009 there was no substantial improvement in the problem of delayed and non-payment in the South African construction industry.
- The statements above are applicable to both the South African construction consultant and the contractor.
- South Africa does not have construction specific legislation to address the need for prompt payment of building and civil engineering contractors.

From the analyses conducted on the four CIDB endorsed construction documents the following general observations are made:

Interim Payment Certificates

- Payment certificates are certified by independent persons.
- The frequency of interim payment certificates is defined.
- It is clear when the interim payment certificates should be issued by the Employer.
- All four CIDB endorsed construction contracts allows for payment of materials on site and only the GCC 2004 does not expressly allow for payment of material of site.
- It is clear when the interim payment certificates should be paid by the Employer.
- It is clear when the final payment certificates should be issued by the Employer.
- It is clear when the final payment certificates should be paid by the Employer.

Interest on Late Payment

- All four the CIDB endorsed construction contracts contain provisions that provides for the payment of default interest (“finance charges” in the case of the FIDIC Redbook).
- The circumstances when default interest may be charged are defined.
- The time from when the interest runs is defined.
- The rate of interest chargeable is defined.

Payment Guarantee

Of the four CIDB endorsed contract documents, only the FIDIC Redbook and the JBCC PBA 2007 contracts expressly provides for the use of payment guarantees.

Some of the most relevant trends indicated by the Questionnaire Survey are:

- 72% of the consultants and 74% of the contractors interviewed never or rarely charges interest on late payments.
- 86% of the consultants and 76% of the contractors interviewed never or rarely insist on the provision of payment guarantees.
- Both groups of consultants and contractors surveyed regard litigation in South Africa as ineffective in securing payment for professional services and construction work duly executed.
- 12% of the consultants interviewed disagreed with the statement that statutory prompt payment provisions will improve late payment practices in the South African construction industry.
- 100% of the contractors interviewed agreed with the statement that statutory prompt payment provisions will improve late payment practices in the South African construction industry.
- 68% of the consultants and 100% of the contractors interviewed agreed with the statement that a commission should be established to investigate errant payments.
- 22% of the consultants interviewed disagreed with the statement that councils/professional bodies for professional consultants in the South African construction industry should be enabled to suspend the licences/memberships of defaulting main consultants (main consultants that do not promptly pay sub-consultants).
- 100% of the contractors interviewed agreed with the statement that the CIDB should be enabled to suspend the registration of defaulting main contractors (main contractors that do not promptly pay sub-contractors).

10. References

Ameer Ali, N.A.N. 2005. *Payment in the Construction Industry – towards Zero – Default*. QS National Convention, 10-11 August 2005, Kuala Lumpur

Butler, D & Finsen, E. 1993. *Arbitration in South Africa: law and practice*. Cape Town: Juta & Co Ltd

CESA, 2009. *CESA Bi-annual Economic and Capacity Survey*. Johannesburg

Cheng, T, Soo, G, Kumaraswamy, M & Jin, W. (n. d.) *Security of payment for Hong Kong Construction Industry Workable alternatives and suggestions*. Hong Kong

Finsen, E. 2007. *The building contract - A commentary on the JBCC Agreements*. 2nd ed. Cape Town: Juta & Co Ltd

Harms, L.T.C. 1998. *Amler's precedents of pleadings*. 5th ed. Durban: Butterworths

Joubert, W.A. 2003. *The law of South Africa*. Vol 2, Part 1. 2nd ed. Durban: Butterworths

Latham, M. 1994. *Constructing The Team. Final Report of the Government / Industry Review of Procurement and Contractual Arrangements in the UK Construction Industry*. London: HMSO

Master Builders South Africa. 2009. *Payment Quality Survey*. Draft report prepared by Industry Insight. Johannesburg

McKenzie, H S. 1994. *The law of building and engineering contracts and arbitration*. 5th ed. Cape Town: Juta & Co Ltd

University of the Free State. Department of Quantity Surveying and Construction Management. Dr Marx, H J. 2009. *Results of the 2008 survey of the Construction Industry Indicators and comparison with the 2005 & 2007 survey results*. Bloemfontein

Legal Considerations in the United States Associated with Building Information Modeling

Lonny Simonian¹ and Thomas Korman¹

¹ California Polytechnic State University,
Construction Management Department
San Luis Obispo, CA 93407-0284, USA

Email: lsimonia@calpoly.edu; tkorman@calpoly.edu

Abstract:

Building Information Modeling (BIM) offers the potential for significant savings in the cost and time required to construct a building; however, there are several legal issues associated with its use:

- ◇ Ownership of the BIM model and data - An owner may believe that they are entitled to own the model; however the rights of the architect, design engineers, equipment suppliers need to be protected as well.
- ◇ Allocation of Risks - Taking responsibility for updating BIM data and ensuring its accuracy involves a significant amount of risk. Issues of complicated indemnities by BIM users and limited warranties and disclaimers of liability by designers will need to be resolved.
- ◇ Privity of Contract and Third Part Reliance - The use of a collaborative model lessens the likelihood that a designer may claim the lack of privity of contract as a legal defense.
- ◇ Professional Design Responsibility - It will be very difficult to ensure that licensed design professionals will always in charge of the creation and modification of the data that forms a digital model.
- ◇ Standard of Care - While it appears that the author of each individual design element should have ownership of the copyright to the information and documentation within the model, many different parties may have the ability to change the design created by another party.
- ◇ Spearin Doctrine - The BIM collaborative process during the design phase of the project could deprive the contractor from the protection of design errors provided by this doctrine.
- ◇ Economic Loss Rule - The use of a collaborative building model will be a factor tending to support a contractor's claim that it should be able to recoup its economic losses absent a contract with the defendant.

This paper will examine the current literature to date on the legal implications associated with the use of BIM, and propose an implementation plan to transition to BIM.

Keywords:

Building Information Modelling, Privity of Contract, Standard of Care, Spearin Doctrine, Economic Loss Rule

1) Introduction

There are several different definitions for Building Information Modeling (BIM). These definitions vary by type of organization and their specific work emphasis. For example, from a design perspective, the American Institute of Architects¹ (AIA) defines BIM as the digital representation of the physical and functional characteristics of a project, and that Building Information Modeling refers to the process and technologies used to create a Building Information Model (also abbreviated as BIM). From the construction side, The Associated General Contractors of America² (AGC) defines BIM more narrowly, as “the development and use of a computer software model to simulate the construction and operation of a facility”

Another definition for BIM includes that proposed by Van Nederveen³:

“a model of information about a building (or building project) that comprises complete and sufficient information to support all lifecycle processes and which can be interpreted directly by computer applications. It comprises information about the building itself as well as its components, and comprises information about properties such as function, shape, material and processes for the building life cycle”.

Finally, the National Institute of Building Sciences⁴ (NIBS) offers the following definition:

“Building Information Model or BIM utilizes cutting edge digital technology to establish a computable representation of all the physical and functional characteristics of a facility and its related project/life-cycle information, and it is intended to be a repository of information for the facility owner/operator to use and maintain throughout the life-cycle of a facility”.

¹ American Institute of Architects (AIA), 2008. *Document E202 – Building Information Modeling Protocol Exhibit*. Washington DC: AIA

² Associated General Contractors of America (AGC), 2006. *The Contractor's Guide to Building Information Modeling*, AGC, Arlington, Virginia

³ Van Nederveen, S., Beheshti, R. and Gielingh, W., 2009. Modeling Concepts for BIM. In: Underwood, J., and Isikdag, U., eds., *Handbook of Research on Building Information Modeling and Construction Informatics: Concepts and Technologies*, Hershey, PA: Information Science Reference, 1-18.

⁴ National Institute of Building Sciences (“NIBS”), 2007. *United States National Building Information Modeling Standard: Version 1 – Part 1: Overview, Principles and Methodologies*. NIBS.

Typically BIM uses three-dimensional, real-time, dynamic building modeling software to increase productivity in building design and construction. BIM is a process which goes far beyond switching to new software. It requires changes to the definition of traditional architectural phases and more data sharing than most architects and engineers are used to performing. The process produces the Building Information Model, which encompasses building geometry, spatial relationships, geographic information, and quantities and properties of building components.

Despite the significant benefits associated with BIM, there are several legal issues and risks which the design and construction industry has not addressed properly. In the United States, the legal system principally is concerned with individual rights and responsibilities, whereas BIM is essentially a collaborative effort amongst several different entities. This difference in focus between personal versus collective betterment results in a tension between an individual firm's need to tightly define their responsibilities and limit their reliance on others, and the need to promote collaboration and encourage reliance on information stored within a shared building model.

These issues have been categorized⁵ into three major groups: commercial, technical, and legal. This paper examines some of the legal issues and risks associated with the use of a BIM.

2) Ownership of the BIM Model and Data

The relationship between engineering and construction firms can vary widely, depending upon past working relationships. In general, however, engineering firms are reluctant to supply reference information, such as take-off data, to a contractor⁶. There is no gain for an engineering firm to share this information, and if there is a problem with the data, this exposes the engineering firm to liability. Therefore, contractors typically perform their own quantity take-offs even though the engineer already has this information. This creates needless duplication of work.

The sharing of information can also enable the reuse of data throughout the life cycle of a project. An investment in BIM software could actually add value to a model, by including space planning, quantities, facility maintenance records, furnishings, and inventory, as well as recording changes throughout the facility life cycle⁷. This includes providing a digital owner's manual of the building upon completion, so facilities managers can take that building and use it to help budget and prepare reports for all maintenance work. Historically, a design remains the property of the designer following the completion of a project⁸, but BIM is capable of providing a wealth of information

⁵ Ashcroft, H. , 2006. Building Information Modeling: A great idea in conflict with traditional concepts of insurance, liability, and professional responsibility. Victor O. Schinnerer and Company, Inc. Chevy Chase, Maryland.

⁶ Ireland, B., 2009. Barriers to BIM: Electrical Industry lags behind other trades in adoption of Building Information Modeling. *Electrical Construction and Maintenance*, March 2009.

⁷ McGraw-Hill Construction, 2008. *Building Information Modeling Trends SmartMarket Report*. New York: McGraw-Hill

⁸ Hurtado, K. A. and O' Connor, P. J., 2008. Contract Issues in the use of Building Information Modeling. *International Construction Law Review*.

which may be beneficial for the maintenance and operations of a building. Owners and building occupants therefore may want to continue to use and develop the model.

Some owners have contractually obligated the architect to treat the model as a deliverable. There are many reasons for this: the owner's desire to get what they believe they are paying for, to know more about the design as it progresses, and to use the model as a tool in management and operation of the facility during its lifecycle. This latter reason is cited as one of the benefits of using BIM for the design and construction of a facility. Obviously, this raises more concerns regarding a design professional relinquishing possession and control of an instrument that could serve as a basis for future liability.

This free flow of information is also a cause for concern, and Ireland⁹ points out the critical issue of ownership of the design data. For example, a specialty mechanical, electrical, or plumbing (MEP) contractor who inputs detailed design into a model (which is shared with all team members) may want to maintain the right to that data when the project is over. If that data is now part of a building model used by an owner, proprietary information could find its way to a competitor upon completion of the project.

Many BIM-related issues stem from concerns about ownership of the model and use of the information that the model contains and generates. The ownership of the intellectual property generated in the BIM process has not been adequately addressed. In the shared design philosophy intrinsic in BIM, there are layers of intellectual property provided by design participants and others that are incorporated into the final model. With project stakeholders all being able to share project information and add details to the project model, disputes may arise over who owns the copyright if any of the design elements are used on future projects. To further complicate this issue, there may be inadvertent sharing of proprietary information, trade secrets, or patented processes. Confidentiality as well as ownership rights may be compromised.

3) Allocation of Risks

The use of BIM substantially alters the relationships between parties and blends their roles and responsibilities. The legal framework in the United States, however, assumes a less collaborative environment with clearer delineation of responsibility. As BIM projects become more commonplace, risks will need to be allocated rationally, based on the benefits a party will be receiving from BIM, the ability of the party to control the risks, and the ability to absorb the risks through insurance or some other means.

Recent advancements in technology have made BIM both available and relevant to the work of all members of a project team. The prevalence of BIM will inevitably change the ways projects are conceived, designed, communicated and constructed. However, the core responsibilities of the members of the project team will not change.

⁹ Ireland, B., 2009. Barriers to BIM: Electrical Industry lags behind other trades in adoption of Building Information Modeling. *Electrical Construction and Maintenance*, March 2009.

Whether the design is issued in the form of 2D printed documents or a 3D electronic medium, or in combination of both, the responsibilities of the members of the project team remain unchanged. It is very important to recognize the difference between design and coordination. The creation of a composite coordination model does not require or supplant a design that is conveyed in 2D printed documents. When a contractor or construction manager creates a “coordination” model, the BIM tool is completely similar to a light table used in the past to overlay mechanical and electrical drawings. Recognizing the validity and value of the information in any BIM is the responsibility of every project team member who utilizes it.

Contractors and construction managers need to understand that coordination, whether through BIM technology or a light table, is their core service to the project. As the leaders of construction coordination, contractors and construction managers have the responsibility to encourage and facilitate the sharing and distribution of BIM technology on a project. Appropriate contract language will guide the open sharing of information between team members. In addition, a design team also must recognize the benefits of sharing all available electronic information with the entire project team. Subcontractors are still responsible for fully conveying their interpretation of the design intent to the design team. They also must coordinate their work with that of other subcontractors by sharing electronic information they have developed in file formats that can be used and combined with the work of others.

4) Privity of Contract and Third Part Reliance

The extent to which third parties may rely upon a designer’s work is a highly contested subject. The use of a collaborative model lessens the likelihood that a designer may claim the lack of privity of contract as a legal defense.

The model designer must be aware that there are other parties relying on the accuracy of the model. It is foreseeable that the main purpose of the model is to provide information for contractors’ and subcontractors’ to use in constructing a facility. Under the Restatement of Torts, Second, issued by the American Law Institute, a person negligently providing information is liable if it is intended that the plaintiff be able to rely on the information. Liability only requires that there be intent to influence and reach a group or class of persons¹⁰. For this reason, contractors’ and subcontractors’ relying on the model will likely be able to bring an action against the designer for damages caused by negligent errors. Therefore, considerations must be given to requiring a waiver of consequential damages as a pre-condition to using the model or otherwise limiting damages due to model errors.

Provisions used by design firms that treat electronic data as inferior representations of 2D drawings no longer makes sense. The idea of obtaining waivers or limitations of liability to control allegations of detrimental reliance is counter to the BIM process. Disclaimers may be ineffective since reliance is implicit. With BIM, there must be a free exchange of data and the ability to rely on such data when incorporated into the

¹⁰ Ashcroft, H. , 2006. Building Information Modeling: A great idea in conflict with traditional concepts of insurance, liability, and professional responsibility. Victor O. Schinnerer and Company, Inc. Chevy Chase, Maryland.

final model. Harm can still occur, however, and whichever party is seen as controlling the information may be seen as the source of the harm. If the model becomes a tool to assist the client in operating or modifying the facility, the question of the rights of the client to use all the information in an unregulated way also becomes paramount.

5) Professional Design Responsibility

For the protection of the general public, many states regulate the professional practices of architecture and engineering. These states require that each project be under the responsible charge of a licensed architect or engineer, and that these personnel are designated as the architect or engineers of record. Additionally, these states require that the seal of such individuals appear on all drawings, specifications, and other design documents issued by the firm for such projects. These requirements are easily understood and followed in the creation of 2D drawings where the design documents are issued only in paper form; however, there are inherent problems when many non-licensed participants may have access to a BIM model.

It would be very difficult to ensure that licensed design professionals will always in charge of the creation and modification of the data that forms a digital model. Design elements are increasingly delegated to unregulated parties such as contractors, fabricators, and manufacturers. With BIM, parties who are supplying design information are not, by contract, under the responsible charge of a prime design professional. BIM may lead to increased decision making not by design firms using professional judgment, but rather by construction entities or by a computer program working on preset rules created by independent organizations not subject to registration laws.

6) Standard of Care

According to the American Society of Civil Engineers¹¹ (ASCE), an engineering standard of care is a relative term and its determination involves experience, insight and common sense. The law provides that an engineer performing professional services has a duty to exercise care and skill to the same degree as that used in like cases by reputable members of the profession practicing under similar circumstances. Also, the engineer has the duty to use reasonable diligence and best judgment in the exercise of skill and the application of learning. The failure to perform any one of these duties is defined as negligence.

Design professional liability is almost always based on the standard of care. Yoakum¹² defines a typical standard of care clause as follows: “The Design Professional’s services shall be performed in a manner consistent with that degree of skill and care ordinarily exercised by practicing Design Professionals performing similar services in the same locality, and under the same similar circumstances and conditions”. A design professional’s agreement should explicitly permit reliance without detailed checking,

¹¹ Bass, Eugene (2006) The Standard of Care - A Moving Target, <http://www.asce-sf.org>, accessed July 1, 2010

¹² Yoakum, S. (2006). “Building Information Modeling (BIM) Risks and Liabilities.”, Donovan Hatem, LLP, Builders Association, Kansas City, Missouri.

but the ability to rely on another's work may be limited by professional registration statutes. This may lead to using risk transfer devices, such as limitations of liability or indemnification agreements, as methods to decrease professional liability.

According to Sieminski¹³ there is a concern that the general use of BIM will alter both the standard of care and historical protections afforded to design professionals by the doctrine of privity. Until recently in some jurisdictions, the doctrine of privity of contract shielded architects and engineers from negligence claims by parties with whom the design professionals did not have a contract. Recent case law has relaxed the privity requirements to a limited degree. Many jurisdictions now allow claims where no contract exists when it is clear that a contractor reasonably relied upon information that a design professional misrepresented, within a context in which it could be anticipated that the contractor would be using this information. This may cause the unraveling of the standard of care where no privity of contract exists.

7) Spearin Doctrine

The *Spearin* doctrine is generally used by contractors as a defense to an owner's claim of defective and non-conforming work. If a contractor builds a structure according to an owner's plans and specifications, and the structure does not function as intended, the contractor is not responsible. When defects in the plans and specifications are the cause of the problem, *Spearin* shifts the responsibility back to the owner, who may then pursue legal action against the project architect or engineer of record. This is often referred to as "the owner's implied warranty" of the adequacy of the plans and specifications. The original *Spearin* case, *United States v. Spearin* (1918), 248 U.S. 132, gave rise to the *Spearin* doctrine.

The implied warranty set forth in the *Spearin* case was twofold: first, that the information contained in the plans and specifications would be accurate, and second, that the plans and specifications, if followed, would be adequate to accomplish the purpose of the project. Most importantly, the *Spearin* case has also withstood challenges that the responsibility of the owner is overcome by the usual clauses contained in construction contracts. This includes requiring the contractor to visit the site, to check the plans, to inform him of the requirements of the work, or to assume responsibility for the work until completion and acceptance. Contractual language obligating the contractor to examine the site do not impose upon the contractor the further duty of making a diligent inquiry into local conditions to confirm whether the owner's representations in the plans and specifications were accurate. The contractor also is not obligated to second-guess the adequacy of the plans and specifications to accomplish the purpose of the project. The United States Supreme Court has emphasized that an owner's warranty was implied by law, and that this did not need to be expressly stated in the contract documents.

This raises two major legal questions that need to be further investigated as a result of the collaboration process enabled by BIM. First, from a contractor's point of view, does the BIM collaborative process during the design phase of the project deprive the

¹³ Sieminski, J., 2007. Liability and BIM. *AIA Best Practices BP 13.01.08*.

contractor from the protection of design error provided by the *Spearin* Doctrine? Second, from the designer's point of view, does the collaboration enabled by BIM during the design phase erode a designer's traditional protection from responsibility for contractor means-and-methods?

The answer to these key legal questions remains in the construction contract itself. Therefore, as long as the parties' roles are appropriately defined and the appropriate control is exercised over the collaborative process, the use of BIM does not necessarily alter the traditional allocation of responsibility among project stakeholders. It is important to emphasize that there is no new legal issue here. Instead, the hypothetical merely involves the application of long standing legal principles to a new context.

The *Spearin* doctrine remains a powerful tool for contractors when faced with problems resulting from defective drawings and specifications. Applied in its basic form, the *Spearin* doctrine frees contractors from having to double-check a designers' work and enables contractors to concentrate solely on the construction's means and methods. Owners who wish to avoid the application of the *Spearin* doctrine need to think ahead and include very clear statements as to the risks that the contractor is accepting. All parties involved in a construction project need to study the contract language carefully to be sure they understand exactly what the contract says about liability for plans and specifications that don't reflect reality or that won't work as intended.

8) Economic Loss Rule

The economic loss rule is another highly contested defense to contractors' actions against design professionals¹⁴. The doctrine is very simple in concept; it provides that where a party sues for purely economic losses, the party that is suing needs to have a contract with the defendant. The utility of the defense varies among jurisdictions and is dependent upon specific facts, similar to defenses used with privity and third-party reliance.

The complexity comes in the application. In states where the economic loss rule holds sway, plaintiffs continually attack application of the rule. For example, many states hold that professionals who commit malpractice may be sued in both tort and contract. Some states require a contract; while other states permit a suit without a contract where the plaintiff could be reasonably expected by the professional to rely upon the professional's information. The ability to sue professionals without a contract has an obvious potential impact on construction projects where information is often provided by architects, civil engineers, geotechnical engineers, land surveyors and other licensed professionals. Furthermore, the use of a collaborative building model will be a factor tending to support a contractor's claim that it should be able to recoup its economic losses.

¹⁴Ashcroft, H. (2006). Building Information Modeling: A great idea in conflict with traditional concepts of insurance, liability, and professional responsibility. Victor O. Schinnerer and Company, Inc. Chevy Chase, Maryland.

9) BIM Implementation

Architects, Engineers, and Constructors must remember that BIM is not just software, but should be considered a process. It must be thought of as a business decision, and the impetus of transition must be a management decision. A recommended course of action implementing BIM is to start by using the BIM process to produce faster, more accurate construction documents. Then, transition to model coordination/collision detection.

A transitioning roadmap must be a commitment from management. A BIM “champion” should be selected. This should be someone who is passionate about implementing the change. Next, an implementation plan should be developed and a pilot project selected. Additional formal training may be necessary initially. Using an implementation plan, a company-specific BIM Manual should be developed based on the experience of those involved in the pilot project. The process should then be repeated on future projects, one by one, until all new projects are using the BIM process. Finally, an ongoing training program should be developed to improve the process.

In developing an implementation plan, it is important to evaluate the current staff’s skills and knowledge and to set measurable milestones. The selection of the pilot projects should be based on size. It may make sense to focus on a particular construction market as well, such as medical office buildings, or science and health care labs, before deploying to other markets. It should not be too big, or too small. It should be a typical project for the company. Additionally, it should be a project that will benefit from the use of three-dimensional modeling.

The creation of a BIM Manual should document BIM procedures that work for the company. These procedures will not be the same for each company. The creation of templates will assist in formalizing office standards and encourages efficiency for future projects. The manual and template can then be used as a basis for training current employees who will transition with future projects, as well as new employees.

By implementing BIM it is conceivable that some participants, such as contractors, will find that they are able to expand their services: providing detailed feasibility studies (pre-design), programming and planning (pre-design), early cost estimating, value engineering (design development), offering pre- and post- construction services, reselling the model data for public relations animation or marketing, model coordination and collision detection, and facilities management. At the very least, the implement of BIM will enhance the delivery and value of their core expertise; streamline their workflow; improve communication with clients, consultants, and contractors; broaden the services they offer to clients; and ideally, increase net revenue per employee.

The work process evolves as information is shared rather than isolated – new possibilities emerge. The ability to model coordination/clash detection with minimal delays increases as design iterations are sent back and forth to engineers as well as other specialty contractors, such mechanical and plumbing. Scheduling and estimating can be imbedded and revised concurrently with the BIM model-work, ultimately allowing the ability to provide for improved construction.

A New Frontier for the Construction Court

Michael Reynolds¹

Faculty of Science and Technology,
Anglia Ruskin University, Chelmsford,
United Kingdom

Email: m.p.reynolds@lse.ac.uk

This paper looks towards a new frontier for civil justice that will extend the boundaries of judicial function to procure earlier settlement. We firstly consider a brief historic review of early judicial intervention in the Official Referees' court and some of my quantitative findings and conclusions on its efficiency in the 20th century. We also glance at comparative systems and key elements of processes from which we may conclude undoubted benefits of mediation: the key driver being cost and time, and whilst time may be available, cost is becoming an ever more critical question in the economic context of financial default and decline following the global depression or Great Recession, as some term it. This gives reason for a degree of lateral thinking in looking at comparative international systems supporting the idea of a multi-door courthouse concept pioneered by Frank Sander of Harvard Law School. Having also glimpsed at adjudication and finding general benefit in terms of cash flow, we may conclude that no particular ADR process provides a panacea for all ills. Thus we may consider that speed cannot be at the expense of justice: courts must act with regard to the factual matrix, the extent of technical complexity, the amount of evidence, the sums involved and importance to the parties. Courts can aim to be more modern and effective with the appropriate level of judicial expertise where initiative and creativity make a more efficient use of judicial time so that cost becomes less important. That must surely be the aim of all engaged in pursuit of justice.

Keywords:

Civil justice, legal process, mediation.

1 Judging the judges²

The publication of the Hamlyn Lectures 2008, *Judging Civil Justice*, in November 2009 together with the recent completion of Lord Justice Jackson's deliberations on costs reinvigorates the debate on ADR and civil justice. Before joining the debate, it is important to define justice in the context in which this paper is written and in which part of it first appeared in the *Construction Law Journal*. In the context of procedural reform, justice is difficult to define, meaning different things to different people. Here we mean procedural justice achieved by efficient and effective processes, which obtain a

¹ PhD (LSE), LL.M (Lond.), M.Sc (Lond.), F.C.I.Arb., F.C.I.O.B. Solicitor and Chartered Arbitrator. Senior Lecturer, Part-time Anglia Ruskin University and Visiting Senior Fellow, Law Department, London School of Economics and Political Science.

² See also: Reynolds, M.P.(2010), 'Crossing the Rubicon', *Construction Law Journal*, Volume 25 (issue no.2), pp 77-87

just resolution within a reasonable time. According to my friend, the learned editor of that journal,⁸⁵¹ might as well have called my article: “A Search of the Holy Grail.” However, I do not perceive any similarity between myself and Sir Percival, nor indeed Sir Galahad, and regrettably even less with Indiana Jones of Last Crusade fame. As elusive as the grail may have been and as elusive justice may seem, no less than the Knights of the Round Table, fable or fiction we must journey towards it. Thus, we may also interpret substantive justice as attaining the correct interpretation of applicable law. The two are not necessarily synonymous, unless the substantive just result is discernable earlier and, with the parties’ consent, judgement can be expedited. In my view, that is the key to a means of resolving this conundrum between those who promote ADR and those who espouse the cause of procedural reform. If only like Merlin we might magically reconcile the problems of time, and delay then the cost conundrum might dissipate. But, there is no easy solution and history teaches us that many of the current *vox populi* pundits propagandising ADR have no idea whatsoever of how we got to where we are. Without an appreciation of history we cannot attempt to predict the future. This is true in legal process as much as it is true in international relations.

Thus, this paper looks first at the historic tradition of effective case management in the Official Referees’ Court, second we look at lessons we might learn from our European friends and neighbours in terms of comparative processes and international arbitration techniques, third we look at an advanced trial centre in the United States and we finally we take a hard look at the practicalities.

2 Newbolt’s Scheme

In the early 1920s, Sir Francis Newbolt³ invented a form of judicial case management in order to progress referrals i.e. actions transferred to his court by the High Court for investigation and report by a referee, more effectively and efficiently. His approach was revolutionary combining judicial activism with traditional judicial passivism which varied according to the particular needs of the case. In the context of the current debate, what is particularly interesting is the first element of my discovery of this scheme, lost to history for 75 years until I came across the Lord Chancellor’s Office files on Official Referees in the National Archive during my doctoral research at the LSE. What I discovered was a rudimentary form of case management: special procedures that Newbolt adopted in chambers enabling informal referee resolution and early settlement. This was facilitated by a first summons for directions hearing which he used more like a commercial arbitrator’s preliminary meeting. The summons for directions procedure had been introduced in *Rules of the Supreme Court 1883* but only to give formal directions as to preparation for trial. Newbolt revolutionised this and used it as a case management tool to encourage settlement. Not only that but he also invented a second stage directions process before trial.⁴

As Newbolt (Newbolt, 1923, p.440) wrote:

³ Sir Frances Newbolt K.C. 1914; Hon. R.A.; J.P., M.A., F.C.S., A.R.E. Hon. Professor of Law in the Royal Academy. Publications included: *Sale of Goods Act 1893*; *Summary Procedure in the High Court*, and *Out of Court*. Official Referee 1920-1936.

⁴ *The Times*. 4th September 1930. p.11. Issue 45609. col. F

..there is no greater check on wasteful expenditure than the arrangement by which the Trial Judge takes his own summonses.

Elsewhere I have described the other elements of his form of judicial case management (Reynolds, 2008). The essential point to grasp in the current debate is Newbolt's informal businesslike approach. This was atypical of the traditional conservative and passive English High Court judge of those times who would under no circumstances be overtly involved in settlement discussions. What Newbolt did was to narrow the issues in the course of discussion so that settlement became inevitable. He did not negotiate in any sense, nor mediate nor conciliate. His "scheme" was not the same as the modern scheme of mediation operated in the Technology and Construction Court today as implied by the Editor of the *Construction Law Journal*.⁵ As a referee Newbolt occupied a unique position in the English judicial hierarchy. His office was that of a bifurcated judicial officer who performed both the functions of a master at interlocutory stage, and of a High Court judge at trial. This gave him and his colleagues a unique advantage in so far as, upon giving early directions after the referral, he could ascertain the true issues in dispute. Having ascertained such matters, and with the parties' consent, he was able to suggest a means of resolution.

None of his interventions or chambers sessions were subject to any appeal, and whilst not entirely approved by the Lord Chancellor of the time (Lord Birkenhead), he was not dissuaded from pursuing such initiatives.

In his seminal article Newbolt (Newbolt, 1923) stated the case for such intervention:

The Defendant who often has good reason to complain of some overcharge, of defective work, swears a vague affidavit, and obtains leave to defend as to part, or all, of the claim. But he may have, in fact, no case... If a few days after an order on the summons before the Master the parties met before the Referee and discussed the position such a miscarriage of justice as appears in the cases described would be impossible. The main source of avoidable waste of money is the occupation of time in Court which a little thought and discussion in Chambers would save, and does save. In matters of account, in kindred cases, much money has been thrown away in the past by discussing in open court matters of pure arithmetic, or the contents of business books which turn out not to be in dispute, or not material to the issue, or fatal to one parties contention. Many other examples might be given. In one case evidence was taken before and also at the trial on both sides to prove the market price of goods at a foreign port. *If a preliminary discussion had taken place* none of this evidence would have been gone into as it was not relevant to any issue on the pleadings.

⁵ Editorial. (2010) *Construction Law Journal*, Volume 25 (Issue no.2),

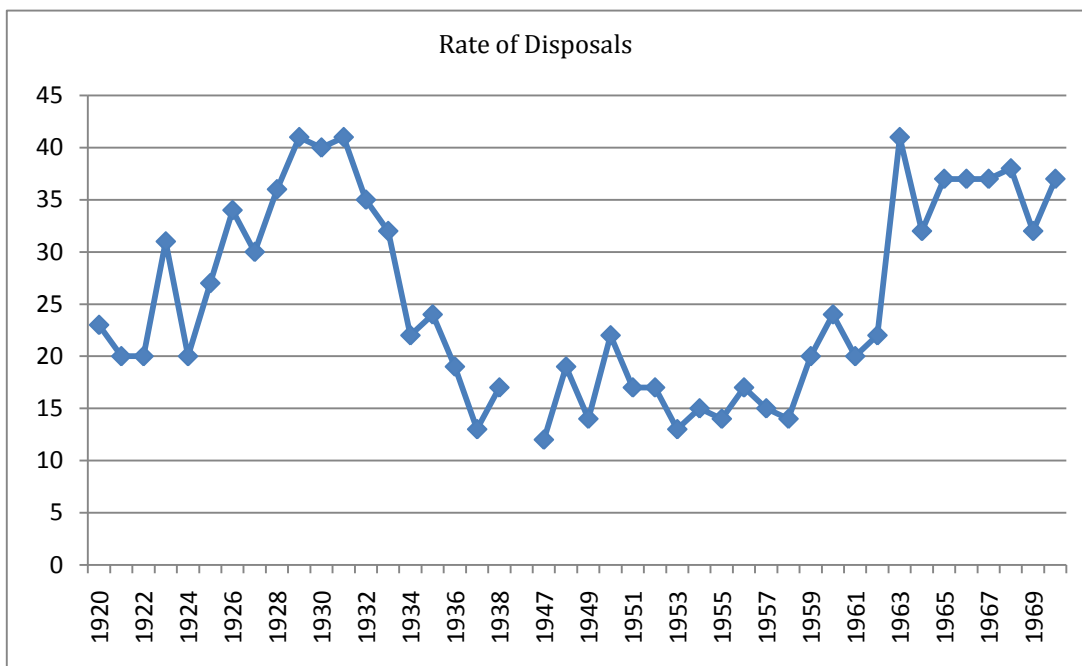
Another instance will strikingly illustrate the point. A mortgagor claimed an account of matters extending over many years: the case was expected to last for a fortnight. After *an informal discussion in Chambers* the parties agreed that an independent accountant should examine the books before trial, as a witness for both sides, and report on the points in difference: so that the issue between the parties should be defined and tried. He reported that having explained the figures to both the Plaintiff and the Defendant there were no points in difference and there was nothing to try. *This is not arbitration or conciliation or concession, but an intelligent use of a Court of justice by business men.* They spent perhaps £50 or less in arriving at a result which would in the ordinary course have cost ten times that sum, and would have worried them for a year.”

That was the problem then and that is how he attempted to deal with it. The “Scheme” had varying degrees of success as may be seen in the following analysis.

Effectiveness of Newbolt’s “Scheme.”

The efficiency of Newbolt’s “Scheme” may be considered from an historical analysis of *Judicial Statistics* in the 51 year period between 1919 and 1970. The following is just one of the statistical analyses that were applied to consider the efficiency of Newbolt’s “Scheme.”⁶ Comparing the numbers of cases withdrawn, or otherwise disposed of, to the number of referrals it is possible to calculate the effectiveness of Newbolt’s “Scheme” as a means of encouraging the parties to earlier settlement. This is represented from the following chart C.1:

Chart C.1 Disposal rates 1919-38 and 1947-70



⁶ Five other tests were applied: M.P. Reynolds, *Caseflow Management: A Rudimentary Official Referee Process 1919-70* Ph.D Thesis (London School of Economics, 2008).

Commentary

2.1.1 Pre-war era:

Following my discovery of a confidential report to the Lord Chancellor dated May 1920 I undertook an analysis of *Judicial Statistics* in the pre and post war periods to see whether the judicial techniques employed by Sir Francis Newbolt had any effect especially his advanced settlement techniques in chambers discussions with solicitors. In this pre-war phase, Chart C.1 illustrates my findings showing a dramatic rise in settlements between 1921 and 1924. Being practically contemporaneous with that report and with his seminal paper in the *Law Quarterly Review* I suggest that Newbolt's activist methods had some effect. This view was substantiated by further analysis of the statistics between 1924 and 1931. Here the Chart demonstrates a greater rise in the settlement/disposal rate which again I suggest is due to some extent to Newbolt's effective interventionist style.

Taking the years 1921-31 there is a 21 per cent increase in the rate of cases settled; i.e. a rise from 20 per cent to 41 per cent. Between 1922 and 1923 there is an increase of 11 per cent in the number of settlements, or a 65 per cent increase in the rate of settlement. A significant factor here is the presence of two inventive referees—George Scott and Francis Newbolt—who were undoubtedly assisted in achieving this record during the crucial period 1923 to 1933 by two experienced judges: Sir Edward Pollock,⁷ and, subsequently, Sir William Hansell. Whilst the chart illustrates considerable fluctuations in settlement rates these tend to support the argument that judicial case management increased in efficiency between 1919 and 1930. After 1932, the efficiency rating slumped and it remained for the most part below the 20 per cent mark until 1958.

The average rate of settlement in the years 1919-38 was 27 per cent, whilst the average number of annual referrals was 348. Newbolt and his colleagues succeeded in reducing a backlog of 226 cases in 1921 to 83 cases in 1928 after which it began to rise to 126 in 1936, the year of Newbolt's retirement.

I have suggested that these effective efficiency trends were due to the operation of Newbolt's approach to case management, but after 1931 there was a period of decline. Chart C.1 illustrates a high point of 41 per cent in 1931. This is followed by a period of decline to a low of 19 per cent when Newbolt retired in 1936. If he was still practising judicial case management then one has to ask, why the decline? The answer may be found in the fact that when Sir Edward Pollock retired he was 86 years old. Lord Cave, the Lord Chancellor, appointed George Hansell to replace him in 1927. Hansell was 71 years old. He retired at the age of 75 in 1931.⁸ Following that Scott retired in 1933. In that year the rate of settlement dropped from 32 per cent down to 22 per cent in 1934. Although it slightly

⁷ Senior Official Referee 1920-27 replacing Sir Henry Verey. K.C.

⁸ By reputation he was a judge of ~~ability~~ and character," *The Times*, April 20, 1937, p.21, col.D.

recovered in 1935 at 24 per cent, it declined further to 19 per cent in 1936. It then fell to a low point³⁵⁵ of 13 per cent in 1937, recovering slightly at 17 per cent in 1939. These lower rates in 1937-38 may be explicable by the managerial inexperience of Eastham and Pitman, Eastham having been appointed in 1937. The latter period therefore presents a difficulty because there appears to be a rise in the number of trials, and a corresponding rise in the backlog with a fall in the number of disposals before trial. If judicial case management was effective then we would expect the disposals to be increasing and the backlog to be falling. That is not the case. But it does appear that the increased efficiency was due to Newbolt and the lower efficiency rates after his resignation illustrate that.

Given the figures in Chart C.1 the average percentile for settlements between 1919 and 1932 is 30 per cent. Between 1933 and 1937 it is 21 per cent. The margin of difference cannot be ignored, and if judicial management did make a difference in the first pre-war period it is difficult to argue that it had a similar effect in the latter pre-war period. Thus, the trends indicated from *Judicial Statistics* and contained in the chart support an efficiency hypothesis up to 1933. Post-1933 the trend up to the outbreak of the war does not support it. What it infers perhaps is that however efficient the judiciary maybe or whatever case management methods are adopted there is a certain minimum manpower level that must be maintained otherwise whatever methods are employed higher rates of efficiency are unattainable.

(b) Post-war period

Looking at Chart C.1 in respect of the post war era we find that in 1947 the disposal rate was 12 per cent. The figures then fluctuate between 19 per cent in 1948, and 14 per cent in 1958. After that, disposal rates rise to 20 per cent in 1959. They do not rise above that rate until 1962 when the rate rises to 22 per cent. In 1963 it rises to 41 per cent; equivalent to the rates in 1929-31. This apex coincides with the appointment of Norman Richards who by reputation was a good judge. During his tenure of office, the rate of settlement did not fall below 32 per cent of referrals and averaged 36 per cent. This compares very favourably with a 27 per cent average rate of settlement achieved during Newbolt's era. The average settlement rate for the post war period was 24 per cent, as compared to 27 per cent in the pre-war period. Whilst the pre-war era was generally more efficient than the post-war era, the Richards era was more efficient per se in terms of settlement than Newbolt's time notwithstanding the increased workload. Taking a wider view, accounting for backlog, we find that whereas in Newbolt's time the referees halved the backlog, in Richards's time it almost trebled. An argument in support of the overall effectiveness of judicial case management in the post-war period is therefore more difficult to sustain although referrals more than doubled between 1960 and 1970, from 425 cases to 901 cases. This had considerable effects on the backlog which increased by 43 per cent in the period 1957-70.

Newbolt's court was relatively efficient both in trials and in earlier resolution, and that in terms of the settlement rate it is arguable that Newbolt's "friendly discussions in chambers" made some difference. Newbolt's settlement rate was double that of his successor Sir Tom Eastham. At best, properly applied judicial case management could cut trial times by one-half to two-thirds (Newbolt said he achieved an 80 per cent reduction where he employed single joint experts) and that possibly a quarter of all referrals had some form of caseflow management. Despite all Newbolt's and his successors' efforts the backlog remained a key problem as demonstrated in the chart below.

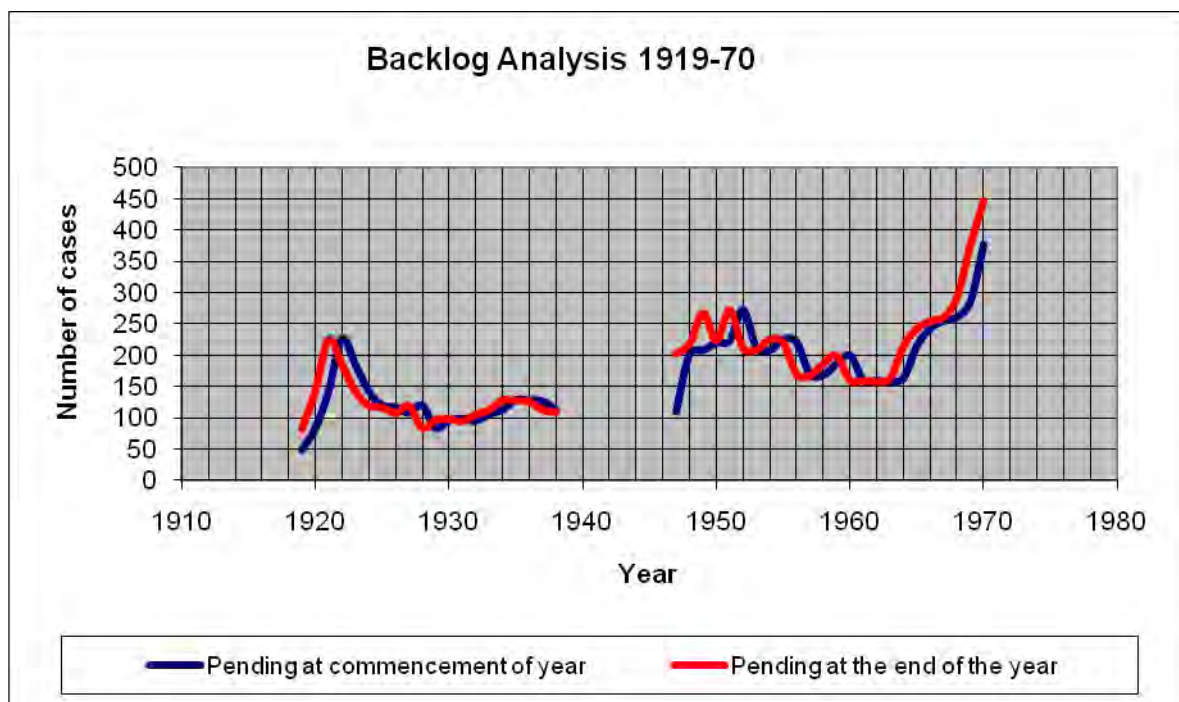
General Conclusions on Newbolt's Scheme."

First we may conclude that a form of case management in the Official Referees' court existed in the 1920s and that it is not a new concept invented by Lord Woolf or his advisors.

Second, we have demonstrated that Newbolt's Scheme" may have been effective in disposing of cases earlier.

Third, backlog like the national deficit can be reduced, but not eliminated. We see from Chart C 2 that whilst backlog was reduced over the periods: 1922-26, 1927-28, 1936-38, 1950-52, 1955-56, 1959-60, it rose understandably during the Great Depression and after 1963 and 1966 when the construction industry witnessed a surge in claims.

Chart C.2



Fourth, that where referrals increase so does the backlog, even though the rate of disposal is also increased. No matter what devices might be employed, manpower it seems is the decisive factor.

Fifth, it does not appear that Newbolt compromised his judicial function or independence and that he used his experimental case management techniques in only a proportion of appropriate cases. My overall calculation was that such measures were used in a quarter of all cases by the referees. In other

words those safeguards of the adversarial system were maintained and that the employment of such methods depended largely on the type of case having regard to the factual matrix, technical complexity especially in relation to expert evidence, and the importance of the issues and the value of the sums involved.

Taking these historical analogies aside is there anything to be learned from the experience of international arbitration and other jurisdictions in particular northern Europe as it is likely that Newbolt and, certainly Lord Woolf, had a clear appreciation of some of these more cost effective regimes?

3 Comparative Studies

3.1 The German Process

If we look at the “episodic” character of German process we see fundamental differences with our system. I do not advocate any adoption but we appreciate our own system better sometimes if we examine others. In Germany the judge investigates the case—a little like the practice of pro-active adjudicators here. Their system as described by (Langbein, 1985, p.823) suggests advantage in the judges’ fact gathering techniques. Whilst in theory this may overcome preliminary preparation and pre-action work as well as reduce expert witness briefing, both lawyers and clients are disadvantaged by a lack of control over the case: a position English lawyers would find intolerable. On the other hand, the German process has the advantage of lessening the risk of surprise. It also imitates Newbolt’s “Scheme” in terms of “discussions in chambers.” The German Code of Civil Procedure (Zivilprozessordnung) [ZPO] s. 278(iv) states:

The whole procedure up to judgement may therefore be viewed as being essentially a series of oral Conferences.

(Kötz, 1982, pp.237,243) describes a procedure where the judge has the file under his control and routinely requests relevant public records of relevance to the case. But the German judge, like Newbolt, as confirmed by Langbein (pp.828,831) discusses matters with the parties in a business-like meeting. It is also clear that like Newbolt, German judges see settlement as the goal as Othmar Jauerling (Jauerling, 1983, p. 171) says:

Settlement is sometimes prized as the crown of the judicial function, as the goal for which a healthy legal system continually strives.”

Whilst there are efficient features of the German system and it is less expensive, as there is less lawyer involvement, it would be difficult to adopt without fundamental changes to the education and training of lawyers and judges in England. So far as the Common Law system is concerned German lawyers might argue that costs are higher in England because of the nature of the English system. Let us take a few examples proffered by Langbein (pp.833-848):

Expert witnesses. Despite the provision in the *Civil Procedure Rules* for a single joint expert most construction cases in the TCC have an expert on each side whereas in Germany the court appoints a single expert.⁹

Witness evidence. This may be overtly partisan as taken by one side's lawyer whereas judicial examination is by an impartial judge.

Equality of arms. There is no such concept applicable where a large law firm is instructed on behalf of a major construction group against a small specialist sub-contractor with no retained lawyer or firm. In Germany the interposition of the judge compensates the balance-*jura novit curia*-the court knows and is bound to apply the law without the parties prompting. (Kaplan, van Mehren and Schaefer, 1958, pp.1193,1443)

As Fuller points out (Fuller, 1961, and Schwartz, 1978, pp.669,672) there is an ever-present danger of a rush to judgement. The fact pattern that appears to emerge may be entirely misconceived and wrong. That is why in England we prefer the adversarial system where the lawyer can protect the client's interest and prevent injustice from an anticipatory judgement. But the question is whether the German model has a point which we can recognise. That point is the stage at which the critical mass of fact is known so that the judgement is not going to be any different from an early judicial evaluation or opinion. Thus, the critical question for the court to address is at which point the court justly recognises that judgement is going to be no different from an interlocutory judgement.

We may surmise that it was in those cases that Newbolt's "Scheme" was effective. The problem is to identify that critical stage and to act impartially and independently so that justice may be done. If the facts can be reconciled and resolved the law follows. As Sir William Blackstone said (Blackstone, 1768, p.330)

Experience will abundantly shew that above a hundred of our lawsuits arise from disputed facts, for one where the law is doubted of.

As a model the German system has many attributes that enable earlier resolution and similarities with other Common Law and Civil Law systems a sample of which follow.

3.2 The Canadian Process

In Toronto and Ottawa in Canada cases are assigned to a standard or fast track. (World Law Group Member Firms, 2003) Cases must be submitted to mediation first unless the judge decides otherwise. At case management stage the judge sets a time limit on procedure to trial. There are two types of case management conference: trial management conferences where admissions can be made and a

⁹ Although parties can also appoint their own expert which may cause the court to appoint a second court expert or *obereparte*.

means found to make the resolution more efficient; and settlement conferences which are automatically scheduled to avoid trial by facilitating settlement.

In British Columbia the Small claims Court has a mandatory settlement process (Chemenki, and Hart, 1996) where judges spend 20 to 30 minutes with the parties facilitating settlement by —assisting them to address the needs and interests in a mediative way and possibly expressing an opinion on the strengths and weaknesses of each sides case.” (Marriott and Brown, 1999)

3.3 The Netherlands Process

This has its roots in the French Civil Code with some influence from the German Civil Code applying what its judges call the —*pleier* model.” Here the judges try to encourage settlement or set a date for final judgment or an interim judgement which may require further evidence. (World Law Group Member Firms, p.459)

3.4 The Norwegian Process

In Norway where the parties do not have legal representation mediation is compulsory before proceedings are undertaken in the District Court. In those cases a judge will act as the mediator. (World Law Group Member Firms, pp. 476,487).¹⁰

3.5 The United States

It is not possible to cover every state court in the Union and the Federal courts in this paper, but suffice to say that ever since Judge Lauer published his paper in the late 1920s containing his address to the American Bar Association (Lauer, 1928, p. 55) the courts in the United States have been far ahead of their English and European counterparts. In the Federal courts ADR is recognised as a tool of caseflow management initiated in legislative form by the Civil Justice Reform Act 1990 and recognition of ADR as one of the six principles of caseflow management. Further judicial recognition gave effect to this alignment of ADR and the courts in the case of *Re Atlantic Pipe Corporation*.¹¹ The judges held that a Federal District Judge had power to compel mediation subject to time and cost limits.

3.6 Conclusion

These are just a few random examples of how ADR processes have complemented the judicial process in civil law and common law countries. Whilst smaller claims tend to have mandatory mediation in some countries, there is no doubt that in many jurisdictions the benefits of mediation can

¹⁰ and see also (the Norwegian) Disputes Act 1915.(“*twistemålsloven*”)

¹¹ 304 F.3d 135 (1st Cir.2002)

be worthwhile provided there is ample safeguard against abuse. The key driver in all this is cost and time, and whilst time to mediate may be available, cost to litigate or arbitrate is becoming an ever more cash critical difficulty for a struggling industry. This was clearly recognised in the mid 1990s when statutory adjudication was introduced in England.

4 Adjudication lessons

One of the leading commentators in England on adjudication Tony Bingham has opined that adjudication is transforming the arbitral domestic scene. This is because adjudication has replaced arbitration as the most popular method of resolving construction claims, especially money claims. Bingham believes that this has affected arbitrators, who having experienced adjudication's fast track process, are getting to grips with cases earlier and being more pro-active utilising the powers of the Arbitration Act 1996 in a more effective manner. It may also be that the combined effects of Woolf (1996) Saville (1996) and Latham (1994) have placed additional challenges on dispute resolvers of whatever complexion to be more inventive and pro-active. According to Bingham a number of arbitrators have been driven to adjudication for fear of being criticised by the courts and their reputations ruined. This is apposite to the objective of the Arbitration Act itself which contrary to that view was aimed at encouraging English arbitration in the commercial world. In that context the court's role was that of supportive jurisdiction, not competitive and not supervisory in a disciplinary sense. If therefore arbitrators do have that "fear" then this something that must be of concern to the Chartered Institute of Arbitrators and other bodies that maintain panels of practising arbitrators. Whilst this presents negative reasons for adjudication we must recognise that cash flow is critical to any business and especially to construction, where work is in progress. Accelerating insolvency is not a sensible answer. This may result in a sub-contractor ceasing to trade, his labour thrown out of work and unemployed, the project programme being delayed, not to mention difficulties in finding a replacement contractor to complete an unfinished job. Thus, the underlying concept of adjudication is a good one and makes economic sense. Whilst in its incubation period there were teething troubles and considerable scepticism (not least from some members of my editorial board¹²) which I shared, those issues have been subsumed by the overwhelming need for early resolution and cash flow stability. It is perhaps due more to economic performance than other factors: the backdrop of a public sector net debt of £890billion is not an encouragement to adventurous dispute processes (Centre for Economic Performance, 2010).

5 The Multi Door Courthouse Concept¹³

Having looked at a variety of disputes processes we may try and reconcile these concepts by suggesting that it would be less confusing and more practical to affect a coalition of these processes.

¹² The Editorial Board of *Construction and Engineering Law* which included an eminent critic of adjudication the late Professor Ian Duncan Wallace Q.C. M.A (Oxon).

¹³ Idea originally described in: Sander, F.E.A. 'Varieties of Dispute Processing' (1976) Volume 70 *Federal Rules Decisions* . pp. 126-127, and 130-132.

Given that we have a plethora of alternatives to litigation and arbitration, and given the expertise that has grown in the RICS, RIBA, ICE, and in the legal profession, including the judiciary, it may be time to debate the concept of what Professor Frank E Sander of Harvard Law School¹⁴ referred to as the multi-door courthouse (Sander, 2004, p.117). That concept has many attractions in my view. This facility would avoid the confusion of a variety of appointing bodies and institutions offering dispute resolution resources and provide a key point of reference for resolving disputes. Bearing in mind the imbalance in public funding the construction industry, which spends a considerable amount of its profits on disputes, might see the sense in such a body. The key components might be the TCC, with disputes resolvers from LCIA, ICC, ICE, RIBA, ICE, RICS and CIOB augmenting the process. This might have some appeal to the political classes who face increasing unpopularity in cutting the public deficit. If such an approach can be adopted in the former English colonies of America by our American cousins, and in particular those in: Cambridge, Massachusetts, and the District of Columbia, and the State of Maine I see little reason why we should not take a much closer look at an alternative sensible, practical, and commercial approach in stringent times.

6 International Arbitration recommendations from CEDR

Having considered the advantages of comparative systems, and particularly having been an advocate of judicial encouragement of settlement for most of my career as a solicitor, it is welcoming to find the most recent proposal from CEDR (2009) promoting settlement in international arbitration disputes. CEDR's Commission was co-chaired by Lord Woolf and an experienced international arbitrator, Professor Gabrielle Kaufmann-Kohler. Significantly the arbitrator may consider whether the dispute can be settled and the method of achieving settlement. This is a useful step forward because it means that the arbitrator like Newbolt can discuss ways and means of settlement by stepping into the arena without prejudice to his duties or function and possibly free from fear of challenge. Taking proactive steps as Newbolt did would therefore be within the parameters of Article 3 of the CEDR Rules of the Draft CEDR Rules for the Facilitation of Settlement in International Arbitration. Under Article 4 the tribunal has a positive duty to ensure that the parties understand that they can settle their dispute, that the parties are aware of the processes to do so, and that the tribunal has explained the costs consequences for the parties if offers to settle are made in the course of the process. More importantly Article 5 goes further and states that the tribunal can provide its views on the dispute and suggest terms of settlement as a basis for further negotiations. Where the process diverges from Newbolt is the dicey situation where the tribunal meets with only one party. That would seem to me to be very risky and beyond my conception of the Newbolt —Scheme" which I adopt as a judicial parameter. Not only that but the idea of arbitrating then mediating and if that fails reverting to arbitration mode is not inviting and fraught with hazard. The concept of pro-active intervention to

¹⁴ Bussey Professor and Co-Director of the Dispute Resolution Program at Harvard Law School.

procure earlier settlement however remains a sound objective even if it follows the quest of Sir Galahad.

7 Common Sense and Commercial Practicality

In this paper we have looked at early judicial intervention in the Official Referees' court and some of my quantitative findings on its efficiency. Having also looked at the German system and others we may conclude that there is no doubt that in many jurisdictions the benefits of mediation can be worthwhile: the key driver in all this is cost and time, and whilst time may be available cost is becoming an ever more critical difficulty. This gives reason for a degree of lateral thinking in relation to the multi-door courthouse concept pioneered by Frank Sander of Harvard Law School. Having also glimpsed at adjudication and finding general benefit in terms of cash flow no particular ADR process provides a panacea for all ills and thus speed cannot be at the expense of justice: courts must act with regard to the factual matrix, the extent of technical complexity, the amount of evidence, the sums involved and importance to the parties. Courts can aim to be more modern and effective with the appropriate level of judicial expertise where initiative and creativity make a more efficient use of judicial time so that cost becomes less important.

To reconcile the differences in culture and approach represents a conundrum as elusive as that Arthurian quest referred to earlier. Is the truth better obtained by the judge's examination of a witness, or by the lawyer taking that evidence? Is judicial intervention a benefit to settlement or detriment? These are just two of the fundamental questions posed by this paper which must be debated if we are to reach that new frontier of civil justice.

Dr Michael Reynolds, Lecturer in Construction Law and Civil Law, Anglia Ruskin University and Visiting Senior Fellow, London School of Economics and Political Science.

8 Acknowledgements

The author would like to acknowledge the kind assistance of Professor Jan Kleinheisterkamp of the Law Department at the London School of Economics and Political Science who referred me to relevant papers on German Law and the role of German judges. I would also like to thank Mr Tony Bingham, Barrister at Law, of 3 Paper Buildings, Temple for discussing the current role and practice of adjudicators in the construction industry.

9 References

Blackstone, W. (1768) *Commentaries on the Laws of England* p.330

CEDR Commission, *Settlement in International Arbitration Consultation Document*, 2009

Centre for Economic Performance. The London School of Economics and Political Science. April 2010 *Macroeconomics and Public Finances: The Worst is yet to come.*

Chemenki, A. and Hart, C.E., (1996) *A Dispute Resolution Handbook*. Butterworths.

Fuller, L. (1961). *The Adversary System.* *Talks on American Law*. Volume 30

Jauernig, O. (1983). *Zivilprozessrecht* s.48 (VII) p. 171

Kaplan, van Mehren and Schaefer. (1958). *Phases of German Civil Procedure (Parts 1 and 2)* *Harvard Law Review*. Volume 71 p.1193,1443.

Kötz, H. (1982), *Civil Litigation and the Public Interest*, *Civil Justice Quarterly*, 237,243.

Langbein, J.H. (1985), *The German Advantage in Civil Procedure*, *The University of Chicago Law Review*, Volume 52 (Issue no. 4) pp. 828 and 831.

Langbein, J.H. (1985), *The German Advantage in Civil Procedure*, *The University of Chicago Law Review*, Volume 52 (issue no. 4) p. 823.

Latham, M. (1994) *Constructing the Team* H.M.S.O.

Lauer, (1928) *Conciliation-A Cure for the Law's Delay*, 156 *ANNALS* 55

Lord Woolf. (1996) *Final Report* The Stationery Office.

Marriott, A. L. and Brown, H.J. (1999) *ADR Principles and Practice* Sweet and Maxwell, London.

Newbolt, F. *Expedition and Economy in Litigation* (1923) 39 LQR 440.

Reynolds, M.P. *Caseflow Management: A Rudimentary Official Referee Process 1919-70* Ph.D Thesis (London School of Economics, 2008).

Reynolds, M.P. (2010), *Crossing the Rubicon*, *Construction Law Journal*, **Volume 25** (issue no.2), pp 77-87

Sander, F.E., (2004) *Some Concluding Thoughts* Volume 54 *Journal of Legal Education* p. 117.

Saville, (1996) M. *Report on the Arbitration Bill and DAC Supplementary Report on the Arbitration Act 1996.* Departmental Advisory Committee on Arbitration Law.

Schwartz. (1978) *The Professionalism and Accountability of Lawyers.* Volume 66 *California Law Review*. p.669,672 n.5.

World Law Group Member Firms. (2003) *International Civil Procedure*. Kluwer, The Hague/London/New York.

Building Information Models as Contract Documents: Common Practice for the U. S. Construction Industry – A Preliminary Report

Darren Olsen, JD¹, J. Mark Taylor, Ph.D., JD¹

¹McWhorter School of Building Science
Auburn University
Auburn, Alabama

Email: dao0002@auburn.edu; taylojt@auburn.edu

Abstract:

Building Information Models (or “BIM”) are being utilized to facilitate better collaboration and coordination on construction projects throughout the building industry. The question this research seeks to answer is, “To what extent is BIM being relied upon as contract documents by the U.S. construction industry?” In the United States, two different contract addenda (E202 Building Information Modeling Protocol Exhibit and Consensus Document 301 BIM addendum) have been introduced to allow the contracting parties to decide whether or not BIM is to become part of the contract. Further, the parties may then decide the order of precedence of the design information which may include both BIM and two-dimensional drawings created independent of BIM. The reasons in favour and against BIM being treated as contract documents should reveal some of the shortcomings and challenges related to BIM which, in part, prevent such models from completely replacing traditional two-dimensional drawings. For this research, architects, contractors and owner representatives were surveyed to discover their methods for dealing with the issue of BIM as contract documentation.

Although BIM has been utilized as a design and planning tool for a number of years, many in the construction industry are more comfortable with the use of traditional two-dimensional environments to define the construction contract. Even on projects that utilize BIM, it appears that BIM has not supplanted the traditional two-dimensional drawings when it comes to defining the scope of work through the contract documents. The goal of this study is, therefore, to determine to what extent BIM contract language is supplementing traditional contract language as used in the U.S. construction industry.

Keywords:

BIM, Building Information Models, Construction Documents, Order of Precedence

1 Introduction/Background

Building Information Modeling is a data-rich, object-oriented, intelligent digital representation of a facility. Two-dimensional drawings, such as lines, arcs and circles, are graphical entities only. In contrast, BIM define objects in terms of building elements and systems like spaces, walls, beams and columns (Azhar et. al., 2008 citing AGC, 2005 and CRC Construction Innovation, 2007). Building information models are much more than just three-dimensional representations of two-dimensional drawings. They also enable users to add cost, schedule, sustainability and other useful information to

the model. Despite all of the potential and possibilities, the legal and business structures for incorporating building information modeling lag far behind the technical advances (Ashcraft 2008). Since its introduction, BIM has been seen as the cornerstone of a promising technology that could revolutionize the construction industry by enabling construction professionals to engage in virtual design and construction. This would allow participants to achieve an unprecedented level of collaboration, and in return, have significant impacts on cost efficiency and effectiveness. One of the biggest obstacles to implementing BIM has been the fact that most, if not all, of the standard form contracts employed by construction professionals completely ignored, or inadequately addressed, how BIM is to be a part of the process (Lowe and Muncey, 2009).

In construction contracts, words alone do not define the agreement. Many documents, including graphical representations, come together to form the contract. These documents are naturally referred to as contract documents and are used to define the scope of work. Traditionally, contract documents include (a) the contract, (b) general conditions, (c) plans, (d) specifications, and (e) any other written conditions upon which the parties agree. The plans included in this traditional definition of contract documents are typically two-dimensional drawings (Hurtado and O'Connor, 2008). BIM potentially could upset the traditional definition of contract documents and replace some or all of the two-dimensional drawings, specifications and other written conditions with BIM based information and data.

Defining the contract documents is much more than a trivial part of the construction process. It is the basis from which the parties will work, coordinate and collaborate throughout the building process. According to the Spearin Doctrine¹, if a contractor follows the plans and specifications provided by or on behalf of the owner, and if those plans or specifications turn out to be defective, the contractor cannot be held liable for any losses related to those defects. Despite the age of the Spearin case, this doctrine remains the law in the United States and, as such, “contract documents” are of critical importance to the entire construction process. To stand in place of traditional two-dimensional contract documents, BIM must contain a tremendous amount of information and be relied upon by the parties as the primary source of information. Depending upon how BIM is utilized on a project, each model may not include enough design information to be referred to as a “contract document.”

Theoretically, each construction project utilizing BIM would have a fully integrated model from which all design and construction information could be obtained. If this were the case, BIM could be used to replace many contract documents and stand by itself as the single source of information for the design, construction and operation of a facility (Post, 2009a). While conceptually possible, several key points have prevented BIM from being fully adopted in this manner. First and foremost, the concept of a single model for a building project is not currently a reality. Often on projects that utilize BIM, several federated models are created for individual components of the facility. Secondly, the federated models are often themselves not the product of shared efforts and information. Instead, each is created independently from the two-dimensional drawings received from the designers (Post, 2009a; Post, 2009b).

From the designer's perspective, the use of BIM often presents increased risk and effort for very little reward. Use of BIM requires a higher level of accuracy and integration from the designers;

¹ *U.S. v. Spearin*, 54 Ct. Cl. 187, 248 U.S. 132, 136-37, 39 S. Ct. 59, 63 L. Ed. 166, 42 Cont. Cas. Fed. (CCH).

however, it does not always entitle them to more compensation. The nature of BIM's three-dimensional environment requires designers to coordinate the drawings much more purposefully and to take on some of the coordination efforts that traditionally have been the responsibility of the contractor (Post, 2009a). Ironically, from a designer's perspective, it is oftentimes deemed better to incorporate as little detail as necessary into the model, so as to not take on additional responsibility and expose themselves to additional risk. With traditional two-dimensional contract documents, designers are able to show certain systems and components schematically (i.e. MEP & structural steel) while leaving the detailed design up to the fabricators and specialty contractors. By introducing the three-dimensional environment, lines and symbols no longer suffice and specific objects have to be added to the model.

On the other hand, contractors typically desire the model to contain as much information as possible to enable them to coordinate the work, get accurate quantity takeoffs, and to perform class detection. Clearly, the full adoption of BIM continues to be hampered by positives and negatives that tend to offset (Post, 2009a; Post, 2009b). Currently, while members of a project team intend to use BIM for collaborative purposes, they also operate independently of one another. The team's collaborative efforts are based primarily on digital models, while the contract documents legally governing the contractors' work continue, for the most part, to be two-dimensional plans and specifications (Larson, 2008).

In 2008, two groups independently proposed contract language to incorporate BIM into the construction process. Both groups referred to their recommendation as a BIM addendum. The first to come out with its recommendation was a conglomerate of trade organizations that collaborated to develop and promote a new set of construction contracts called ConsensusDOCS. This group referred to its proposal as ConsensusDOC 301 - BIM Addendum (CD301). Later that year the American Institute of Architects came out with its BIM addendum which is called AIA-E202 "Building Information Modeling Protocol Exhibit" (E202). Both of these, while significantly different, were created to provide a contractual framework for the use of BIM on a specific project. Neither of these addenda was intended to replace the standard prime contracts; but rather created as contract riders to deal specifically with the issues surrounding BIM in the building process. The intent of these BIM addenda is to "create a document that would enable parties to easily and effectively introduce 21st century technology, namely virtual design and construction or building information modeling, into construction projects utilizing standard form documents" (Lowe and Muncey, 2009).

The BIM Addendum is an *addendum*, and is not intended to be used as a substitute for other standard form agreements between owners, design professionals, and contractors. Rather, an identical copy of the BIM Addendum is required to be appended to the contract between the owner and design professional and to the contract between the owner and the contractor. The effect of this arrangement is that the contractual relationships among the three principal parties (owner, design professional and contractor) are largely preserved, and any significant shift in contractual responsibility among these parties is avoided. At the same time, by implementing the BIM Addendum, all three of the principal parties agree to perform certain BIM-related tasks and assume certain BIM-related responsibilities, all of which must be addressed in order for a BIM project to be successful (Lowe and Muncey, 2009).

While there is a great deal of similarity between the approaches each of the BIM addenda takes, there is also a fundamental difference. While CD301 advocates the implementation of a BIM Execution Plan, E202 introduces the concept of Level of Development. The BIM Execution Plan of CD301 requires that all “Project Participants” meet within 30 days of contract signing and “use their best efforts to agree upon the terms of or modifications to the BIM Execution Plan” (ConsensusDOCS, 208). As a part of this process a checklist is used to guide the parties in the plan development. The checklist begins by identifying what models are to be developed, who will be responsible for developing the model and for what purpose each is to be used. The checklist is not intended to be exhaustive but rather a framework from which to develop the plan.

The E202 Level of Development (LOD) requires the parties to assign a LOD from between one to five for each model used in the process. For each LOD there is a description detailing exactly what that particular LOD will include and for what purpose the model can be used. For example, a LOD100 (first level) model element only provides overall building massing information such as area, height, volume, location and building orientation. A LOD500 (fifth and highest level) model is made up of constructed assemblies which are accurate to the extent that they can be relied upon for purposes of construction, maintaining, altering or adding to the project (Haynes, 2009).

The very nature of BIM introduces additional risks that must be allocated among the Project Participants. The BIM Addendum attempts to allocate these risks in the most fair and efficient manner possible. At the same time, the BIM Addendum attempts to deal with these additional BIM-related risks in such a manner as to not upset the typical allocation of risk on a project utilizing two-dimensional drawings and specifications. One of the risks unique to a BIM project is the risk that project participants may rely on the contribution of another project participant as accurate when in fact that contribution is not accurate. CD301 handles this risk by making each party responsible for any contribution that it makes to a model or that arises from that party’s access to that model. E202 addresses this potential problem with indemnification language whereby each party agrees to indemnify the other against claims resulting from modifications or unauthorized use of a model. The approach adopted by each BIM Addendum is simple and straightforward—each party is responsible for any contribution made by it or by any party for whom it is responsible (Lowe and Muncey, 2009).

The emergence of the CD301 and E202 in 2008 can be seen as a major step forward in addressing some of the contractual risks which have previously impeded the implementation of BIM. While these documents have differences, experience and further development of BIM will undoubtedly result in even further modifications to their respective formats and contents. At this stage it is too early to assess which form will become the more preferred as continued development and use of BIM will undoubtedly raise new issues and suggest new solutions. The purpose of this research is to explore some of the issues that have emerged as BIM addenda have been incorporated into construction contracts to get a preliminary understanding of how the project participants chose to address the problem as it applies to the commercial construction industry in the United States.

2 Methodology

This research was conducted to explore how the construction industry in the United States is incorporating the issue of BIM into their contract documents. The research was conducted by sending a survey to designers and contractors requesting their opinions, based on their own personal experience, on how BIM is being incorporated into their contract documents. The survey was an online survey facilitated through Zoomerang. Nineteen complete responses were obtained. One reason for the limited number of responses is simply a reflection that few in the building industry are truly engaged in the use of BIM. Three of the nineteen responses were from designers while the remaining sixteen were from contractors. This is not a statically significant population size but it did allow inferences to be ascertained which were later confirmed via interviews with respected industry professionals who were known to be heavily involved with BIM.

3 Results/Analysis

The first question was simply designed to classify the respondents by profession. As stated above, complete responses were obtained from three designers and sixteen contractors. These two respondent groups were split them into two different tracks in order to track their responses to the questions separately. The first question to the architects' group related to how they initiated their design on projects which utilized BIM. Most responded that they began their design with two-dimensional drawings (See Figure 1).

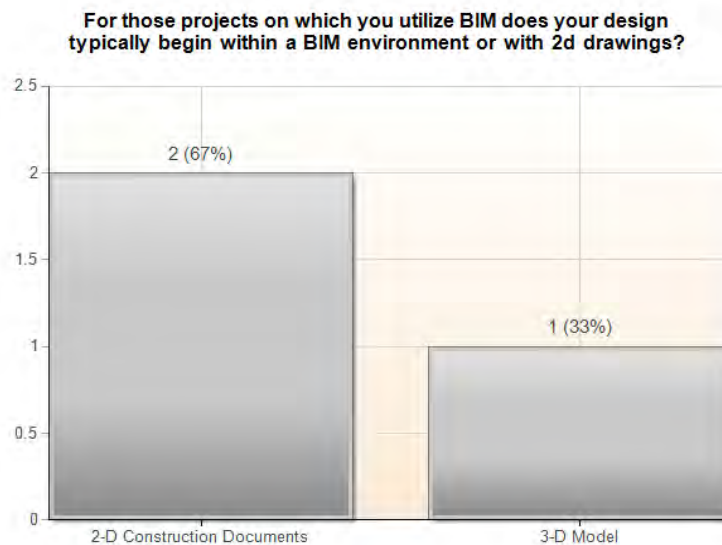


Figure 1.

The next question, asked to both designers and contractors, was which party they saw as responsible for creating the model. This question allowed the respondent to select all that apply but generally the designers (see Figure 2) and the contractors (see Figure 3) agreed that it was primarily the architect's responsibility to create the model, but other parties shared some modeling responsibilities. It appears that the responses to this question reflect the reality that multiple models are produced on a project but participants believe that it is the architect's primary responsibility to create a BIM model.

In your experience on projects which utilize BIM which project team member is typically responsible for creating the BIM model?

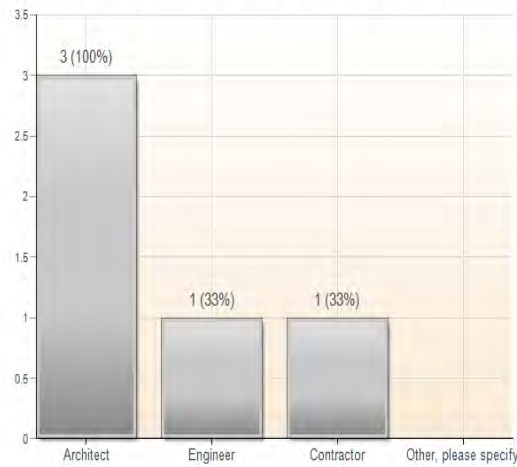


Figure 2.

In your experience on projects which utilize BIM which project team member is typically responsible for creating the BIM model?

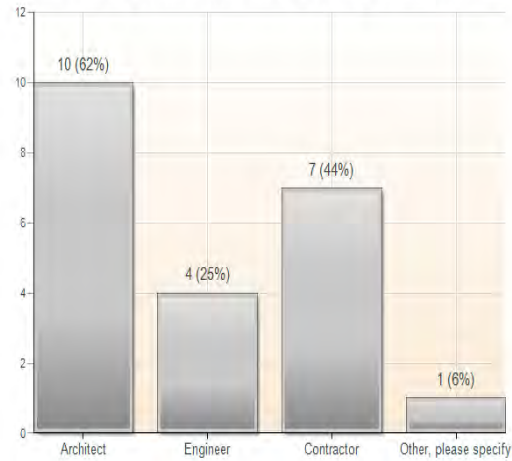


Figure 3.

The next question asked both the designers and contractors to comment on what they see as the primary role of BIM on a construction project. This question allowed the respondent to select all that apply but generally the designers (see Figure 4) and the contractors (see Figure 5) felt that the primary role of the model was to use it for coordination drawings followed by marketing.

What do you see as the primary role of BIM models on construction projects?

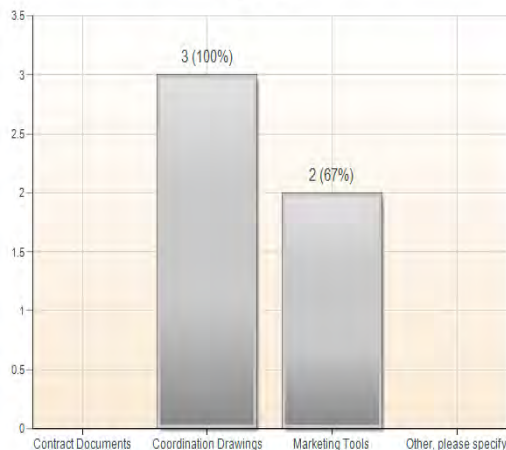


Figure 4.

What do you see as the primary role of BIM models on construction projects?

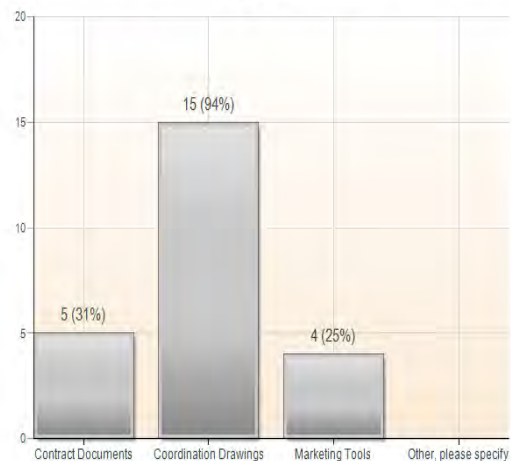


Figure 5.

The next question asked to both the designers and contractors was in their experience whether they had ever used the model as a contract document. This question allowed the respondent to select only one response and generally the designers (see Figure 6) and the contractors (see Figure 7) agreed that the model is typically not defined as contract documents.

On projects in which you were part of a project team and which utilized BIM models, how often does the contract define the BIM model as a contract document and allow all parties to the contract to rely on it as such.

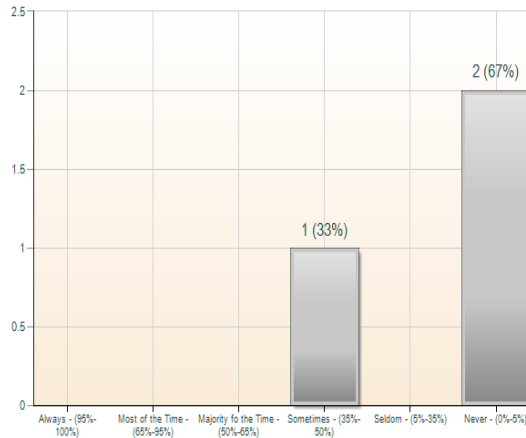


Figure 6.

On projects in which you were part of a project team and which utilized BIM models, how often does the contract define the BIM model as a contract document and allow all parties to the contract to rely on it as such.

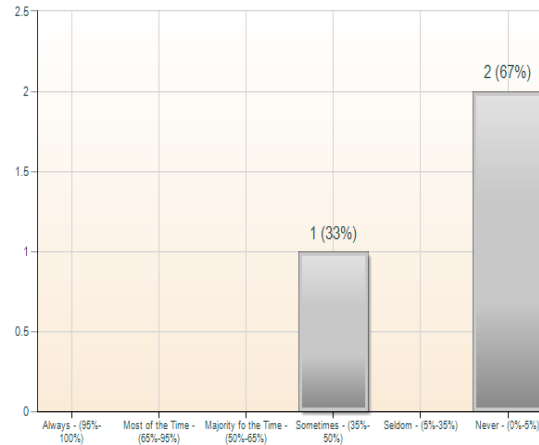


Figure 7.

The next question was asked to both the designers and contractors as a follow up to the previous question to clarify some of the inevitable confusion related to 2-D documents which were derived from the BIM. This question allowed the respondents to distinguish between 2-D documents derived from BIM and those created independent of BIM. Respondents were allowed to select all that apply and generally the designers (see Figure 8) and the contractors (see Figure 9) agreed that two-dimensional drawings created independent of BIM are most often defined as the contract documents.

For projects in which design information is created and communicated through a combination of BIM models, 2D drawings derived from BIM models, and 2D drawings created independent of a BIM models, which design information is primarily defined as the "Contract Documents."

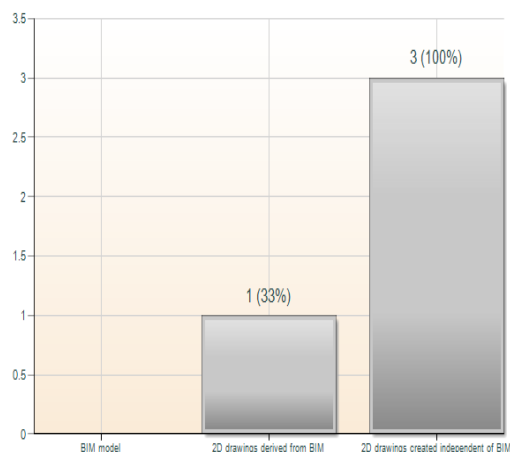


Figure 8.

For projects in which design information is created and communicated through a combination of BIM models, 2D drawings derived from BIM models, and 2D drawings created independent of a BIM models, which design information is primarily defined as the "Contract Documents."

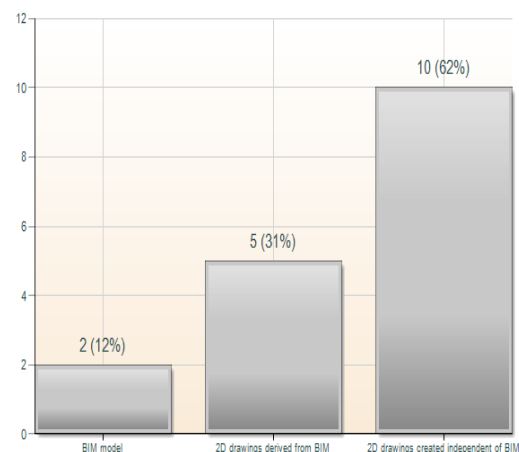


Figure 9.

The last question was to determine order of precedence on projects which defined models derived from BIM and two-dimensional drawings as contract documents. It asked that in the event of an ambiguity which instrument governs. This question allowed the respondent to select single response and generally the designers (see Figure 10) and the contractors (see Figure 11) agreed that if projects used both 2-D drawings and BIM as contract documents, the two-dimensional drawings would govern.

If you have been on projects which utilize a combination of BIM models and 2D drawings to define the "contract documents", typically how has the contract resolved an ambiguity (Order of Precedence).

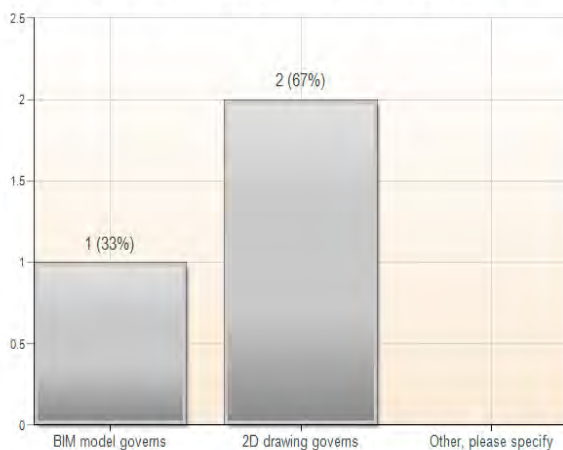


Figure 10.

If you have been on projects which utilize a combination of BIM models and 2D drawings to define the "contract documents", typically how has the contract resolved an ambiguity (Order of Precedence).

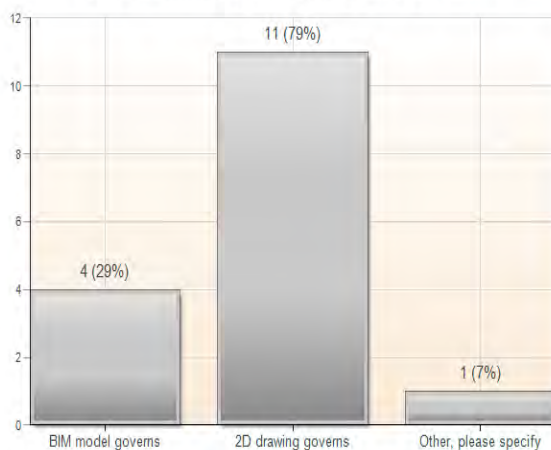


Figure 11.

Due to the limited number of responses and the potential for misunderstanding in a blind survey, personal interviews were conducted with five industry professionals who had significant experience and expertise with BIM. Four of the five worked for contractors and held positions related to their BIM expertise, and the remaining individual worked for the General Services Administration and had significant knowledge of, and experience with, BIM. The comments received from these five individuals were generally in line with the responses received from the survey participants. The following is a summary of the interview responses from the five industry professionals.

Two-dimensional drawings are still used as the primary contract documents.

For different reasons this group of industry professionals agreed that two-dimensional drawings remain as the primary source of contract documents. Even if the two-dimensional drawings were derived from models created with BIM, the 2-D drawings were favoured as the primary contract document. The general consensus was that on every project team there were still members who were not completely comfortable with the BIM environment and for that reason they continued to use the two-dimensional drawings as the primary source of information.

If both two-dimensional drawings and models developed with BIM are defined as the contract documents the two-dimensional drawings will govern.

Universally our respondents told us that if BIM models and two-dimensional drawings are defined as contract documents, the two-dimensional drawings will govern. This is simply

reinforcement that the traditional contract documents remain the favoured mechanism for determining a party's role and their responsibilities.

In the United States the primary use of BIM has been for clash detection and coordination.

Some are exploring different uses for BIM, such as estimating, scheduling, facility management, and as a contract document. However, it is widely held that the most common uses and benefits of BIM have been for clash detection and coordination purposes.

4 Conclusion

As with the advent of most new processes or technologies, it normally takes a period of time before a new process or technology is fully embraced. While BIM has been around for a number of years, it has only gained a much more widespread following over the past few years. Some of the reasons have to do with the cost of implementation; some have to do with learning curves and technical skills; and some are simply a reflection of the construction industry's reluctance to change. Each of these should go away over time as industry members become more knowledgeable of the benefits that can be achieved through the use of BIM and more comfortable with the use of the underlying technology.

However, other factors which contribute to this reluctance to take advantage of BIM include the shifting roles and increased risks and liability for some of the project participants. These inherent, underlying factors may not be as easy to overcome. The next step towards incremental change would seem to be that models derived from the BIM process and two-dimensional drawings created independent of BIM would both be defined as contract documents, and share an equal footing. It may take a number of years before the U.S. building industry adopts this position as significant contract changes are normally not forward looking but are a reflection of a legal system based on precedence. Eventually the need for efficiency and the desire to avoid duplication should dictate that the contract documentation be delivered via a single medium (BIM); and the traditional two dimensional contract documents will become obsolete.

References

- American Institute of Architects (AIA) (2008), "Document E202 – Building Information Modeling Protocol Exhibit", American Institute of Architects, Washington DC.
- Ashcraft, H.W. (2008), "Building Information Modeling: a Framework for Collaboration", *Construction Lawyer*, **Volume 28** (3).
- Azhar, S., Hein, M. and Sketo, B. (2008), "Building Information Modeling (BIM): Benefits, Risks and Challenges", *International Proceedings of the 44th Annual ASC Conference*, Associated Schools of Construction, Auburn University, Auburn, AL, US, April 2-5, 2008.
- ConsensusDOCS (2008), "ConsensusDOCS 301 - Building Information Modeling Addendum", ConsensusDOCS, Arlington, VA.

- Haynes, D. (2009), “A Tale of Two Exhibits – A Comparison of the CD301 and E202 BIM Exhibits”, *Construction Watch*, **Volume 3** (09), McElroy, Deutsch, Mulvaney & Carpenter, LLP, Boston, MA.
- Hurtado, K. A. and O’ Connor, P. J. (2008), “Contract Issues in the Use of Building Information Modeling”. *International Construction Law Review*, **Volume 25** (3).
- Larson, D. A. and Golden, K.A. (2008), “Entering the Brave New World: An Introduction to Contracting for BIM”, *William Mitchell Law Review*, **Volume 34**.
- Lowe, R. H., and Muncey, J. M. (2009), “ConsensusDOCS 301 BIM Addendum”, *Construction Lawyer*, **Volume 29** (1), p 17.
- Post, N. M. (2009a), “Building Team Members See Progress and Problems”, *Engineering News Record*, **Volume 262** (12), p 28.
- Post, N., M. (2009b), “Leading-Edge Collaboration Hurt by Lots of Software Workarounds”, *Engineering News Record*, **Volume 262** (14), pp 28-29.

Legal issues of partnering in construction: an analysis of case law

Faraziera Mohd Raslim¹, Nur Emma Mustaffa¹ and Jamaluddin Yaakob¹

¹Department of Quantity Surveying
Faculty of Built Environment,
Universiti Teknologi Malaysia Skudai, Johor

Email: farazi3ra@yahoo.com.my; b-nuremma@utm.my; jbyaakob@hotmail.com

Abstract:

Partnering has been acknowledged by the team players in construction industry as one of the approaches to tackle the adversarial culture in the industry. It effectively lowers the risk of cost escalation and delays, as a result from better time control and cost control over the construction project. Through trust, transparency, open communication, co-operation and mutual objectives, it introduces a win-win approach to team players in the construction industry. Though it has been debated that partnering is not about contracts, there are legal issues to be measured from the adoption of partnering concept. This paper examines all the relevant case law which issues are on partnering in the construction industry. Legal issues pertaining to partnering will be determined from these cases and adopted as the ground for the formation of partnering contract. The ultimate aim of this study is to produce a framework of the partnering contract which is suitable for the Malaysian construction industry to be as a reference in the implementation of partnering.

Keywords:

Construction projects, legal issues, Malaysia, partnering.

1 Introduction

Dispute resolution has become an ordinary characteristic of the usual construction project process.¹ Time and money have been lost in defensive posturing, developing claims and litigating. Normally, the building team members make decisions based on their own goal and objectives without considering the impact on the other party. This can lead to an adversarial relationship. Various studies agree that this situation significantly affects the productivity and efficiency of the construction industry.²

¹ Cornick, T and Marther, J (1999) "Construction project teams: making them work profitably." London: Thomas Telford.

² Kodefors A. (2004) "Trust in project relationships-inside the black box". International Journal Project Management 22 (3), 175–182; Scott B. (2001) "Partnering in Europe: incentive based alliancing for projects". London: Thomas Telford; Murphy, Michael (1995). "A contractor's Partnership Baptism". In Godfrey, Kneeland A. (1995). Partnering in Design and Construction. New York: McGraw-Hill.

Consequently, this relationship has contributed to the need for a better approach, and one of it is partnering; which is an innovative concept to construction organizations, which traditionally rely heavily on contracting to bind the parties together.³

The formal introduction of partnering into the Malaysian Construction Industry was done in 2007 through the Construction Industry Master Plan (CIMP),⁴ but before this, in 2000, a joint seminar on project based partnering in Public Works Department (PWD) was organized by the department and the Master Builders Association Malaysia titled “Introducing the Concept of Partnering in the Construction Industry”

2 Definition of partnering

Partnering, a form of relationship contracting, was pioneered in the United States by the United States Army Corps of Engineers in the late 1980's and is now gaining popularity in the United Kingdom, Australia, Canada and Hong Kong. The Dispute Avoidance and Resolution Task Force (DART) of the American Arbitration Association defined partnering as follows:

“Partnering is a voluntary, organized process by which two or more organizations having shared interests perform as a team to achieve mutually beneficial goal. Typically, the partners are organizations that in the past worked at arm's length or may have had competitive or adversarial relationship with one another. Generally, the more partners involved, the better the overall results. Partnering is also a collaborative process that focuses on co-operative solving reconciliation (win-win) as opposed to either compromise (lose-lose) or concession (win-lose). It is not a social process that simply promotes courtesy and politeness among participants, but rather good faith joint resolution of problems. Partnering is a non binding process. It neither alters the contract documents nor the relationships between the parties to use the agreed-upon partnering process and to deal with one another as true partners.”

Partnering has thus been described as a ‘procedural overlay’⁵ to the underlying contractual relationship between the parties. While the contractual provisions deal with rights, obligations and the allocation of risk between the contracting parties, the partnering charter seeks to address the way parties deal with each other and interact to reduce conflict. While the legal relationships are established by the contract, the partnering process establishes the working relationships between the parties based on trust, co-operation and honesty.⁶

³ Cheng E.W.L and Heng Li (2004) “Development of a practical model of partnering for construction projects.” *Journal of Construction Engineering and Management*, 130(6), 790-798; Latham Sir M (1994) “Constructing the team: Final report of the government industry review of procurement and contractual arrangements in UK industry”. London: HMSO.

⁴ Construction Industry Development Board (2007) “Construction industry master plan 2006-2015.” Kuala Lumpur: Construction Industry Development Board, (ISBN 978-983-2724-44-5)

⁵ Heal, A. J. (1999). “Construction partnering: Good faith in theory and practice.” *Construction law Journal*, 167–198.

⁶ Chan, A.P.C, Chan, D.W.M, Fan, L.C.N, Lam P.T.I and Yeung J.F.Y (2007) Achieving partnering success through an incentive agreement: lessons learned from an underground railway extension project in Hong Kong. “*ASCE Journal Management in Engineering*”, 24(3), 128-137.

3 The partnering process

The formal partnering process begins after award of contract with an initial 'team-building' training workshop led by a neutral facilitator. This workshop should be attended by the key representatives of all partnering organizations to help establish commitment. The workshop generally begins with activities designed to help participants feel comfortable with each other followed by training on the principles of effective partnering and communication skills. Participants then develop common goals and agree on a process to resolve disputes quickly, in a co-operative manner and effectively. Approaches on how to remove barriers to communication, how to delegate decision making to members of the team and how to bring about continuous quality improvement are also discussed. A Partnering Charter is generated setting forth the intent of the parties to work together towards a successful project. All parties should sign and display the Charter on a wall to remind team members of the commitments made. Subsequent workshops are held to improve teamwork, maintain the partnering spirit and involve all team players at the working level. As the program evolves, periodic follow-up workshops are held to reinforce the desired behaviour and introduce new team members.⁷

4 Case Law

Regardless of the acknowledged benefits of partnering in construction,⁸ there are certain legal concerns as to the effect of the partnering arrangement on the rights and obligations of the parties as well as on the question of risk allocation between the parties to the contract. As always, these issues will be much of concern in the event that the relationship between the parties breaks down and dispute arise. Therefore, this paper will examine all the relevant case law which issues are on partnering in the construction industry. Legal issues pertaining to partnering will be determined from these cases and adopted as the ground for the formation of partnering contract since the ultimate aim of this study is to produce a framework of the partnering contract which is suitable for the Malaysian construction industry to be as a reference in the implementation of partnering.

4.1 Birse Construction Ltd v. St David Ltd [1999] BLR 194

This is the first judicial opinion on the enforceability of partnering agreements was seen in the first instance decision by the Judge. Although the decision was overturned on appeal, neither the court dealt substantively with the partnering charter. Therefore, the first instance decision stands as the best guideline to the court's approach to partnering for the time being. In this case, the initial project was intended to lead to a long term

⁷AGC (1991) "Partnering: concept for success". Washington DC: Associated General Contractors of America; CII Construction Industry Institute (1991). "In search of partnering excellence". Publication no.17-1, Report CII, Austin, Texas; Eriksson, E and Nilsson, T (2008) "Partnering the construction of a Swedish pharmaceutical plant: case study". Journal of Management in Engineering; Chan, A.P.C, Chan, D.W.M, Fan, L.C.N, Lam P.T.I and Yeung J.F.Y (2007) Achieving partnering success through an incentive agreement: lessons learned from an underground railway extension project in Hong Kong. "ASCE Journal Management in Engineering", 24(3), 128-137.

⁸ Chan APC, Chan DWM, Ho KSK (2003) An empirical study that of the benefits of construction partnering in Hong Kong. "Construction Management and Economics", 21(5), 523-533.

relationship. The case involved building construction by contractor Birse for developer St David and was concerned with establishing whether a formal JCT form of building contract existed between them for the initial work done over a fifteen month period. They agreed a partnering charter and under the heading “relationships” they agreed to “promote an environment of trust, integrity, honesty and openness”. However, in just over a year, Birse had issued proceedings against St David claiming that there was no contract and issued a writ for payment on a quantum meruit basis for work done on the project. St David argued that there was a contract incorporating the JCT Contract. This contained an arbitration agreement and therefore the proceedings should be stayed for arbitration.

Subsequent to the partnering agreement, Judge Humphrey Lloyd was satisfied that a JCT contract had been formed as a result of the course of dealing, together with the fact that the parties had never excluded the formation of contract unless and until formal documents were prepared and executed. The Judge also commented;

“I have a little doubt that the parties considered that the ‘partnering’ arrangement that they had made, as exemplified by the Charter, made it unnecessary. People who have agreed to proceed on the basis of mutual co-operation and trust, are hardly likely at the same time to adopt a rigid attitude as to the formation of a contract.”

Whilst the Judge rejected any legal status for the partnering charter itself he described it as follows;

“Even though the terms of the Charter would not alter or affect the terms of the contract (where they are not incorporated or referred to in the contract or are not binding in law in their own right) an arbitrator (or court) would undoubtedly take such adherence to the charter into account exercising the wide discretion to open up, review and revise, etc which is given under the JCT conditions.”

Judge Lloyd appeared to take the view that if a party had agreed to co-operate by way of partnering charter, it could not then go back on its word by failing to co-operate. So, while the partnering agreements are non-binding, it appears open for the courts to take these into account and imply high standards of moral conduct into the contract.

4.1.1 Principles of interpretation of contractual terms

Following Birse’s case, the courts may now use provisions of a non-binding partnering agreement to interpret the contractual terms of the parties under the substantive construction contract. It is established that although it is not open to look at parties’ subjective intentions, regard can be had to the ‘matrix of fact’ or the factual background in interpreting the terms of contract.⁹ The principles of interpretation were summarized by Lord Hoffman in the recent House of Lords decision of *Investors Compensation Scheme Ltd v West Bromwich Building Society*¹⁰ as follows;

⁹ *Prenn v Simmonds* [1971] 1 WLR 1381; *Reardon Smith Line Ltd v Yngvar Hansen-Tangen* [1976] 1 WLR 989; *Investors Compensation Scheme Ltd v West Bromwich Building society* [1998] 1 WLR 896.

¹⁰ [1998] 1 WLR 896 at p 912

1. Interpretation is the ascertainment of the meaning which the document would convey to a reasonable person having all the background knowledge which would reasonably have been available to the parties in the situation in which they were at the time of the contract;
2. Subject to the requirement that it should have been reasonably available to the parties and to the exception to be mentioned next, it includes absolutely anything which would have affected the way in which the language of the document would have been understood by a reasonable man;
3. The law excludes from the admissible background the previous negotiations of the parties and their declaration of subjective intent;
4. The background may not merely enable the reasonable man to choose between the possible meanings of words which are ambiguous but even (as occasionally happens in ordinary life) to conclude that the parties must, for whatever reason, have used the wrong words or syntax.

It is likely that a partnering agreement may be considered as a component of this 'background' as it is a document which should have been reasonably available to the parties and does not fall within the two exceptions enunciated by Lord Hoffman, i.e. previous negotiations and declarations of subjective intent, as it embodies the mutual commitments made by the parties. Thus, the terms of partnering charter may have a bearing on the meaning of the word 'reasonable' satisfaction of the architect or engineer.¹¹

4.2 Thiess Contractors Pty Ltd v. Placer (G ranny Smith) Pty Ltd [1999] WASC 1046

This is an Australian case relating to principle of good faith. Placer employed Thiess as earthmoving contractors in an opencast mining operation under a partnering contract. This arrangement included an 'open book' system for costing the mining operations with payments to Thiess being based on these costs plus an agreed profit margin. This system "involved the disclosure by Thiess of confidential information concerning the way in which it derived its rates for carrying out various mining operations". The contract stated that "the successful operation of this Contract requires that (Thiess) and (Placer) agree to act in good faith in all matters relating both to the carrying out (of) the works, derivation of rates and interpretation of this document". This clause, therefore, placed a specific good faith requirement on the derivation of plant rates and general good faith requirements on the works and the interpretation of the contract as a whole.

A dispute arose over the plant rates whereby Placer alleged that they had not been calculated on the agreed costing basis and had been inflated both during the negotiations leading up to the execution of the partnering contract and during the period of contract. Placer subsequently invoked a general termination clause in the contract. Thiess sued claiming that Placer had breached a fiduciary duty by terminating the contract. Placer counterclaimed for damages for the inflated plant rates and claimed that Thiess had owed it fiduciary obligations in respect of both the pre-contract negotiations on plant

¹¹ Pradhan, V., (2002) "Partnering in construction contract" Seminar on 'Partnering & Best Practice in Construction & The Work of Sir Michael Latham'. James R Knowles.

rates leading up to the partnering contract and on the subsequent assessment of those rates after execution of the contract.

Judge Lloyd found that during the pre-contract negotiations the parties were 'at arms length' and was not persuaded that the relationship between Thiess and Placer up to the point at which they entered in to the contract, was a fiduciary one. However after execution of the partnering contract the Judge gave extensive consideration to good faith and fiduciary duties. He pointed to contract requirements stated in general terms "that are typical of many....which do not define fine rights and obligations with any precision. Their implementation clearly requires goodwill and co-operation on the part of both parties. "good faith" must include those matters." Using this basis he went on to interpret the wording and application of the good faith clause in detail.

The judge limited the expression "all matters relating to the works" to meaning all matters related to the "carrying out" of the works and he interpreted the good faith obligation as "requiring the parties to act honestly with each other and to take reasonable steps to co-operate in relation to matters where the contract does not define rights and obligations or provide any mechanism for the resolution of disputes"

It is submitted that 'reasonable steps' in this context would at least imply a duty to negotiate in relation to these undefined rights and obligations. It is also submitted that such a duty must also involve the implication of a duty of good faith in relation to the negotiation process because, as observed by Einstein J in relation to negotiation and mediation processes in another Australian case of *Aiton Pty Ltd v Transfield Ltd*¹², "without it there is no chance of reaching a mutually satisfactory conclusion". In that case a duty to negotiate and mediate in good faith was expressly stated in the contract and was generally considered to be certain and enforceable by the Judge. His definition of the content of the express duty of good faith in this situation amounted to the display of an appropriate level of pro-active participation in the specified negotiation and mediation process.

4.2.1 Principle of good faith

The conclusions from this case seem to be that in Australian law general statements of good faith may require the parties to take reasonable steps to co-operate in relation to matters where the contract does not define rights and obligations or provide any mechanism for the resolution of disputes; this suggests the implication of a duty to negotiate in good faith; such a duty would be rejected in English law as a result of *Walford v Miles*¹³ and it is submitted that this would also be the case in Malaysia.

It may be argued that a duty to act in good faith is implied into the contract by the parties' decision to employ the partnering arrangements which in itself is based on principles of good faith- mutual trust and objectives, open communication, honesty, confidence and dispute minimization.¹⁴ Further, partnering is a concept developed and imported from the United States of America which recognizes the doctrine of good faith in contract law. There is thus far no recognized general duty to perform contractual

¹² [1999] SC NSW 236

¹³ [1990] 12 EG 107

¹⁴ Colledge, B. (2000) „Obligations of Good Faith in Partnering of UK Construction Contracts., "International Construction Law Review", 17:1, 175-201.

obligations in good faith in Malaysian or English law.¹⁵ There are however instances of judicial observations¹⁶ that a discretion should not be exercised in bad faith, on the basis that there is a negative duty not to act in bad faith as opposed to a positive duty to act in a good faith. With reference to the recognition of civil law systems of a general principle of good faith, Bingham LJ in the Court of Appeal decision of *Interfoto Picture Library Ltd v Stiletto Visual programmes Ltd*¹⁷ stated;

“English law has, characteristically, committed itself to no such overriding principle but has developed solutions in response to demonstrated problem of unfairness”

Certainly, in respect of negative obligations, the courts would appear to be willing already to uphold these principles implicitly. In doing so, they make reference to general standards of honesty,¹⁸ fair dealing and reasonableness to determine abuse or otherwise of discretionary powers.¹⁹ The existence of a partnering charter provides further evidence of the background to the contract and together would suggest enhanced expectations between the parties. Furthermore, the nature of the parties themselves may have an impact upon the imposition of obligations, with the weaker parties being protected by the courts from opportunism by the stronger party.²⁰

For the avoidance of doubt as to whether a positive duty to act in good faith is unintentionally implied into a contract as a result of the partnering arrangement to avoid any unforeseen shifting risk allocation provisions of the contract.

4.3 Amalgamated Investment & Property Co. Ltd. (In Liquidation) v Texas Commerce International Bank Ltd [1982] Q.B. 84

The plaintiffs, a property company registered in England, had a subsidiary company, A.N.P.P., registered in the Bahamas. The defendant bank, who were registered in England had early in 1970 lent the plaintiffs U.S. \$3,000,000 secured by mortgages on properties in England. By letter of September 23, 1970, the bank agreed to make a further loan of U.S. \$3,250,000 to A.N.P.P. to be secured by a mortgage on a building on property which A.N.P.P. owned at Nassau in the Bahamas together with a guarantee from the plaintiffs for the sum loaned. On September 28, 1970, the plaintiffs executed a guarantee to pay to the bank on demand all moneys which might at any time be due or owing to the bank by A.N.P.P. For exchange control purposes the bank arranged for the loan to A.N.P.P. to be channelled through Portsoken, a subsidiary company which the bank acquired in the Bahamas. On December 31, 1970, A.N.P.P. executed a mortgage on the Nassau building in favour of Portsoken to secure an advance of U.S. \$3,250,000.

¹⁵ Such a general duty has been postulated by some legal commentators: eg Lucke, 'Good Faith and Contractual Performance' in Finn, *Essays on Contract* 1987 Law Book Company at 161; O'Byrne, 'Good Faith in Contractual Performance: Recent Developments' *Canadian Bar Review* Vol LXXIV [1995] 71; Carter and Furmston, 'Good Faith and Fairness in the Negotiation of Contracts', [1995] 8 *Journal of Contract Law* 93. See also dicta of Proestly JA in *Renard Constructions (ME) Pty Ltd v Minister for Public Works* [1992] 26 NSWLR 234 at p 268.

¹⁶ See *Chay Chong Hwa & Ors v Seah Mary* [1984] 2 MLJ 251 at p 256. In *Bains Harding (Malaysia) Sdn Bhd v Arab-Malaysia Merchant Bank Bhd & Ors* [1996] 1 MLJ 425, Talalla J at p 443 observed that where a termination for convenience clause in a contract is exercised not for convenience but for the purpose of employing another sub-contractor so that the other sub-contractor can enjoy whatever profits remain in the contract, then such use of the clause amounts to the employer acting in bad faith and in an unconscionable manner.

¹⁷ [1989] 1 QB 433 at p 439.

¹⁸ *Neodox Ltd v Borough of Swinton and Pendlebury* [1958] 5 BLR 34

¹⁹ *Renard Construction (ME) Pty Ltd v Minister for Public Works* [1992] 26 NSWLR 234

²⁰ *Timeload Ltd v British Telecommunications plc* [1995] EMLR 459 (CA)

That sum was entered in the books of the bank and Portsoken respectively as a loan by the bank to Portsoken and then by Portsoken to A.N.P.P., but the contemporary and subsequent negotiations and transactions between the parties in respect of the loans were upon the basis that the guarantee relating to the Nassau loan of U.S. \$3,250,000 made the plaintiffs liable to the bank in respect of that loan. In reliance upon the parties' belief that under the guarantee the plaintiffs were bound to discharge any indebtedness of A.N.P.P. to Portsoken in respect of the loan, the bank between 1974 and 1976 allowed the loans to remain outstanding despite opportunities to call them in.

By 1976 the plaintiffs were in severe financial difficulties. The loan of U.S. \$3,250,000 was not repaid and the Nassau building was sold for U.S. \$2,500,000, leaving a deficit of \$750,000 for which the guarantee was the only security. After the plaintiffs had defaulted on the English loan, the English properties were also sold leaving a surplus of about U.S. \$750,000 which the bank purported to apply in discharge of the deficit on the Nassau loan.

The plaintiffs, who in May 1976 were ordered to be compulsorily wound up, acting through their liquidator claimed a declaration that they were under no liability to the bank either under the guarantee or otherwise in respect of the sums owing by A.N.P.P. under the loan to A.N.P.P. Robert Goff J. dismissed the plaintiffs' claim holding that although the bank could not rely directly upon the guarantee the plaintiffs were estopped from contending that it did not cover A.N.P.P.'s liability in respect of the Nassau loan.

The plaintiffs appealed on the grounds that the judge was wrong in law and/or in fact in holding that by reason of the conduct of the plaintiffs they were estopped from denying that they gave a binding and effective guarantee to the defendants of money owed by a subsidiary of the plaintiffs to a subsidiary of the defendants.

The court explained the nature of an estoppel by convention. Lord Denning MR said: "When the parties to a transaction proceed on the basis of an underlying assumption – either of fact or of law – whether due to misrepresentation or mistake makes no difference – on which they have conducted the dealings between them – neither of them will be allowed to go back on that assumption when it would be unfair or unjust to allow him to do so. If one of them does seek to go back on it, the courts will give the other such remedy as the equity of the case demands." and "If parties to a contract, by their course of dealing, put a particular interpretation on the terms of it -- on the faith of which each of them -- to the knowledge of the other -- acts and conducts their mutual affairs -- they are bound by that interpretation just as much as if they had written it down as being a variation of the contract. There is no need to inquire whether their particular interpretation is correct or not -- or whether they were mistaken or not -- or whether they had in mind the original terms or not. Suffice it that they have, by their course of dealing, put their own interpretation on their contract, and cannot be allowed to go back on it."

The Court of Appeal held that since at the relevant time the parties had acted upon the common assumption that the plaintiffs were liable under the guarantee in respect of the

advance which was now in dispute, the claimants were estopped by convention from denying their liability under the guarantee.

4.3.1 *Estoppel and waiver*

It is possible that during the course of the partnering arrangements, statements will be made by the parties to each other that may not be in tandem with the strict provisions of the contract between them. In the event of dispute, it would thus be open to the party to whom such statements are made to rely upon the doctrine of estoppel to argue that the party making the statements cannot now seek to enforce its strict contractual rights in the light of the statements made during the course of the performance of the contract.²¹

In another law case related to estoppel, the Malaysian courts, in the landmark Federal Court decision of *Boustead Trading Sdn Bhd v Arab-Malaysian Merchant Bank Bhd*,²² have recognized that the doctrine of estoppel is a flexible principle by which justice is done according to the circumstances of the case and that it is a doctrine of wide utility which has been resorted to in varying fact patterns to achieve justice. The doctrine may operate to bind parties as to the meaning or legal effect of a document or a clause in a contract which they have settled upon or which one party to the contract has represented or encouraged the other to believe as the true legal effect or meaning. Further, the doctrine may be applied to enlarge or reduce the rights or obligations of parties under a contract. Both plaintiffs and defendants may have recourse to the doctrine which is not confined to cases where a representation of fact has been made.²³

To determine the applicability of the doctrine of estoppel to a particular set of facts, the question that a court would ask is whether it would be just that a particular litigant should, in the light of his conduct, succeed in that action given the peculiar facts of the case. The doctrine would only apply if the answer to this question is negative.²⁴ It is very likely that partnering arrangement would be an important element of the circumstances of a case that the court would have regard to in deciding on the applicability of the doctrine of estoppel to that set of facts. This is because parties often have to act in a manner which is inconsistent with the strict provisions of the contract so as to achieve their mutual objectives in accordance with the spirit of the partnering charter.

In the circumstances, parties to a partnering arrangement must be properly advised on the very real possibilities that their contractual rights may have been compromised owing their acts and/or omissions in carrying out the shared objectives of partnering. A suggested solution is to include in the partnering agreements provisions to the effect that representations, in whatever form by the parties made during the course of the partnering arrangements, will not be contractually binding in the absence of an appropriate amendment to the contract in writing. Thus, the necessary amendments to the relevant contractual provisions must be made in writing by the parties in the event that it is the intention of the parties that the said representations should affect or alter their contractual obligations pursuant to the contract.

²¹ Jamieson B.K (2001) "Partnering-Some legal issues" *Contrax Asia Ltd*.

²² [1995] 3 MLJ 331

²³ *Teh Poh Wah v Seremban Securities Sdn Bhd* [1996] 1 MLJ 701

²⁴ *Chor Phaik har & Ors v Choong Lye Hock Estates Sdn Bhd & Ors* [1996] 2 MLJ 206

4.4 Eastover Ridge, LLC v. Metric Constructors, Inc., 2000 N.C. App.

This case applied the fiduciary relationship's principles. Here is the fact of the case, on 22 July 1994, owner entered into an agreement with the contractor for the construction of 216 apartment units in nine buildings, a clubhouse/leasing building, pool, tennis courts, maintenance building, certain landscape features, and associated site work. The contractor accepts the relationship of trust and confidence established by Agreement and covenants with the owner to cooperate with the Architect and utilize the contractor's best skill, efforts and judgment in furthering the interests of the owner. The owner filed suit against the contractor, alleging unfair and deceptive business practices because the contractor breached its fiduciary duty to the owner resulting in constructive fraud. According to the owner, the contractor did not properly perform all of its contractual duties. Owner initiated this action on 22 October 1996 and filed an amended complaint four days later, asserting claims for breach of contract, breach of fiduciary duty and constructive fraud, unfair trade practices, and equitable relief of recoupment and setoff.

The Judges stated: "Plaintiff alleges in its complaint that defendant failed and refused to perform its obligations under the agreement and lists examples of defendant's breaches. Although plaintiff has made numerous allegations that defendant breached its agreement regarding the construction project, we note that a certificate of substantial completion was signed by the architect on 27 March 1996 and that the construction project was subject to local government inspection. Therefore, we conclude that plaintiff has failed to show sufficient aggravating circumstances to establish a claim for unfair and deceptive trade practices." The court held that because the architect was hired to perform construction administration and was to keep an eye on the contractor, there was no fiduciary relationship between the owner and this contractor.

Although the court's decision rejected the fiduciary duty of the contractor towards the owner, still the issue of fiduciary relationship is questioned in this case.

4.4.1 Fiduciary relationship

The question of this issue is whether the existing partnering arrangement imposes the fiduciary obligations to be owed by the parties to each other? Is a partnering agreement akin to a joint venture agreement where such fiduciary obligations may arise? ²⁵ A relationship is termed a fiduciary relationship if one party owes a duty to act with care and good faith in the interests of the other party. For example, the relationship between trustee and beneficiary is fiduciary as is the relationship between an employer and employee, principal and agent, director and company and amongst partners. ²⁶

²⁵ *Newacres Sdn Bhd v Sri Alam Sdn Bhd* [1991] 3 MLJ 474; *Hartela Contractors Ltd v Hartecon JV Sdn Bhd & Anor* [1999] 2 CLJ 788 where Gopal Sri Ram JCA found that in view of the mutual trust and confidence between parties to a joint venture essential for a proper working of the relationship, where there was reliance by one party upon the skill or expertise of the other, there was a duty upon the other to use his best endeavours to ensure the success of the venture. Equity would imply such an obligation in the absence of an express term in the joint venture agreement.

²⁶ It has been said that the categories of fiduciary relationships are not closed: per Mason J in *Hospital Products v United States Surgical Corporation* (1984) 156 CLR 41. An argument that project managers owed fiduciary obligations to property developers was rejected by Judge Hicks QC in *Chesham Properties Ltd v Bucknall Austin Project management Services Ltd* (1996) 82 BLR 92.

Consequently, a party under a fiduciary obligation must avoid any conflicts between its duty to others and its own self-interest and it must not misuse its fiduciary position for personal gains.

Generally, parties to an ordinary commercial contract owe no fiduciary obligations to each other. However, owing to the very nature of partnering agreement, which is based on principle of trust, fairness and mutual co-operation for the benefit of the project, it is arguable that parties have put themselves in a position where they are placing reliance upon each other to act in each other's interest.

The parties to a partnering arrangement must consider whether they owe fiduciary obligations to each other which impinge upon their freedom to act in their own self-interest.²⁷ Parties who wish to avoid the consequences of fiduciary obligations should do so by express provisions in the partnering charter. A provision could be inserted to totally exclude the possibility of those obligations from arising. Alternatively, if the participants decide that fiduciary obligations are consistent with the spirit of the agreement, the charter could define the scope of the obligations for the purposes of the partnering process.

4.5 P Ward v Civil and Civic [1999] NSWSC 727

These proceedings relate to a dispute concerning the earthworks component for the construction of the Prospect Water Filtration Plant from 1993 to 1996. Civil and Civic Pty Limited [now known as Lend Lease Property Services Pty Limited] was the design and construct contractor for that plant and were negotiating a sub-contract for earthworks with Ward which included a design development aspect. During these negotiations the parties had attended a partnering meeting which had resulted in the signing of a partnering charter. In this charter the parties expressly agreed that their relationship would be "to work together to achieve our mutually developed goals via the collective utilization of our joint skills in an environment of open and honest communication". This charter was signed before the conclusion of the formal sub-contract but its legal status does not appear to have been specifically defined by the parties.

When the formal sub-contract was eventually prepared, it did not contain any references to the partnering charter, but it did include a definition development which Ward subsequently claimed was different from that discussed during negotiations. Ward had apparently failed to notice this change before the sub-contract was signed and they claimed that this placed an additional financial liability on them.

A dispute arose over payment for the costs of design carried out by Ward during the course of the sub-contract. Ward claimed that Civil and Civic had represented to them that the partnering relationship between them "would be in the nature of a partnership and that they would co-operate to ensure that the project was a financial success for both the plaintiff and the Defendant" and that in this context Civil and Civic's failure to draw Ward's attention to the changed liabilities amounted to misrepresentation.

²⁷ Jamieson B.K (2001) "Partnering-Some legal issues." Contrax Asia Ltd.

However the Judge advised that “the most obvious of the difficulties is the fact that Wards seeks to disavow a formal written contract signed by both parties following close dealings between them over a period of almost a year, which dealings were, on my findings, clearly understood by both parties as intended to culminate in the execution of the Subcontract”.

4.5.1 *Misrepresentation*

From the case before it seems that the issue of misrepresentation arose between the parties, where one party failed to inform the other parties the addition of financial liability.

Given the wide ambit of the Misrepresentation Ordinance, it is foreseeable that disenchanted participants in partnering will seek to assert that, the representations made in the course of the process amounted to misrepresentation conduct.²⁸

When a statement during negotiations leading to a contract is deemed not a term of the contract or a warranty collateral to the contract, it is a ‘representation’. Such pre-contractual statements, though not intended to have contractual effect, may have seriously affected the inclination of one party to enter into the contract. If the representation is false, the party that was misled cannot claim damages for breach of contract because no term of the contract has been broken. However the misled party may claim relief accorded by the law for misrepresentation or fraud.²⁹

In other case, *Thiess Contractors Pty Ltd v Placer (Granny Smith) Pty Ltd* [1999], the Judge found that during the pre-contract negotiations the parties were at ‘arm’s length’ and was not persuaded that the relationship between Thiess and Placer up to the point at which they entered into the contract, was a fiduciary one. He considered that “there has been a misrepresentation: not a breach of fiduciary duty”.

In partnering, the team members rely on information supplied by the other parties, each and every party in the team will provide information where they think useful and important to conduct the construction project. As the team agreed, any critical problem will be discussed and the team will try to find a suitable solution, basically, to ensure the benefits of the partnering process are maximized, team member will enter into partnering’s relationship on the basis of full and frank disclosure. Problems occurred when the team members did not reveal everything regarding to the project, they might think that the information is not important to be shared, while the other members need to know about it in order to avoid losses. When this situation occurred, it is foreseeable that the disenchanted team members in partnering will seek to involve the course of misrepresentation conduct. It is therefore critically important that the representatives are fully aware of the import of their conduct and any statements made during the development and conclusion of any partnering arrangement.

²⁸ Jamieson B.K (2001) “Partnering-Some legal issues.” *Contrax Asia Ltd*.

²⁹ Wallace, I.N.D., (1999) “The forgotten client : A history of construction contract”, in twelfth annual construction centre, centre of construction law and management, King’s College London, UK. Pt. VIII. : 1-53.

5 Conclusion

The need for change within the construction industry is paramount to its survival. Those who fail to recognize and act within this scenario will struggle to succeed.³⁰ So, partnering activities by public and private organizations today are starting to explore that territory and are demonstrating that change is possible. Partnering has proved successful in the United Kingdom, the United States, Hong Kong and Australia. The lessons learnt from those who have adopted it will aid local construction practitioners to minimize the risks brought by possible difficulties from implementing the concept of partnering; in which legal issues is one of them. This paper has discussed the legal issues that occur from this concept; the issues derived from the research will later be used to fulfill the ultimate objective of the study, to produce the guidelines of partnering contract which will serve as a suitable reference in the implementation of partnering in the Malaysian construction industry since there is no contract for partnering created yet.

6 Acknowledgements

My appreciation goes to Universiti Sains Malaysia (USM) and the Ministry of Higher Education Malaysia for their financial support given to me throughout my studies.

7 References

- AGC (1991) "Partnering: concept for success". Washington DC: Associated General Contractors of America.
- Chan APC, Chan DWM, Ho KSK (2003) "An empirical study that of the benefits of construction partnering in Hong Kong." *Construction Management and Economics*, 21(5), 523-533.
- Chan A.P.C, Daniel W.M Chan, Y.H Chiang, B.S Tang, Edwin H.W Chan and Kathy S.K Ho (2004) "Exploring critical success factors for partnering in construction projects." *Journal of Construction Engineering and Management*, 130(2), 188-198.
- Chan, A.P.C, Chan, D.W.M, Fan, L.C.N, Lam P.T.I and Yeung J.F.Y (2007) "Achieving partnering success through an incentive agreement: lessons learned from an underground railway extension project in Hong Kong." *ASCE Journal Management in Engineering*, 24(3), 128-137.
- Cheng E.W.L and Heng Li (2004) "Development of a practical model of partnering for construction projects". *Journal of Construction Engineering and Management*, 130(6), 790-798.
- CII Construction Industry Institute (1991). "In search of partnering excellence". Publication no.17-1, Report CII, Austin, Texas.
- Colledge, B. (2000) "Obligations of Good Faith in Partnering of UK Construction Contracts.", "International Construction Law Review", 17:1, 175-201.
- Construction Industry Development Board (2007) "Construction Industry Master Plan Malaysia 2006-2015." Kuala Lumpur: Construction Industry Development Board, (ISBN 978-983-2724-44-5).

³⁰ Walker, I and Wilkie, R (2002) "Commercial management in Construction." UK: Blackwell Publishing.

- Cornick, T and Marther, J (1999) "Construction project teams: making them work profitably." London: Thomas Telford.
- Eriksson, E and Nilsson, T (2008) "Partnering the construction of a Swedish pharmaceutical plant: case study". *Journal of Management in Engineering*.
- Jamieson B.K (2001) " Partnering-Some legal issues." *Contrax Asia Ltd*.
- Kadefors A. (2004) "Trust in project relationships-inside the black box". *International Journal Project Management* 22 (3), 175–182.
- Latham Sir M (1994) "Constructing the team: Final report of the government industry review of procurement and contractual arrangements in UK industry". London: HMSO.
- Murphy, Michael (1995). "A contractor's Partnership Baptism". In Godfrey, Kneeland A. (1995). *Partnering in Design and Construction*. New York: McGraw-Hill.
- Pradhan, V., (2002) "Partnering in construction contract" Seminar on 'Partnering & Best Practice in Construction & The Work of Sir Michael Latham'. James R Knowles.
- Scott B. (2001) "Partnering in Europe: incentive based alliancing for projects". London: Thomas Telford.
- Walker, I and Wilkie, R (2002) "Commercial management in Construction." UK: Blackwell Publishing.
- Wallace, I.N.D., (1999) "The forgotten client : A history of construction contract", in twelfth annual construction centre, centre of construction law and management, King's College London, UK. Pt. VIII. : 1-53.

Dispute Resolution: A Preliminary Report Regarding Recent Changes to the AIA Documents

J. Mark Taylor¹

¹McWhorter School of Building Science
Auburn University, Auburn, Alabama
United States of America

Email: taylojt@auburn.edu

Abstract:

The commercial construction industry has steadily increased in complexity over the last few decades. Construction contracts have become extremely complex in their language, often making interpretation very difficult. While construction contracts often contain anticipatory language for unanticipated occurrences, not all eventualities can be prepared for. Consequently, disputes are inevitable in the construction process. The construction industry thrives on building lasting relationships, so it is absolutely critical that disputes are handled appropriately and expeditiously. If disputes are not dealt with expedience, the project itself can be delayed, leading to claims, which could ultimately destroy professional relationships. Dispute resolution, if handled properly, can have a major impact on the success of a project.

Dispute resolution mechanisms should be decided upon at the time the parties enter into the contract. These are important decisions for construction participants to make. Fortunately, there are a number of options the parties can utilize in dispute resolution including: litigation, arbitration, negotiation, mediation, and dispute review boards. Recent changes in the 2007 American Institute of Architects (AIA) construction contract documents reveal a potential changing attitude towards dispute resolution. The purpose of this study will be to provide a preliminary look into which dispute resolution methods are most effective for commercial contractors.

In addition, as a means for providing background, this study will look at the various dispute resolution practices used in the commercial construction industry and attempt to determine which methods offer the best solution. Data will be obtained from attorneys, construction executives (general contractors and subcontractors), and project managers in order to gain the broadest perspective on the issue. By doing so, the study will be able to present excellent supportive data for the most desired resolution outcome.

Keywords:

Arbitration, Contracts, Disputes, Dispute Resolution, Mediation

1. Introduction

The commercial construction industry has steadily increased in complexity over the last few decades (Cheung, 1999). Construction contracts have become extremely complex in their language, often making interpretation very difficult. While it is true that construction contracts often contain anticipatory language for unforeseen occurrences, not all eventualities can be prepared for. Consequently, disputes are inevitable in the construction process (Cheung, 1999). The construction industry thrives on building lasting relationships, so it is absolutely critical that disputes are handled appropriately and expediently. Dispute resolution, if handled properly, can have a major impact on the success of a project, as each party involved is able to stay focused on the actual construction of the project itself (Essex, 1996).

Dispute resolution mechanisms are an important decision for construction participants to make. Dispute resolution proceedings can be costly and exhaust a tremendous amount of resources (Cheung et al., 2002). Subsequently, the construction industry has begun to look for the best possible solution in settling disputes. The uniqueness of every construction project demands that every aspect of resolving potential disputes be examined thoroughly to determine which method is the most applicable. Fortunately, dispute resolution as a whole offers many different options to choose from including: litigation, arbitration, negotiation, mediation, and dispute review boards (Essex, 1996).

All parties involved with commercial construction contracts should make an effort to use prevention techniques prior to construction in order to minimize the impact of disputes that may arise. Dispute prevention techniques include an understanding of pre-construction risk mitigation by all parties including: contractors, owners, designers and subcontractors (Cheung, 1999). However, despite all of the pre-construction efforts, it is likely that disputes will arise during the construction process. Once a dispute arises, the construction industry has vacillated over the most effective means of resolving the dispute. The private commercial construction sector has most recently favored mediation and arbitration as means for resolving the dispute. This is reflected in the most popular construction contract document forms; mainly those promoted by the American Institute of Architects (AIA). However, in the latest version of the AIA documents (2007) a significant change has occurred whereby mediation and arbitration is no longer the default mechanism for resolving disputes. Our research is aimed at looking at how the industry views dispute resolution mechanisms and if, indeed, there are changing attitudes in this arena.

2. Background / Literature Review

Negotiation is almost always the first step in dispute resolution. However, when negotiation fails, other methods are quickly imposed. The use of litigation has historically been the favored option for resolving ones dispute and is often looked upon as the traditional method. However, as a consequence of the enormous cost and resource drain that litigation requires, the United States construction industry began looking to alternative measures to deal with disputes. Alternative dispute resolution, or ADR, has been proven to be a highly effective alternative to litigation since the mid-1970's even though it was not widely adopted in the construction industry until the mid-

1980's (Weidner, 2006). ADR refers to any means of dispute resolution that takes place outside of a court room (ADR, 2008). ADR is often thought of as a means of resolving disputes more efficiently and also can assist the parties in understanding the issues between them more clearly, minimizing the adversarial atmosphere that can be created in the midst of a dispute (Essex, 1996).

The application of dispute resolution practices in the commercial construction industry will vary depending on the project in question. Past research indicates that the resolution of potential conflicts between contractors, owners, designers, and subcontractors has several conclusions. Primarily, the dispute resolution discussions focus on: (a) when to resolve disputes (Cheung, 1999; Groton, 1991; Cheung et al., 2002), (b) what type of dispute resolution is appropriate (Groton, 1991; Widiss, 1979; Moffitt & Bordone, 2005; Tarlow, 2008), (c) what has worked and not worked in the past (Korn & Pallas, 2007; Hinchey & Schor, 2002; Keil, 1999; Widiss, 1979), and (d) what forms dispute resolution may take in the future (Essex, 1996; Groton, 1991; Kane, 1992).

Commercial contractors should always make an effort to use prevention techniques prior to construction, in order to minimize the impact of disputes that arise. Dispute prevention techniques include an understanding of pre-construction risk mitigation by all parties including: contractors, owners, designers, and subcontractors (Cheung, 1999). In other words, all parties who have a part in the project should meet before breaking ground, and make sure that everybody understands what is expected of them. However, despite all of the pre-construction efforts, it is likely that disputes will arise during the construction process. The question then becomes, when and how do we resolve these disputes?

The American Institute of Architects' (AIA) construction contract documents are widely used throughout commercial construction. Dispute resolution has been a concern expressed in the past from owners, architects, and contractors (Gibbs, 2007). Up until the latest version of the AIA documents were issued in 2007, the architect was identified as the initial arbiter of claims. This posed a number of problems and potential areas of conflict. Owners don't want their architects making decisions against them and contractors often believe that architects cannot act impartially when they are being paid by the owner. In addition, architects are reluctant to make decisions that may affect their own liability. This often resulted in the Architect being placed in a "Catch 22" scenario; i.e., simply being caught in the middle. As a result, the latest revisions to the AIA documents have allowed for several changes to their contract documents as related to dispute resolution.

The AIA101-2007 and AIA201-2007 (Owner-General Contractor Agreement and General Conditions to the Contract, respectively) family of documents advocate the concept of a third party neutral as the Initial Decision Maker (IDM). The third party neutral is agreed upon by the parties at the time of contracting and if a third party neutral is not named, the architect takes on the role of IDM. Another important revision in the 2007 documents, as related to dispute resolution, is the freedom to choose between arbitration and litigation. This is often referred to as the "check box" system; and if neither option is selected the default option becomes litigation. As a reflection of

the success of mediation in assisting the parties in resolving their disputes, the 2007 documents maintain mediation as the first step in dispute resolution, as did the 1997 documents (Gibbs, 2007).

Some industry professionals believe that there is a tendency to resolve monetary disputes after construction is complete (Groton, 1991). This idea stems from the belief that the cost would be greater to everybody in the long term, if project resources were reallocated to dispute resolution during the construction process. The problem with this belief is that uncertainty about the outcome of the project is increased, creating an adversarial relationship. However, the truth is that the vast majority of industry professionals advocate prompt resolution to construction disputes as soon as they arise (Cheung et al., 2002). These construction leaders recognize the value added in dealing with disputes expeditiously. Research suggests that early dispute resolution practices will help contractors, owners, designers, and subcontractors maintain and preserve their business relationships.

Mediation remains a popular method of dispute resolution. However, in order for mediation to be effective all parties must be willing to listen and try to work things out amicably. Many of the contracts today stipulate that parties must try mediation prior to arbitration or litigation (Tarlow, 2008). The problem with this is that some disputes may already be to the point where a binding resolution may be the only way to resolve the dispute. Alternatively, mediation has proven to be an excellent dispute resolution option if all parties have the right attitude regarding the desirability of resolving the dispute (Tarlow, 2008). Unfortunately, since mediation is non-binding, it may fail to produce a resolution and parties may have to look to binding resolutions, such as arbitration and litigation. Historically, contractors have looked at arbitration as a favored mechanism under the guise that it is an expeditious, efficient, and economical solution. One distinct difference between arbitration and litigation that is critically important is that arbitration is agreement based and not always provided with standards, as is litigation (Widiss, 1979).

The true benefit of litigation is that legal principles and the right to appeal are what uphold the standards, making litigation an attractive option since previous court rulings provide a solid foundation for future proceedings (Widiss, 1979). Commercial contractors must look for the dispute resolution practice that is most likely to resolve the conflict as economically as possible. Contractors can start the selection process by asking three simple questions: what are the goals of the processes, what aspects of the dispute in question make it resolvable using one practice over another, and what contributions might each practice make to overcoming problems that would prevent an effective resolution from coming to fruition (Moffitt & Bordone, 2005).

Architects and contractors are both under increased pressure today from owners to design and build projects quickly. This pressure only increases the likelihood that disputes will arise. Mediation often proves to be a successful mechanism for resolving disputes quickly but is dependent upon having a mediator that is proactive. A mediator must ask the tough questions so that there is no doubt as to where the proceeding is heading. If a mediator fails to be proactive, then the probability of a successful outcome is diminished (Korn & Pallas, 2007). Mediation, however, may not be the best method

to use when parties are trying to resolve a case that deals with a distinct issue of law. Hinchey and Schor contend that litigation establishes a legacy from which attorneys may advise their clients. Furthermore, contractors who build large projects with enormous amounts of money at risk, usually spend a tremendous amount of time in pre-construction negotiation in order to mitigate their risks. These contractors may not be willing to give up their interpretation of these negotiations to a mediator or arbitrator, who is not legally bound to adhere to the terms of the agreement (Hinchey & Schor, 2002).

Whichever dispute resolution method is selected should aim to resolve the dispute as efficiently and effectively as possible. All professionals involved in a construction project typically only have two resources that they can contribute to a construction process: time and talent. Consequently, if the wrong method for dispute resolution is chosen, these two resources may be siphoned away from the project which would open the door for project failure, and the development of an adversarial environment (Keil, 1999).

As the commercial construction industry continues to advance in technology and complexity, the need for advanced dispute resolution mechanisms will increase. One emerging concept finding favor in some segments of the industry is dispute review boards. These boards allow a pre-selected panel, normally consisting of three independent parties, to advise disputing parties on a resolution solution specifically tailored to their dispute (Essex, 1996). The review board is often engaged at the outset of the contract and makes a commitment to meet as necessary during the construction process to handle any disputes in a timely fashion. Some commercial contractors are finding this method to be extremely effective. It is designed to keep the panel abreast of all ongoing construction proceedings so that they may offer immediate advice should a dispute arise. This method may also be helpful, particularly for contractors and subcontractors, since field occurrences are difficult and costly to recreate. Research has shown that dispute review boards also help to minimize actual disputes from arising in the first place. This finding is directly attributable to the real-time project knowledge that the panel members possess (Kane, 1992). As project complexity increases dispute resolution may also move towards hybrid forms in the future.

3. Methodology

The purpose of this study is to look at various dispute resolution practices used in the commercial construction industry and to ascertain which methods offer the best solution. Data was obtained from owners, attorneys, construction executives, and senior project managers in order to gain the broadest perspective on the issue. By focusing on the aforementioned parties for direct information, the study was able to present supportive data to determine which methods offer the best solution for the commercial construction industry.

Data was gathered from industry professionals using a qualitative research approach. Qualitative research can be defined as subjective, and is often structured into two categories of research: exploratory and attitudinal (Coles & Naoum, 1998). The purpose of exploratory research is to understand a situation, look for alternatives, and to

propose new ideas (Zikmund, 1997). Exploratory research was used as construction professionals were interviewed regarding their opinions on dispute resolution practices. Attitudinal research was used to evaluate the opinion, view, or perception of an individual, towards a certain object (Coles & Naoum, 1998). The interviews provided substantive information as to the current views associated with dispute resolution in order to provide the reader with contemporary information regarding practices that are best suited for resolving commercial construction disputes.

4. Data Analysis and Results

The data compiled was analyzed using the qualitative approach of triangulation. The data obtained from the interviews was interpreted using a three step process. First, the data was analyzed separately according to professional position to ensure that all data has been properly catalogued. Second, the data was compared using the triangulation method, and extensions thereof, to look for commonalities between the various professionals interviewed. The triangulation method is used to integrate data from multiple sources (Univ. of California; Triangulation, 2008). The goal of triangulation is to find recurring themes that are prevalent throughout the interviews. Extensions of the triangulation method were also used to further develop a rich understanding of dispute resolution. For example, the theory of complementarity was used with the construction executives and attorneys as they were both asked questions that are unique, in the sense that they are overlapping as well as different (Gaber & Gaber, 2007). The goal of using the complementarity approach is to understand the difference in thinking between business owners and the attorneys who represent their interests. In addition to using the complementarity approach, the researcher used an additional method of triangulation with the senior project managers known as expansion (Gaber & Gaber, 2007). The expansion method was used to extend the range of conceptual understanding by asking questions that target different components of the same framework of questioning. Although the questions are different in their own right, the underlying concepts are similar to those asked the executives and the attorneys. The expansion process allows the researcher to gain a deeper understanding of the affects that the dispute resolution process has regarding various construction professionals.

These methods are of particular importance to this data analysis, since the individuals interviewed were asked a variety of questions. The variety of questioning will ensure that potential biases are reduced, by allowing industry leaders to answer the questions in a manner that corresponds to their professional capacity. Consequently, the triangulation method will serve as a filter, allowing the professionals to provide the reader with a strong interpretation of dispute resolution opinion, relative to their role in commercial construction.

Initially data was coded a priori using themes and keywords derived from the interview questions (Gibbs & Taylor, 2006). The a priori codes were used to label the questions asked. In addition, for analysis purposes, the questions were categorized by which participant was asked the question. The ultimate goal of this research is to provide construction professionals with an informed outlook of dispute resolution methods.

Twenty seven questions were used to explore the attitudes of the parties involved with the process. Initially a simple question with an expected outcome was asked and all of the participants agreed that negotiation was the most non-adversarial method of resolving a dispute. All participants agreed that dispute resolution methods made a significant impact on operations with the exception of a single subcontractor.

Next, the parties were asked about the root cause of most disputes. Money and communication breakdowns were factors that result in disputes entering an impasse from which the parties would seek outside assistance from their attorneys.

The next series of questions were aimed at the contracting parties and their interaction with attorneys. General contractors were much more willing to seek attorney involvement than were subcontractors.

Questions were then asked about the importance of having dispute resolution methods contractually defined prior to construction. In a related question, the parties were also asked about the differences in the 1997 and 2007 versions of the AIA documents. The contractors interviewed had not yet been personally involved with the 2007 documents. However, the majority of parties agreed that dispute resolution methods should be defined in the contract and that the Architect should not be the Initial Decision Maker.

A series of questions were asked to explore the attitudes of the respondents to the services of the American Arbitration Association (AAA) or other arbitration facilitators. Also explored in this line of questioning is whether there is a preference for a single arbitrator or a panel of arbitrators. To further explore arbitration, participants were questioned about the lack of appeal in arbitration. The majority of the respondents favored arbitration over litigation and had a preference for a single arbitrator, unless the particular situation dictated a panel. No preference, one way or the other (i.e., positive or negative) was evident as it related to the AAA or any other arbitration facilitator. A number of the respondents stated that their experience was generally a joint administration of the arbitration proceedings by the parties; i.e., using no formal entity or association as a facilitator. General contractors and subcontractors had mixed responses about arbitration's lack of appeal while attorneys stated that there should be no concern since the parties interested in a right to appeal should favor litigation in the first place.

Questions were asked to determine if the participants felt that mediation was an effective form of dispute resolution. There was some concern among all parties interviewed that since mediation is not binding it is often not taken as seriously by all parties and some may try to play the system. However, a great majority of the respondents stated that mediation was effective in resolving disputes. Attorneys stated that the success of mediation is, to a large degree, dependent upon the skill of the mediator. A mediator who takes charge of the proceedings and acts in a forceful manner is much more likely to end with a settlement than one that is more passive. In addition, the attorneys offered that parties often appear to have a preconception regarding the likelihood of a settlement being reached prior to the mediation. Parties that go in with a positive attitude that the dispute will be resolved during the mediation are much more likely to succeed than those going in without a positive approach.

Furthermore, some of the respondents indicated that mediation was treated simply as a precursor to arbitration or litigation.

The attorneys unanimously agreed that litigation offered more legal challenges than alternative methods. They stated that litigation could be very complex and time consuming, and that judges and juries often had limited construction knowledge. Attorneys also reported that most executives were knowledgeable about dispute resolution methods and the distinctions between the alternatives; but that it really depended upon the specifics of the project and the dispute as to which method was most appropriate. All of the attorneys interviewed felt that dispute resolution methods were making positive improvements, but they went on to say that mediation, arbitration, and litigation were here to stay. Attorneys pointed out that advances were being made in dispute review boards. They opined that while many in the industry believe that dispute review boards were appropriate only for federal projects they are seen more frequently in private commercial construction than just a few years ago. The attorneys also stated that some are experimenting with hybrid forms of dispute resolution such as a mediation-arbitration hybrid.

Attorneys were also questioned about the changes in the most recent version of the AIA documents (2007). All of the attorneys agreed that the change in the consolidation clause was important. The 1997 AIA forms require disputes involving the architect to be held at a hearing separate and apart from other parties; i.e., that claims against architects cannot be consolidated with claims against other parties. To illustrate, an owner has a potential dispute that involves both the architect and the general contractor. The owner would be required to have two hearings in order to resolve this matter under AIA 1997. Many owners and contractors were frustrated with this clause and were threatening to discontinue use the AIA contract documents altogether if it was not modified. The change is reflected in the 2007 AIA documents in that their no longer is the “no consolidation” clause.

Most of the attorneys stated that it was too early to adequately assess the changes in the dispute resolution clauses from the 1997 to 2007 AIA documents. That we are still at least a couple of years away from being able to see if the “check box” system and the default provision for litigation will have any effect on how the parties handle disputes.

Project managers were also questioned and all reported that when a dispute arose on one of their projects they were intimately involved in dispute resolution process. They also reported that there was a heightened degree of tension on a project during the course of the dispute resolution. The project managers also reported that they had worked with IDMs other than an architect and these always turned out to be independent construction managers. They also reported that they were personally involved with the data collection process during disputes and this required a significant amount of their time.

5. Conclusion

Not surprisingly, the research results indicate that all parties prefer to settle disputes by negotiating. All agreed that maintaining relationships and keeping dispute resolution costs to a minimum are very important factors in how they conduct business and that

negotiating a settlement is the best way to accomplish this goal. Mediation is seen as an effective form of dispute resolution by all of the general contractors and attorneys. However, subcontractors offered different opinions on the matter with some preferring mediation while others found it futile since it is non-binding.

All parties essentially agreed that litigation is the worst form of dispute resolution mechanism. However, an interesting part of this research is the attorneys' attitude towards dispute resolution. Attorneys have similar opinions to the construction executives, in that they believe that the quicker disputes can be resolved, the better for everyone; but they take a legalistic approach towards the subject, essentially stating that any dispute resolution is fine as long as the parties have the freedom to choose during the time the contract is negotiated.

There may not be one definitive dispute resolution method that is best suited for all situations, but the research clearly indicates that negotiating disputes and dealing with them quickly is the best course of action. More research will be available in the coming years once the latest versions of the AIA documents become more widely used. It will be interesting to see if the new AIA contractual language (e.g., check box system) leads to a decrease in arbitration and an increase in litigation. The attorneys also agree that advancements are being made in dispute resolution methods with the dispute resolution boards and some attempts at hybrid systems. However, they also point out that mediation, negotiation, and litigation practices were here to stay. The research clearly indicates that disputes will always be prevalent in commercial construction, and they must be resolved as quickly as possible to preserve professional relationships and the perpetuity of the commercial construction business.

References

- Cheung, S.O. (1999), 'Critical Factors Affecting the Use of Alternative Dispute Resolution Processes in Construction', *International Journal of Project Management*, **Volume 17** (3), pp 189-194.
- Cheung, S.A., Suen, H.C.H., & Lam, T-I. (2002), 'Fundamentals of Alternative Dispute Resolution Processes in Construction', *Journal of Construction Engineering & Management*, **Volume 128** (5), pp 409-417.
- Coles, D.H., & Naoum, S.G. (1998), *Dissertation: Research and Writing for Construction Students*, Oxford: Butterworth-Heinemann.
- Cornell University Law School (n.d.), 'Alternative Dispute Resolution (ADR): an Overview', *Topics.law.cornell*, <http://topics.law.cornell.edu/wex/ADR>, viewed: October 15, 2008.
- Essex, R. (1996), 'Means of Avoiding and Resolving Disputes During Construction', *Tunneling and Underground Space Technology*, **Volume 11** (1), pp 27-31.
- Gaber, J., & Gaber, S. (2007), *Qualitative Analysis for Planning & Policy*, Chicago: American Planning Association.
- Gibbs, G.R., & Taylor, C. (2006), 'How and What to Code', *Onlineqda*, http://onlineqda.hud.ac.uk/Intro_QDA/how_what_to_code.php, viewed May 3, 2009.
- Gibbs, K.C. (2007), 'Dispute-Resolution Changes Coming', *ENR: Engineering News Record*, **Volume 259** (13), p 67.

- Groton, J.P. (1991), 'Alternative Dispute Resolution in the Construction Industry', *Dispute Resolution Journal*, **Volume 52** (3), pp 48-58.
- Hinchey, J.W., & Schor, L. (2002), 'The Quest for the Right Questions in the Construction Industry', *Dispute Resolution Journal*, **Volume 57** (3), pp 8-20.
- Kane, C. (1992), 'Mitigating Construction Contract Disputes (power plant construction)', *Public Utilities Fortnightly*, **Volume 130** (1), pp 11-13.
- Keil, J. H. (1999), "'Hybrid ADR" in the Construction Industry (alternative dispute resolution)', *Dispute Resolution Journal*, **Volume 54** (3), pp 14-22.
- Korn, R., & Pallas, G. (2007), 'Managing Dispute Resolution Options in the Construction Industry', *Dispute Resolution Journal*, **Volume 62** (2), pp 22-34.
- Moffitt, M.L., & Bordone, R.C. (2005), *The Handbook of Dispute Resolution*, Jossey-Bass, San Francisco.
- Tarlow, A.L. (2008), 'Mediation of Construction Disputes', Ramco Insurance, <http://www.ramco-ins.com/resource-library/risk-management/08-atarlow.htm>, viewed: October 20, 2008.
- University of California, San Francisco, 'Global Health Sciences: Prevention and Public Health Group – Triangulation', *Institute for Global Health*, <http://www.igh.org/triangulation/>, viewed November 5, 2008.
- Weidner, L.E. (2006), 'Recent Development: Model Standards of Conduct for Mediators', *Ohio State Journal on Dispute Resolution*, **Volume 21** (547), pp 15-32.
- Widiss, A.I. (1979), *Arbitration: Commercial Disputes, Insurance, and Tort Claims*, Practising Law Institute, New York.
- Zikmund, W. (1997), *Business Research Methods*, Dryden Press, London.

Appendix A

A. Interview Questions Posed to Construction Executives

1. What types of dispute resolution methods do you feel are most non-adversarial, and are best suited for relationship preservation?
2. When choosing a dispute resolution option, what factors are you most concerned with in terms of the impact on operations? (Do you decide or do you rely on advice from your attorney; do you use different methods for different clients; do you use different methods for different types of work)?
3. Is there one method that stands out among construction executives (or for your company) as the best all around?
4. When do you consider a dispute at an impasse and solicit outside assistance?
5. What individual is responsible for dealing with the attorneys?
6. How important is it that dispute resolution options be contractually defined prior to construction?
7. Do you like the change from 1997 to 2007 AIA Docs re dispute resolution? (i.e., Initial Decision Maker not having to be the Architect, and the new “check box” option in the 2007 version to allow for arbitration or litigation, should mediation attempts fail.)
8. If you do like the check box option in the 2007 version, why?
9. How would you choose the Initial Decision Maker? What is important to you when making this decision?
10. Have you used AAA’s arbitration services before, or any other facilitator?
11. If no to question 10, have you been involved in deciding who will be the arbitrator with the other party involved? In other words, were both sides able to work together in picking their arbitrator?
12. In arbitration, do you favor a single arbitrator or a panel of arbitrators (i.e., 3)?
13. Is the lack of appeal in arbitration a concern?
14. Do your contracts require mediation prior to arbitration (according to the 1997 AIA version)?
15. Since mediation is non-binding, do you feel that it is effective? Do you feel that the environment offered by mediation is non-adversarial and conducive to an amicable solution?

B. Interview Questions Posed to Attorneys

1. How important is it that dispute resolution options be contractually defined prior to construction?
2. What options present the most legal challenges for an adequate resolution?
3. How knowledgeable are construction executives about dispute resolution as a whole?
4. What does the future of the construction industry look like in terms of dispute resolution advancement?
5. Do you like the change from 1997 to 2007 AIA Docs re dispute resolution? (i.e., Initial Decision Maker not having to be the Architect, and the new “check box” option in the 2007 version to allow for arbitration or litigation, should mediation attempts fail).
6. What are your thoughts in regards to the Initial Decision Maker not having to be the Architect?
7. Why did the architects agree to being “consolidated” with other parties in construction related disputes in the 2007 version of the AIA documents?
8. If you do like the check box option in the 2007 version, why?
9. In arbitration, do you favor a single arbitrator or a panel of arbitrators (i.e., 3)?
10. Is the lack of appeal in arbitration a concern?
11. Do you feel that contractual consistency is important between the different parties in construction? (i.e., between Owners and GC’s and GC’s and Subcontractors).
12. Since mediation is non-binding, do you feel that the odds of a successful resolution are worth the effort, or do you believe that arbitration or litigation offer a better solution all around for construction related disputes?

C. Interview Questions Posed to Senior Project Managers

1. How involved are you with dispute resolution?
2. Where do you see the biggest impact of a selected dispute resolution option in terms of worker productivity? Is there a negative impact on the jobsite as a whole? (i.e., attitude, schedule concerns, etc.). Do superintendents express the significance of these issues to you?
3. Can you sense a heightened degree of tension on a project if a particular dispute resolution option is in use? Can you elaborate on any experiences that you may have had?

4. What are your thoughts in regards to the Initial Decision Maker not having to be the Architect?
5. Have you had to deal with an IDM other than the Architect?
6. What type of information are you responsible for collecting when disputes arise?
7. How does the data gathering process impact your productivity, in terms of your day to day professional responsibilities?
8. Do you find it difficult to get other parties to be helpful in data gathering once a dispute arises? Do you sense an attitude shift in the professional relationships with subs, owner, etc. once a dispute is ongoing?
9. Have you ever personally been involved in an actual mediation, arbitration, or litigation proceeding? (i.e., had to testify or participate directly).
10. If yes to number nine, what resolution method have you found to be the most productive and palatable for all parties involved?

A conceptual model for effective maintenance of productivity records: a tool for facilitating the analysis of disruption claims

Nuhu Braimah¹

¹ Civil Engineering, School of Engineering and Design Environment,
Brunel University, Uxbridge, UB8 3PH,
United Kingdom

Email: Nuhu.Braimah@brunel.ac.uk

Abstract

Productivity of construction resources, especially labour, is a fragile variable readily affected by myriad of factors. This vulnerability, combined with difficulties associated with measuring, tracking and controlling site productivity has made the resolution of associated lost productivity (or disruption) claims an extremely challenging task for contracting parties to grapple with. The difficulties involved are exacerbated by lack of accurate and adequate productivity records needed for an amicable resolution. As a contribution toward addressing this problem, the author is currently embarking on a research aimed at developing a scientific system/tool for effective monitoring and maintenance of productivity records. This paper reports on the work done so far. With the aim of highlighting the nature of the problem, the paper first gives an overview of the available methodologies for resolving disruption claims and their use in practice based on the author's previous survey. The survey results indicate that although simpler methodologies such as the Global Method and Modified Global Method are ineffective due to their numerous weaknesses, they are more widely used in practice than the more robust ones such as the Measured Mile and Earned Value Techniques, which require detailed site-productivity data for implementation. On the premise that lack of adequate productivity records is responsible for this practice, a preliminary conceptual model is proposed as an initial step towards developing the scientific system/tool, that will hopefully facilitate the analysis of disruption claims. The model has a number features that enhance its ability to promote this function than what conventional cost control systems offer. The paper finally identifies areas of further research required for refining the model in order to make it more applicable in practice.

Keywords:

Delay and disruption, construction claims, loss of efficiency, labour productivity.

1 Introduction

Effective management of resources (labour, equipment, etc) is a basic requirement for timely and on-budget completion of construction projects. The productivity of these resources is, however, highly influenced by myriad of factors, such as, mismanagement, inclement weather, change orders, differing site conditions, to mention a few (Hanna and Heale, 1994; Makulsawatudom and Emsley, 2001). The effect of these factors is often to disrupt the planned means, methods or schedule for executing the project, resulting in contractors expending more hours for a given work item than originally planned for (Schwartzkopf et al., 2001). Moreover, the ability to capture and account for these factors in bids or subsequently control them during construction is often not possible. This has consequently made cost overruns and delays almost inevitable for even the most carefully managed projects (Ibbs and Liu, 2005).

Loss of productivity claims (commonly termed “disruption claims”) tend to arise when productivity is impacted by factors for which the contractor is not responsible and is entitled to additional compensation (Klanac and Nelson, 2004). Although most forms of contract have provisions for dealing amicably with such claims during the course of construction, they often develop into disputes, commonly resolvable by courts and other expensive forms of dispute resolution (Pickavance, 2005). The issue in dispute often concerns the cause-effect relationship between the events alleged to have impacted on productivity and the amount of compensation being claimed as a result. Establishing this link with absolute certainty as required for meeting the burden of proof, is an extremely difficult task, partly because lost productivity rarely results from a single event. More often than usual, multiple events caused by the different project parties, tend to impact on productivity simultaneously (Leonard *et al.*, 1988; Hanna and Heale, 1994).

The issue of disruption claims resolutions has thus exercised the attention of many researchers and the studies so far has largely focused on two main areas. The first category of studies has concentrated on the development of statistical charts/models as tools for assessing the impacts of certain delay/disruption events on productivity (for example, Ibbs (1997); Thomas et al. (1999); Hanna et al. (1999a, 1999b); Ibbs and Liu (2005) and Moselhi et al. (2005)). The second category has focused on developing improvements to the *Measured Mile Technique*, which is the widely accepted methodology for resolving disruption or lost productivity claims— see for instance, Zink (1990), Presnell (2003), Gulezian and Frederick (2003), Ibbs and Liu (2005) and Merritt (2008). This method compares actual productivity of impacted operation with the actual productivity achieved while doing the same work but in an un-impacted mode (Zink, 1990). Such comparison has the key advantage of factoring out any issues on the validity of costs estimates and programme and contractor’s inefficient working. Nonetheless, the Measured Mile Technique is only as good as the available productivity records used for analysing the claims (Merritt, 2008). Sadly, there is often lack of such records in practice as commonly decried in the literature (see for example, Scott and Assad, 1999; Pickavance, 2005). Despite this being a well known problem, there has been very little effort made towards addressing it. As a contribution in this direction, the writer has embarked upon a research study aimed at developing a scientific system or tool for proper monitoring and maintenance of productivity records to facilitate the

successful resolution of disruption claims. This paper reports on the initial work done so far. It first presents a brief introduction on the existing Disruption Analysis Methodologies (DSAMs). With a view to provide an understanding of the nature of the problem, Section 3 presents a survey results on the extent of use of these methodologies based on the writer's previous study. In Section 4, a preliminary conceptual model of the scientific system/tool aimed at is proposed and described, based on the author's experience and recommended best practice from literature. Section 5 highlights the issues to be investigated further in research and the last section provides a summary and conclusion of the paper.

2 Disruption Analysis Methodologies (DSAMs)

Various methodologies are available for analysing disruption claims. Table 1 provides a list of the most common ones with their brief descriptions, as described in the literature (see for example, Schwartzkopf et al., 2001; Klanac, and Nelson, 2004; Pickavance, 2005). Their detailed descriptions including their advantages and disadvantages are not within the ambit of this paper and can be obtained from the cited literature.

Table 1 List of common DSAMs

Methodology	Brief Description
Measured Mile Technique	Compares actual productivity of impacted operation with actual productivity achieved while doing identical or substantially similar work but in an un-impacted mode.
Industry Standards & Charts	Uses model productivity curves and factors developed by a number of industry organisations to estimate the loss of productivity.
Global (or Total cost) Method	The claimant entitlement is estimated as the difference between the actual cost and the contractual cost without establishing any causal link between the reason for their entitlement and the quantity of the corresponding loss.
Modified Global Method	Similar to the total cost except that with this method, the contractor's bid estimate is adjusted to account for activities that were underbid or deemed to be his responsibility.
System Dynamics Modelling	In this, a dynamic computer simulation model is use to replicate on day by day basis, the impact of streams of delays and disruptions events on project progress. By this, the time and cost consequences of the various events are determined.
Earned Value Management	Cumulative curves of Budgeted Cost of Work Schedule, Budgeted Cost of Work Performed and Actual Cost of Work Performed are overlapped on the same diagram. Changes in forecasts, which are often due to reduction in productivity are identified and related to certain causative events using contemporaneous records.
Time and Motion Studies	All disruptive conditions experienced on the job are simulated; preferably at the actual job site if possible to demonstrate lost productivity attributable to the actual disruption.

All the methodologies have a common purpose of ultimately determining the amount (*Quantum*) of Productivity Lost (PL) as caused by productivity impact events. The main

sources of information for working this out is the Actual Productivity (AP) records and that of original or Planned Productivity (PP). How the former deviates from the latter gives an indication of the amount of PL experienced, i.e., $PL = PP - AP$.

The analysis process involved is not as simple as the above equation makes it look like. Coupled with lack of productivity records needed for complete and robust analysis as highlighted previously, the standard of proof set by the courts, arbitration panel, etc, for claimants to satisfy is also burdensome (Klanac and Nelson, 2004).

PP is usually obtained from contractors' original bid estimate or is represented by the productivity achieved for an unimpacted period of the project, whilst AP is the actual productivity experienced in executing the work. Factors that impact negatively on productivity but were known at the time of bidding are often not compensable, according to most forms of contract. The various DSAMs differ in their approaches, mainly on different information sources each relies on for the data on PP and AP . As Figure 1 below shows, three main sources are often used:

- project-specific data: records from the actual project in dispute;
- similar-project data: records from project with good resemblance to the project in dispute; and
- industry-wide project data: records estimated or determined analytically from models/ charts developed from various productivity studies.

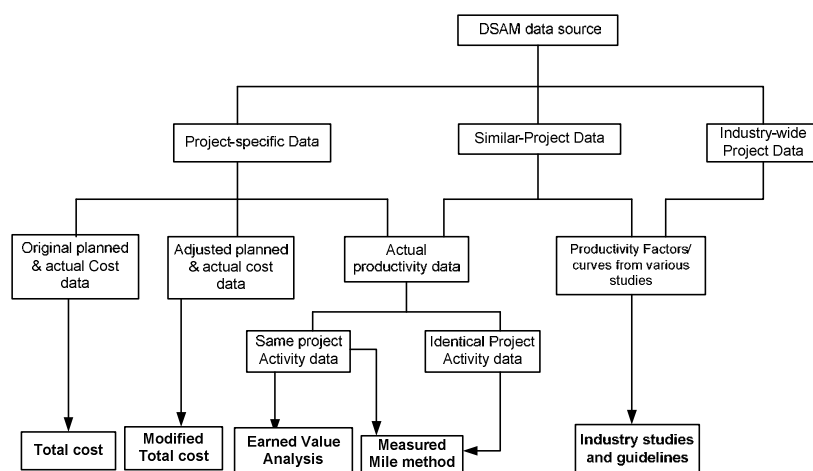


Figure 1 Sources of data used by various DSAMs

(Source: Braimah, 2008)

By virtue of these different information sources, the methodologies tend to produce results of different levels of accuracy. In fact, no single methodology is generally applicable for use in all cases and suitability of use depends on the amount and type of productivity record available (Schwartzkopf et al., 2001). Methodologies utilising Project-specific data are often regarded as being the most reliable, followed by those

utilising Similar-project data, with Industry-wide project data being the least reliable information to use. (Schwartzkopf et al., 2001, Wickwire et al., 2003, Roy, 1998).

3 An investigation into DSAMs in use

As part of the writer's previous research, a survey was carried out in the UK on practitioners views with regard to the usage of DSAMs in practice. The survey employed a semi structured questionnaire and required respondents to score on a 5-point Likert scale, on among others, their level of awareness and extent of use of these methodologies. Six hundred construction organisations were carefully sampled, using quota and purposive sampling techniques, from the Kompass Register (Kompass, 2006), NCE Consultants' file (NCE, 2006), and 2002 RICS Directory (RICS, 2002). From this, only 130 responded well to the survey - consisting of 63 Contractors and 67 consulting firms.

The survey results were analysed using descriptive statistics, relative index analysis and Kendall's Concordance. The analysis was done separately for the two main target groups and then for the whole group. There was a significant degree of agreement between contractors and consultants on the issues investigated at 95% confidence level, as indicated in Tables 2 and 3.

3.1 Level of Awareness of DSAMs

Respondents ranked the Global Method as the most well known methodology, followed by the Modified Global method, and then the Measured Mile Technique (see Table 2 below). This was not surprising given that the Global method has long been the subject of many court cases and is thus more commonly commented upon in the literature. The methodologies with the least level of awareness are the Systems Dynamics, followed by Earned Value Management and then Time and Motion Studies. These are relative new approaches for analysing disruption claims (Pickavance, 2005) and probably the reason for their less popularity among practitioners.

Table 2 Level of awareness of DSAMs

Methodology	Contractors		Consultants		Overall	
	Awareness index	Rank	Awareness index	Rank	Awareness index	Rank
Global Method	76.2	1	79.0	1	80.6	1
Modified Global Method	71.3	2	69.5	2	70.4	2
Measured Mile Technique	61.6	4	66.1	3	66.2	3
Industry Studies and Guidelines	63.7	3	63.8	4	63.5	4
Time and Motion Studies	46.1	6	56.7	6	55.1	5
Earned Value Management	48.7	5	57.4	5	53.1	6
Systems Dynamics	27.0	7	38.7	7	32.9	7
Test Statistics						
Kendall's W = 0.98 and $\chi^2_{critical} (\alpha=0.05) = 12.59$; df = 6; $\chi^2_{sample} = 764.4$						

3.2 Extent of use of DSAMs

On the whole, the most widely used methodologies are Modified Global method, followed by the Global method and then Industry Studies and Guidelines. Although the first two make use of direct project records, they are very simplistic and are unable to accurately determine, with certainty, the lost productivity costs resulting from individual impact events (Klanac and Nelson, 2004). Their higher rate of use is mainly because their implementation requires relatively less amount of project records, which is much in common with inadequate record keeping on sites. The Measured Mile Technique ranked 4th despite its endorsement by US courts and UK's Society of Construction Law (Merritt, 2008). The System Dynamics was the least used methodology followed by the Time and Motions studies. Such ranking was expected as these are relatively new techniques to practitioners.

Table 3 Extent of use of DSAMs

Methodology	Contractors		Consultants		Overall	
	Usage index	Rank	Usage index	Rank	Usage index	Rank
Modified Global Technique	62.3	2	56.8	1	58.8	1
Global Method	65.3	1	54.5	2	50.9	2
Industry Studies and Guidelines	46.7	3	52.3	3	49.5	3
Measured Mile Technique	45.5	4	48.0	4	46.8	4
Earned Value Management	36.4	5	39.0	5	37.7	5
Time and Motion Studies	32.6	6	35.5	6	34.0	6
Systems Dynamics	23.9	7	28.4	7	26.1	7
Test Statistics						
Kendall's W = 0.925 and $\chi^2_{critical} (\alpha=0.05) = 12.59$; df = 6; $\chi^2_{sample} = 721.5$						

4 A conceptual model for facilitating effective assessment of productivity claims

There is evidence in the literature (see for example, Brown, 1995; Gaitskell, 2003) that UK courts do not offer support for DSAMs which, according to the survey results, are common employed in practice, particularly the Global method. As highlighted previously, the key problem here is the fact that productivity records maintained on projects are often not adequate. This has compelled practitioners to rely on simplistic and less rigorous methodologies such as the Global method, which carries much room for challenge and disputes. The implication is that disruption claims will continue to be a major source of dispute unless contracting parties are able to adopt the more acceptable approaches for resolving claims. As part of the writer's wider research aimed at addressing this issue, a preliminary conceptual model is proposed.

The model is designed as a cost control system with a number of features for facilitating the maintenance of contemporaneous productivity data. Conventional cost control systems are often not able to capture comprehensive productivity records on various work items under the cost codes used. They are also backward-looking by providing information that only relates to some past (Harris et al., 2006), with little emphasis on contemporaneous productivity records and monitoring. This limitation increases the need for claims submission long after the completion of the work when much of the records have faded into history, thus risking disputes. In addition to addressing these short comings, the proposed model (Figure 2 below) is intended to provide valuable information for: future bidding, instituting corrective actions and facilitating early determination of cause and effect analysis of any potential claims.

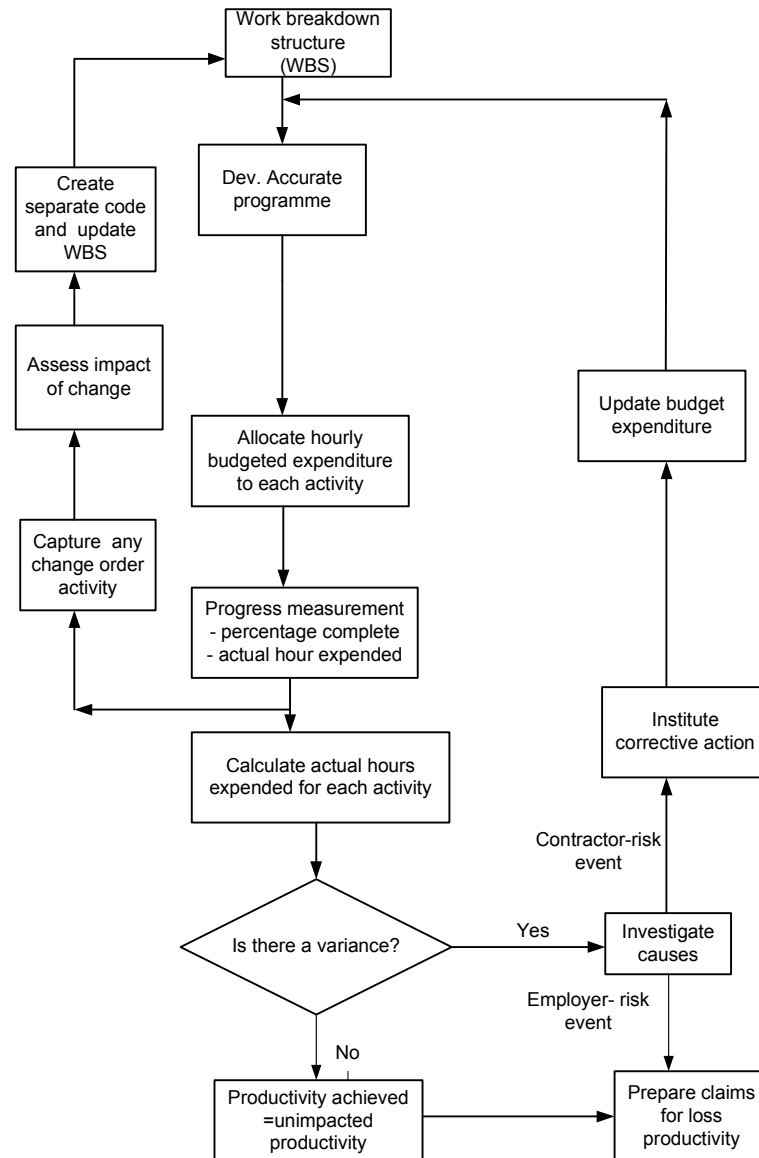


Figure 2 Preliminary process model for effective productivity claims analysis

The key components of the proposed model are explained briefly as follows.

4.1 Work breakdown Structure (WBS)

An essential first step is to develop a flexible and usable WBS. The estimates used for bidding should aid this process but the caveat is not to rely only on this since such estimates are often designed to win the job. The importance of starting with a suitable WBS includes its ability to identify any activity overlooked in the estimate during bidding (Foster, 2000). The key challenge here, though, is determining the level of detail that the WBS should be broken down into (Jung and Woo, 2004). Keeping the level too general would not provide suitable data for effective progress monitoring, whilst too much details, though provides more minute information for detailed analysis, requires more extensive effort for managing the data. For this reason, it is recommended

that the level of detail should be planned and controlled based on how efficient data can be handled (Yang *et al.*, 2007).

4.2 Develop programme and allocate hourly budget

The WBS should form the basis of the usual contract programme that has to be produced. The activities forming the programme should, as much as possible, match the breakdown of work items under the WBS (at least for the critical path activities). For each activity, the quantity of resources required (crew size, material and equipment) and the corresponding man-hours have to be determined accurately. Hours allocated should also take into account expected inefficiencies to be experienced during the first and last 10% of duration of activities, as caused by start-up and tail-off fluctuations in labour. For purposes of quantifying lost productivity these are often disregarded (Merritt, 2008).

4.3 Progress measurement and reporting

Progress should be recorded for each activity by measuring its percentage completion. Corresponding value hours are then computed by multiplying the percent completion by budgeted hours. Site manager should be responsible for this and has to report accurately the hours used for each activity. Not only does labour forms a very high percentage of the total job cost, it is also often the most difficult resource to control and keep track of. Therefore a particular feature of the model is to have a WBS and cost codes that are able to capture separately the hours and cost expended on labour, material and equipment.

4.4 Investigation of man-hour variance

At each reporting stage the value hours of each activity should be compared with its actual hours in the usual way and any shortfall investigated qualitatively in the first instance for the possibly productivity impact factors that caused it. The most common factors reported in various research (for example, Olomolaiye *et al.* (1987), Zakeri *et al.* (1996), Kaming (1997), Makulsawatudom and Emsley (2001)), include: material shortages, change or delays in supervisions, labour shortages/turnover/absenteeism, overtime, capabilities of supervisors, adverse weather, poor scheduling, rework, equipment shortage/breakdown, interference at work, learning curve, variation order/design and drawing errors, accidents, poor communication, fatigue/work overload and overcrowding/trade stacking. Whether a contractor is entitled to compensation for loss and expense resulting from the occurrence of any of these events is entirely a matter of the contract. The relevant contract documents should therefore be consulted in deciding on which of the contracting parties is responsible for the reduced productivity. When establishing this, it is also essential to the reliability of the results for the contractor to also consider and account for any mismanagement on its part or for possible errors in its bid. Such errors often relate to underestimation of labour, miscalculation of work volume and duration and wrong assumptions on method of work.

4.5 Institute corrective action and/or prepare disruption claims

Once the potential factor(s) responsible for the reduced productivity has been identified, the model recommends the contractor, in liaison with relevant project parties or otherwise, to pursue main 2 actions:

- to investigate the best possible way of mitigating and recovering the loss in productivity, if caused by contractor-risk events, i.e. events which, according to terms of the contract, the contractor bears the risk.
- to begin the process of assessing and quantifying the amount of lost productivity incurred, if caused by neutral or employer-risk events, i.e. events which, according to terms of the contract, the employer bears the risk.

In instances where there is insignificant or no variance in the expected productivity, the contractor has to keep record of the actual productivity for its possible use in measured mile application or as a source of valuable information for future bids.

4.6 Monitoring productivity of change orders

Since change orders are often not programmed at the planning stage, the model recommends contractors to, as much as possible, consider and implement all change order requests at the progress review stage. At this stage, the contractor should plan-out and assess the time and cost impacts for possible agreement with the employer before executing it.

5 Issues for further Research Consideration

The implementation of any cost control/performance monitoring system involves additional overheads (Harris, et al., 2006). Systems that are relatively expensive to operate are therefore likely to be ignored, notwithstanding whatever their strengths and good intentions may be. Failure to implement such systems has often been attributed to excessive management demands of collecting and maintaining detailed project data (Jung and Woo, 2004). To avoid this potential barrier, the proposed model will need further design and refinement and this can only be achieved through empirical research. Issues for consideration in this further investigation will cover;

- existing cost control systems in use, effort involved in operating them and the barriers to their utilisations;
- organisation of cost control and record keeping function within construction companies;
- requirements for the implementations of the proposed model in practice.
- validation of the model; and
- possible integration of the model with other project management functions or ICT systems.

6 Summary and Conclusion

Like in any other industry, lost productivity is an inevitable occurrence in construction projects and tends to result in project cost escalations and delays. Resolving the associated time and cost compensation claims among project parties in the course of a project is an extremely difficult undertaking, often leading to disputes that are resolvable in courts and other forums at high expense. An earlier study by the author, part of which is reported in this paper, suggests that a major reason for the difficulty is the fact that practitioners make use of less reliable methodologies (such as the Global Method) for the resolution. More rigorous ones such as the Measured Mile technique are less popular mainly because they require detailed productivity records, which are often lacking in practice.

This paper hypothesises that the key to addressing this difficulty is the need for a better improved means of keeping productivity records. As an initial effort towards contributing in this direction, this paper proposes a preliminary conceptual process model. In addition to offering a means of effective productivity record keeping, the model also functions as a forward-looking tool for facilitating contemporaneous productivity claims assessment. The model has other features that enhance its ability to promote these functions than what conventional cost control systems can offer. In order to fully realise its usefulness and applicability in practice, the next stage of the author's research will be devoted to an empirical research on issues covering existing cost control systems in use, their barriers, the requirements for smooth use of the proposed model in practice and its validation.

7 References

- Braimah, N. (2008) An investigation into the use of Construction delay and disruption analysis methodologies. *PhD Thesis*, School of Engineering and the Built Environment, University of Wolverhampton. Wolverhampton, UK.
- Brown, J. C. (1995) Prolongation and Disruption- problems of causation for the construction industry. *Construction papers- CIOB*, No. 56, pp.2-7.
- Foster, B. (2000) Monitoring Job-Site Productivity. *The Revay Report*, Vol. 19, No. 2.
- Gaitskell, R. (2003) Global Claims – A sympathetic hearing. *IEE Engineering management Journal*, Vol. 13, No. 4, p. 8.
- Gulezian, R. and Samelian, F. Z. (2003) The Productivity Baseline. *AACE International Transactions*. AACE International, Morgantown, WV. EST.02.1- EST.02.4.
- Hanna, A. S. and Heale, D. G. (1994) Factors affecting construction productivity: Newfoundland versus rest of Canada. *Canadian Journal of Civil Engineering*, Vol. 21, pp. 663-673.
- Hanna, A. S., Russell, J. S., Gotzian, T. W. and Nordheim, E. V. (1999a) Impact of Change Orders on Labour Efficiency for Mechanical Construction. *Journal of Construction Engineering and Management*, Vol. 125, No. 3, May/June, pp. 176-184.

- Hanna, A. S., Russell, J. S., Nordheim, E. V. and Bruggink, M. J. (1999b) Impact of Change Orders on Labour Efficiency for Electrical Construction. *Journal of Construction Engineering and Management*, ASCE, Vol. 125, No. 4, July/August, pp. 224-232.
- Harris, F and McCaffer, R. with Edum-Fotwe, F. (2006) Modern Construction Management, 6th Edition, Blackwell Publishing, Oxford.
- Ibbs, C. W. (1997) Quantitative impacts of project change: Size Issues. *Journal of Construction Engineering and Management*, Vol. 123, No. 3, pp. 308-31.
- Ibbs, W. and Liu, M. (2005). Improved Measured Mile Analysis Technique. *Journal of Construction Engineering and Management*, Vol. 131, No. 12, pp.1249-1256.
- Jung, Y. and Woo, S. (2004) Flexible Work Breakdown Structure for Integrated Cost and Schedule Control. *Journal of Construction Engineering and Management*, Vol. 130, No. 5.
- Kaming, P. F. (1997) Factors influencing craftsmen's productivity in Indonesia. *International Journal of Project Management*, Vol. 15, No. 1, pp. 21-30.
- Klanac, G. P. and Nelson, E. L. (2004) Trends in construction Lost productivity Claims. *Journal of Professional Issues in Engineering Education and Practice*. Vol.130, No.3 Jul. pp. 226-236.
- Kompass (2006). Register of Product and Services, Reed Business Systems in Association with the confederation of British Industry (www.kompass.com, accessed 28/09/2006).
- Leonard, C. A., Fazio, P. and Moselhi, O. (1988) Construction productivity: Major causes of impact. *Transactions of American Association of Cost Engineers*. pp. D.10.1- D.10.7.
- Makulsawatudom, A and Emsley, M. (2001) Factors affecting the productivity of the construction industry in Thailand: The craftsmen's perception. *Proceedings of the RICS Research Foundation, COBRA 2001*. Glasgow Caledonian University, UK, 3 - 5 September 2001, Vol. 2, pp. 478 - 489.
- Merritt, D. (2008) Quantifying Disruption – The Measured Mile, *The journal of the Institution of Civil Engineering Surveyors*, April 2008. pp. 28-32.
- Moselhi, O., Assem, I. and El-Rayes, K. (2005) Impact of change orders on construction productivity. *Journal of Construction Engineering and Management*, Vol. 131, No. 3, Mar., pp 354-359
- NCE (2006) New Civil Engineer's Supplement: *NCE's Consultant File*, Institute of Civil Engineers, Emap, London.
- Olomolaiye, P. O., Wahab, K. A. and Price, A. D. F. (1987) Problems influencing craftsmen's productivity in Nigeria. *Building and Environment*, Vol. 22, pp. 317-323.
- Pickavance, K. (2005). Delay and Disruption in Construction Contracts, 3rd Ed., LLP Reference Publishing, London.
- Presnell, T. W. (2003) Measured Mile Process-Project Controls to Support Equitable Recovery of Production Inefficiency claims. *Journal of Cost Engineering*, Vol. 45, No. 11, pp.14-20.
- RICS (2002) Royal Institute of Chartered Surveyors members directory, UK.
- Roy, S. C. (1998) Survey of Courts Reactions to claims for loss of Productivity and inefficiency, Session 612, *ABA Public Construction Superconference*, December 10, 1998.

- Schwartzkopf, W. and John, J. M. (2001) Calculating Construction Damages. 2nd Edition, Aspen Law & Business, New York.
- Scott, S. and Assad, S. (1999) A survey of the site records kept by construction supervisors. *Journal of Construction Management and Economics*, Volume 17, Issue 3, pp. 375 – 382.
- Thomas, H. R., Riley, D. R. and Sanvido, V. E. (1999) Loss of Labour Productivity due to delivery Methods and Weather. *Journal of Construction Engineering and Management*, Vol. 125, No. 1, pp.39-46.
- Wickwire, J. M., Thomas, J. D., Stephen, B. H. and Scott, B. H. (2003) Construction Scheduling: Preparation, Liability, and Claims. Aspen Law & Business, New York.
- Yang, Y-C., Park, C-J., Kim, J-H. and Kim J-J., (2007) Management of Daily Progress in a Construction Project of Multiple Apartment Buildings, *Journal of Construction Engineering and Management*, Vol. 133, No. 3.
- Zakeri, M., Olomolaiye, P. O., Holt, G. D. and Harris, F. C. (1996) A survey of constraints on Iranian Construction Operatives' Productivity. *Journal of Construction Management and Economics*, Vol. 14, pp. 417-426.
- Zink, D. A. (1990) The measured Mile: Proving Construction Inefficiency Costs. *Journal of Cost Engineering*, Vol. 28, No. 4, pp. 19-21.

Investigating the Relationship Between Construction Contract Documentation Incompleteness and Project Transaction Characteristics: Practical Considerations

Florence Walker¹ and Stephen Pryke¹

¹Bartlett School of Graduate Studies
UCL, London, WC1E 7HB,
United Kingdom

Email: florence.walker@ucl.ac.uk; s.pryke@ucl.ac.uk

Abstract:

Research on Documentation Incompleteness (DI) and loss-and-expense claims (LEXs) is currently underway. As part of this research, an experiment must be designed to define the secondary relationship between DI and project transaction characteristics (Project Uncertainty and Asset Specificity) and determine the optimal level of DI for a specific type of building project construction. To maximize reliability and validity in investigating this secondary relationship, experimental design must ensure that only Uncertainty and Asset Specificity (the main independent variables of interest) are varied and consequences on DI (the dependent variable) are monitored and measured accurately. Generally, however, in construction research, it is neither practical nor possible to ensure that only one of the many independent variables affecting construction project outcomes varies in value; nor is it usually possible to isolate individual dependent variables (Fellows and Liu, 2008; Creswell, 2009). The usual approach is to devise quasi-experiments in which only the main independent variables, except the one of interest, are held *approximately* constant and consequences for the major dependent variable are measured.

This paper proposes a method of comparative study of the construction documentation used and transaction characteristics occurring on similar building projects, and executed around the same time, by similar firms employing similar organizational arrangements. This paper reports on the work in progress by reviewing the methodology and rationale for adoption of the methods proposed for the experiment. Consideration is given to how the method would answer the associated research questions, accomplish the research objectives, and interpret the results. Issues reviewed include: research approach and design and considerations for ensuring acquisition of high-quality, relevant data for analysis. Study limitations and strategies employed to increase confidence in the inferences drawn about the relationship between DI and project transaction characteristics are also presented.

This paper aims to raise awareness of: the nature of available building construction documentation and data on Uncertainty and Asset Specificity; practicality constraints of data acquisition; and appropriate analytical tests and techniques. It should serve as a useful guide for future researchers tasked with determining most appropriate methods for carrying out construction contract documentation research under similarly less than ideal circumstances.

Keywords:

incomplete contract documentation, econometric analysis, transaction costs

1 Introduction

1.1 The Research Problem

Many construction industry observers and recognized conflict management researchers (Diekmann and Nelson, 1985; Hughes and Barber, 1992; Semple et al., 1994; Yogeswaran, Kumaraswamy, and Miller, 1997; Kumaraswamy, 1997; Seeley, 1997; Vidogah and Ndekugri, 1997; Vidogah and Ndekugri, 1998; Tilley, 1999; Hoxley, 2000; Jannadia et al., 2000; Abdul-Malak et al., 2002; Yates and Hardcastle, 2003; Winch, 2010; Chan and Henry, 2005; Love, 2005; Love, 2008; Love, 2009; Murdoch and Hughes, 2008; Itzer and Dikbas, 2009) note that increased commercial pressure on clients and reduced design fee levels lead to poor quality construction documentation. In turn, there is increased commercial pressure on contractors to work to tight time schedules using this poor quality documentation. Inadequate, incorrect, or incomplete construction documentation is considered to be a root cause of conflict, claims and disputes (Yates and Hardcastle, 2003). Often, poor quality documentation necessitates design changes and rework, which contributes to inefficient delays, disruption, reduced efficiencies, and cost increases, and ultimately conflict, claims, and disputes (Hughes and Barber, 1992; Yogeswaran et al., 1997; Kumaraswamy, 1997; Love, 2005). All of this contributes to the construction industry still having one of the worst public images of all British industries (Ball, 1988). A good starting point for dispelling this view of high cost, low productivity, poor quality, and chaotic working practices that has become synonymous with building work, would be at the poor quality documentation root. Through implementation of more modern and appropriate procedures to prevent situations deteriorating into destructive conflict, the image can be improved.

Repeatedly, for the UK construction industry, proposed solutions include: strategies for producing ‘improved’ quality procurement documentation and new policies (Clamp et al., 2007; Bingham, 2009; Love, 2009). Most endorsements, recommendations, and exhortations seem to use relatively arbitrary procedures for deciding documentation quality, and are largely based on subjective information and knee-jerk reactions to unpleasant experiences rather than on rigorous scientific analysis of the evidence. There are many procurement documentation policy debates across the private and public sectors (Bingham, 2009; Klein, 2009; BTET, 2009), but there seems to be no tried and tested scientific procedures for formulating sound construction procurement documentation policies. Essentially, the industry cannot truly know whether there may be project circumstances that would result in desirable outcomes (less costly, more efficient, and free of negative claim impacts) even if procured with ‘less complete’ (even if regarded as ‘less than perfect’ by some) documentation.

The many and varied experiences with documentation produced by the same combinations of documentation preparers (CDPRS) (e.g. quantity surveyors) and construction teams on different projects highlights the elusive nature of the right documentation formula to successfully eliminate unnecessary, and minimize unavoidable conflict (Whitfield, 1994; Yates and Hardcastle, 2003). Therefore, the industry should benefit from a tool designed to aid those responsible for documentation preparation in their decision-making on appropriate levels of documentation completeness to minimize claims, under the varying circumstances that different construction projects present.

The tool should consider the answer to the long-standing question:

‘What is the minimum information required in contracts so that those concerned know what is expected of them?’

(Higgin and Jessop (1965), as cited in Boyd and Wild (2003))

Also, removing documentation formulation guesswork should:

- bring more efficiency to documentation production processes;
- promote sound best practices and recommendations; and
- contribute to reducing frequencies/incidence, magnitude/costs (impacts) and negative effects of conflict, claims, and disputes.

The knowledge to be generated in developing this tool should be of interest to many members of the building industry. Clients will be able to apply a ‘star rating’ to the documentation that their design team professionals were commissioned to produce. Also, they will be made aware of marginal costs for improving the documentation quality and, therefore, be guided in their decisions about whether to authorize further work to improve the documentation produced. It should also be possible to use the tool for adjudication of tenders. Building team professionals should also benefit from removal of the guesswork in determining this unknown level of completeness to ensure that they are not inputting a disproportionate amount of work, aiming for an unachievable target. Essentially, the industry will gain a documentation quality assurance rating tool that can be used as a decision aid to documentation preparation policy.

Ultimately, use of this tool could lead to a better construction industry image because it would free more of project personnel time from attending to expensive and time-consuming conflict resolution activities; and allow more time to concentrate on developing modern and appropriate procedures for planning, scheduling, monitoring, controlling, and administering projects.

1.2 Proposed Solution to the Research Problem

Now, examining the research design behind the rating tool will contribute to the knowledge base on documentation preparation by examining the influence of project transaction characteristics on Documentation Incompleteness (DI) levels chosen by parties. Therefore, this paper takes a detailed look at the modelling processes behind the tool, details the model specification, choice of project transaction variables, and their theoretical and operational constructs. Finally, an explanation is given on how parties arrive at their contractual choice of the optimal level of DI, and how the model can be used to aid in documentation policy decisions.

Relative to other topics, there is very little coverage of inter-firm contractual relationships in the construction industry. Also, contract documentation issues do not feature prominently in the construction project management literature. This is very surprising, given that the contract documentation preparation stage is recognized as the ‘seedbed’ for claims, with nutrients essential for their development to be found in the contract documentation and the information supplied (or not supplied, as the case may be) prior to contract execution. Moreover, the greatest opportunity to prevent claims is during these early project phases, before tenders are invited and before contracts for construction of the works is let (Hughes and Barber, 1992; Smith, 2007). Yet, very few studies specifically examine levels of DI as a documentation attribute. To date, the researchers have not encountered any studies that apply econometric testing of the transaction cost economics (TCE) theoretical propositions to construction documentation. Analysis conducted in other industries has yielded encouraging results (Crocker and Reynolds, 1993; Saussier, 2000) and we draw inspiration from this work to study the contribution of DI to construction project management. The life cycle of claims could be beneficially completed by feeding back lessons learned from this study on optimal levels of construction DI and from the ongoing research on DI impacts on claims management, into the claims seedbed.

We believe our research will make a valuable contribution to understanding claims management in building projects for the following reasons. First, we are unaware of any study to test TCE propositions using data from building projects in the UK construction industry. Second, we examine the entire contract documentation package, whereas other studies usually only attempt to review contract forms. The other studies use subjective measures to assess the quality (Walker and Pryke, 2008), whereas more objective measures are to be deployed in developing this tool. Moreover, direct measures of project transaction Uncertainty and Asset Specificity will be used. Such improvements enable testing of a structural form of the relationship between primitive construction project transaction characteristics and DI. These improvements also facilitate the modelling DI as an endogenous choice by parties aiming to minimize transaction costs (Masten and Saussier, 2000).

1.3 Outline of Paper

The paper is organized as follows. First, we will look at the theories providing a basis for the modelling work. Next, there will be discussion how the philosophical worldview influences research design. Then we will explore the practical considerations in designing the model and developing its constructs, and finally, the econometric analysis methodology for analyzing the problem that prompted the design of the tool. We conclude with a discussion of the contributions that this research will make to the body of knowledge and understanding on DI. Recommendations for future research are also provided. A list of abbreviations of abbreviations used throughout this paper is provided in Section 5.

2 The Theories Providing A Basis For the Work

To enable better appreciation of the importance of the issues motivating this study, the research problem is identified and placed within the larger context of the literature on the topic. The literature was consulted to determine the theories to be used to explore the research question. The law and economics literature on contract theory and (TCE) proved instructive. From here, a better understanding of the incomplete contract documentation topic, and of the most commonly-invoked frameworks for analysis, was developed. Review of these and other closely-related theories also enabled better appreciation of the genealogy of methods for resolution, which informed this study design.

2.1 Incomplete Contract Theory vs Transaction Cost Economics

Initially, theorists used Incomplete Contract Theory (ICT) to develop models for exploring their questions about documentation completeness; and to illustrate and aid thinking on reasons for contractual incompleteness. Under the ICT modelling framework, parties sign incomplete contracts (ICs) because they would like to add contingent clauses, but are unable to because states cannot be verified or because they are too expensive to describe *ex ante* (Hart, 1995). Though theorists were generally satisfied with the ability of ICT to explain motivations for ICs, they deemed the ICT framework models to be oversimplifications of real-life contracting problems. ICT does not contemplate contract drafting costs, legal system constraints for enforcement, and the fact that parties cannot instantly determine complex optimal long-term contracts (Salanié, 2005; Bolton and Dewatripont, 2005). Because of Bounded Rationality (BR) ((parties not knowing precise conditions

under which contracts will be executed (Winch, 2010)), contracting parties must determine some transactions and decisions later. In criticizing efforts to model incompleteness level problems within the ICT framework, Saussier (2000) explained that ICT models generally make the following simplifying assumptions:

- information symmetry exists between contracting parties;
- contractual incompleteness is due to external constraints; and
- incompleteness is assumed as a basis for reasoning, rather than actually explained.

The concern is that these assumptions are not entirely satisfied in practice and, the difficult to formalize, BR concept is excluded from ICT analysis. Some theorists, therefore, turned to TCE for analysis of contractual choice. Under TCE, ICs are modelled as transaction governance mechanisms wherein construction project attributes are evaluated and governance instruments combined to minimize transaction costs (Williamson, 1985).

2.2 Transaction Cost Economics

Before proceeding with the explanation of the application of TCE to solve the research problem, it would be helpful to introduce the TCE theory. The TCE conceptual framework applied in this study was derived from institutional economics, and the work of Oliver Williamson, in particular. Winch (2010) further developed the TCE theory for application to construction project transactions. Williamson's basic proposition, according to Winch (2010), is that total costs of supply are derived from two main components – production costs and supply costs. Production costs are associated with efficient transformation of inputs into outputs and prices are used to signal the most efficient choice of technology. Transaction costs are the costs of coordinating any complex production process and occur when a good crosses a 'technologically separable interface'. For example, in the supply of concrete structures, manufacturing concrete and placing formwork are technologically separable. As the first half of a concrete pour is not technologically separable from the second half, the choice of the most efficient method for pouring is a production cost problem. On the other hand, the choice of the most efficient method for placing formwork – readying it to receive the separately-manufactured concrete is a transaction cost problem.

2.2.1 Transaction Cost Economics and Contract Governance

Winch (2010) conceptualizes contracts for the supply of construction services (usually relational contracts) as transactions and outlines procedures, based on TCE, for determining the most efficient governance mode for construction project transactions. Here the assumption is that the most appropriate mode of project governance is determined within a 3-dimensional space (*See Figure 1*), based on the three main transaction characteristics: Uncertainty, Asset Specificity, and Frequency (Williamson, 1985).

2.3 Transaction Characteristics

2.3.1 Uncertainty

Winch (2010) defines Uncertainty as the lack of information required to take a decision at a given time and attributes Uncertainty to two sources, namely:

- Complexity* – the condition where information is, in principle, available, but too costly or time-consuming to collect and analyze; and
- Predictability* - the condition where the past is not a reliable guide to the future.

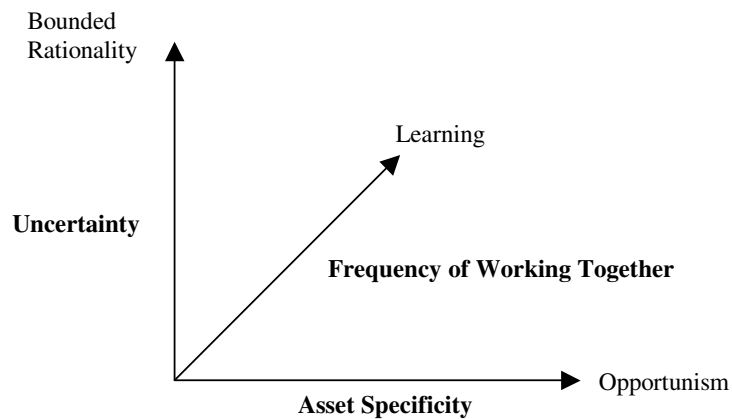


Figure 1 Transaction Cost Economics (TCE) and governance of contracts (adapted from Winch (2010))

Uncertainty affects transactions because it creates BR. BR forces parties to neglect some variables for which the effect on the relationship is difficult to evaluate and makes it impossible to fully measure documentation performance.

2.3.2 *Asset Specificity*

According to Winch (2010), Asset Specificity is the condition where either the client or contractor is limited in their choice of transaction partner because of the specific nature or resources to be supplied. On construction projects, Asset Specificity often takes the form of withholding information from the other party. This Asset Specificity may be: generated *ex ante*, in which case, the problem is one of monopoly or monopsony in the market; or it may be generated *ex post*, because contract-specific investments are made by one or more of the parties. Asset Specificity generates the possibility of Opportunism, as one party may try to exploit the other's disadvantage. For construction projects, specific assets supplied could be: physical; site-specific investments; human; dedicated assets that cannot be redeployed because of the market size; brand name/reputation capital; or temporal specificity (Williamson, 1985; Masten, Meehan, and Snyder, 1991; Yates and Hardcastle, 2003; Winch, 2010).

2.3.3 *Frequency*

Frequency is considered to be a transaction characteristic because repeated transactions allow learning about the behaviour of the other party, and therefore generation of trust in them as a transaction partner. One-off transactions offer no such learning opportunities.

Yates and Hardcastle (2003) also explained the relevance of transaction Frequency and used investments in markets of varying sizes to illustrate their discussion of Frequency and transaction costs. They state that: due to potential difficulty in recovering specialized asset costs in small markets (as compared with large markets), one would expect to observe general purpose investments in small markets. Applying similar reasoning to the study of transaction costs, they point out that the benefits of specialized governance structures are expected to be greatest for transactions that are supported by significant investment in transaction-specific assets (nonstandard transactions) (Williamson, 1985). The expectation is for the cost of specialized governance

structures to be more easily recoverable for large transactions of a recurring kind. However, specialized governance structures come at a great cost and these additional costs should therefore, be justified through consideration of whether the transaction volume processed through a specialized governance structure utilizes the specialized structure to capacity.

Although recent construction clients have been making changes to increase Frequencies (through initiatives such as partnering and prime contracting) to achieve learning benefits, typically, transaction Frequency is low in construction - often effectively, unity for most client/contractor relationships (Winch, 2010). Also, in construction the different options for governance structures is manifested in the various procurement routes (e.g. traditional, construction management, design-and-build) and project performance comparisons can be made both within the routes and across routes. For this study, comparisons would be between routes. In this study, analysis based only on levels of Uncertainty and *ex post* Asset Specificities under which the transaction is made. Frequency is not considered for this study.

2.4 TCE and Documentation Incompleteness

The literature also demonstrates the usefulness of the TCE theory in understanding the diversity of DI levels used to govern inter-firm relational contracts in the construction industry.

Here, BR is invoked to explain the thought processes involved in documentation preparation for ICs. Now, negotiating construction projects is costly, engaging managers and lawyers. At some point during negotiations, costs of contemplating specific arrangements to cover unlikely contingencies outweigh benefits. Also inability/unwillingness of courts and third parties to verify *ex post* values of certain observed variables is costly. Therefore, parties decide on a level of incompleteness for the contract documentation (the legal basis of the relationship) that would minimize these costs (Salanié, 2005).

The TCE view also assumes that there are alternative governance mechanisms and behaviour prescription. These alternative governance mechanisms and behaviour prescription capture interpersonal and inter-organizational aspects of transaction governance arrangements by considering issues such as trust and perceived equity in governance. Alternative governance mechanisms include implicit social conventions and the parties adapt their performance to contingencies that actually arise during the life of the contract. Baum and Oliver (1991) observed this alternative governance phenomenon for childcare service organizations in Metropolitan Toronto, Canada between 1971 and 1987. To use a construction industry example: In a close-knit community, a contractor may be reluctant to act opportunistically in pursuing a claim against the client of a school that his child attends. Al-Najjar (1995) illustrates the TCE view of the contractual arrangement.

The boundaries will vary across transactions to reflect differences in project characteristics. Boundaries will also change in response to the changing environments (e.g. market conditions, technical innovations, etc.) Therefore, differences in the incompleteness levels that can be observed from project transaction to project transaction represents different combinations of governance mechanisms (*See Figure 2*). The chosen governance mechanisms take the transaction characteristics into consideration and are combined in such a way as to minimize overall transaction costs. (*See Walker and Pryke (2009) for a more detailed treatment*).

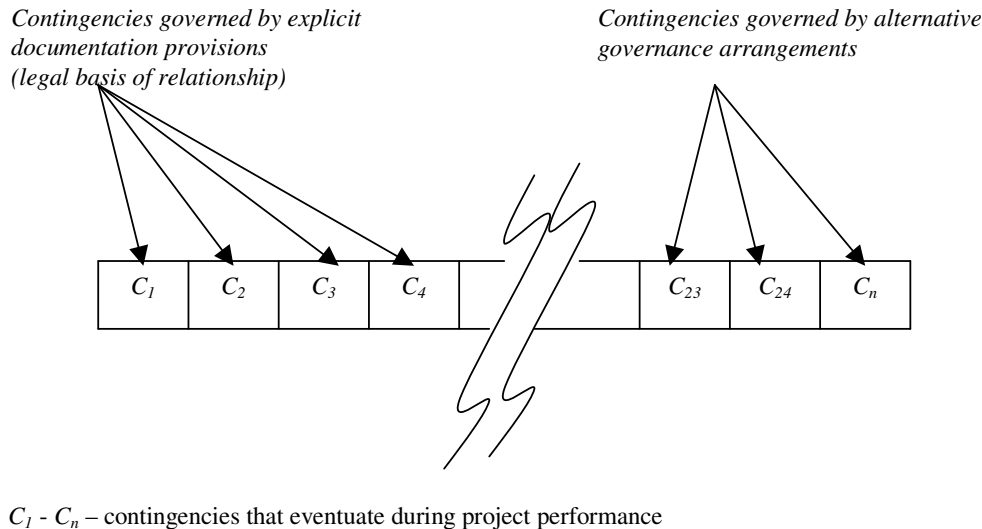


Figure 2 Combination of Governance Mechanisms in Relational Contracts (Conceptualized from Al-Najjar's (1995) description)

2.5 Modelling the Documentation Incompleteness Choice Using TCE

We now return to the discussion of the use of TCE in determination of the optimal level of DI for construction building projects. Because the principal investigator (PRI) holds a pragmatic worldview, the research problem, was emphasized and a search was carried out for all available approaches to understand the problem. Two empirical studies (Crocker and Reynolds, 1993; Saussier, 2000) were found to provide the best application of the TCE framework from which knowledge of the research problem could be derived. In these studies, the choice of the efficient level of DI is represented as an endogenous trade-off of costs and benefits in transactions between parties aiming to minimize transaction governance costs. There was one main shortcoming in both studies when they were considered for use as a guide to the research problem. Specifically, these studies were used to analyze only the contract forms whereas our research problem examines entire documentation packages, including the contract forms. Nevertheless, the studies were still deemed instructive because the principles are still seen as being applicable to entire documentation packages. This is because it is, in fact, the entire documentation package that is used to govern the legal basis of the relationships in the construction industry and more often than not, contractual effect is generally given to all of the contract documents in the package (Hughes and Barber, 1992). Moreover, documents are chosen simultaneously and may interact, so empirical studies should, ideally estimate all of the contractual provisions (Saussier, 2000) Therefore no modifications should be necessary to extend them to this study, which forms part of a larger mixed methods design. Saussier (2000) was deemed the better guide for this study of construction documentation packages because Crocker and Reynolds (1993) recommended the use of contractor opportunism as a proxy for the Asset Specificity project characteristic and only focused on price provisions in contracts. Moreover, Crocker and Reynolds (1993) modelled DI as an exogenous choice, which is not the preferred approach for this study. Saussier (2000) was considered an improvement upon the concepts introduced by Crocker and Reynolds (1993) because DI was modelled as an endogenous choice, direct measures of Asset Specificity were employed and the entire contract form was reviewed for the analysis.

Both approaches proposed a quantitative method – econometric analysis using (Ordinary Least Squares Regression) (OLSRA) for data collection, analysis and interpretation. A single equation model is used to allow empirical verification of the hypothesized (mathematical) relationship between transaction characteristics and DI . The quantitative method was seen as appropriate for this study because the TCE paradigm translates the research problem into one of understanding the relationship between project transaction characteristics and DI. Essentially, the focus is on the relationship between DI, as the dependent variable, and Uncertainty and Asset Specificity as independent variables. The TCE theory propositions were reassessed to come up with hypothesized relationships between the variables. In keeping with the advice for when quantitative methods are chosen (Creswell, 2009), specifying this relationship then became central to answering the research question. The aim here was to enable generalization of the findings to a population of the specific type of building project that will be identified for the larger study.

The research problem was then restated for solution within the TCE framework using this quantitative approach. When clients and contractors decide to collaborate on any project of modest size, usually they prepare documentation to create a contractual interface, to guide the transaction (Hughes and Barber, 1992; Ramus et al., 2008; Smith, 2007). The purpose of this documentation is to create legally-enforceable agreements defining rights, obligations, and procedures for project contributors (Abdul-Malak et al., 2002; Semple et al., 1994; Furlotti, 2007). Ability to clearly and fairly apportion and assign risks between parties is also a measure of contractual efficiency (Hughes and Barber, 1992; Hartman and Snelgrove, 1996; Furlotti, 2007). Additionally, construction documentation should provide incentives to effectively motivate the contractor to make full use of his expertise and overcome difficulties expeditiously. According to Saussier (2000), contract documentation is designed to facilitate transactions between parties, and the documentation achieves its purpose to the extent that it:

- aligns parties' expectations with regard to the other's obligations under agreement;
- provides incentives for parties to fulfill these obligations;
- prevents costly disputes from arising; and
- provides a basis for resolving disputes that arise despite best efforts, whether the disputes arise from opportunism or from honest misunderstandings.

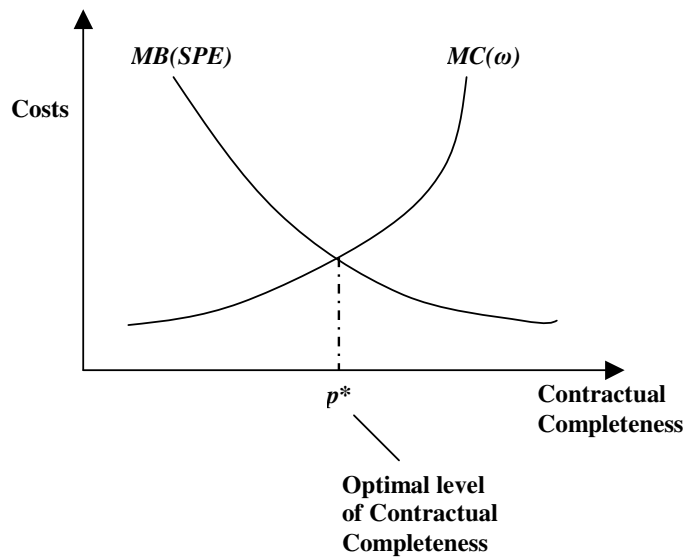
The documentation must be correctly designed and adequately prepared in order to maximize the gains from the project transaction. The level of DI required is difficult to evaluate and, therefore, some contracting parties deliberately prepare incomplete documentation with provisions that are vague or unspecified. To answer the research question:

‘What is the minimum information required in contracts so that those concerned know what is expected of them?’

(Higgin and Jessop (1965), as cited in Boyd and Wild (2003)),

the TCE presentation tests the theory that this deliberate incompleteness reflects the parties' aim to minimize the project transaction costs.

Those adopting a TCE view of contractual choice (Crocker and Masten, 1991; Crocker and Reynolds, 1993; Saussier, 2000) explain that if contracting were costless, it would be, in principle, possible to design arrangements complete enough to circumscribe all surplus-eroding redistributive tactics and intricate enough to mitigate investment distortions. In practice, however, the costs of identifying contingencies and devising responses increase rapidly in complex or uncertain environments, placing economic limits on abilities to draft and implement elaborate contractual agreements.



where
 p^* = Efficient Level of Documentation Completeness
 ω = Degree of Project Transaction Uncertainty
 SPE = Degree of Project Asset Specificity
 MC = Marginal costs of contracting
 MB = Marginal benefit of increased completeness

Figure 3 **The Relationship between Degree of Contractual Completeness and Marginal Costs and Benefits of Contractual Completeness**
 (adapted from Saussier (2000))

When designing a contract, the parties mitigate *ex post* opportunism and investment distortions by using more complete documentation, but at the cost of increased resources dedicated to crafting the documents *a priori*. As a consequence, transaction characteristics that generate increased contracting costs should result in efficient contract documentation being less complete, whereas conditions that exacerbate the potential for *ex post* inefficiencies should lead to more exhaustive agreements (See Figure 3).

In keeping with the theoretical principles identified to guide this study, clients and contractors were regarded as being confronted with a trade-off between *ex ante* costs of crafting more complete documentation and *ex post* inefficiencies associated with less exhaustive arrangements. (See Figure 3) This lead to testable implications about an efficient level of project documentation. To test these implications, this study will examine a panel dataset consisting of a specific type of publicly-procured building projects to determine whether to accept or reject the reassessed TCE theoretical propositions as hypotheses, and determine whether the degree of contractual incompleteness actually chosen by the parties in practice, reflects a desire to minimize the cost of contractual exchange. The TCE theory propositions will be tested econometrically, using a building contract documentation database for projects procured by a single UK public sector client. The level of completeness will then be analyzed as an endogenous choice of clients and contractors.

3 Research Strategy

3.1 Data Requirements

3.1.1 *Contract Documentation Package Dataset*

The guiding studies (Crocker and Reynolds, 1993; Saussier, 2000) employ a panel dataset to gain better insight into the dynamics of documentation development processes for a single client (Dougherty, 2007; Gujarati and Porter, 2009). Efforts will be made to assemble a panel dataset for analysis. Because the success of the study depends on acceptable participation rates, efforts will be made to maximize these rates. In this light, Saussier (2000) recommended that the following be taken into consideration in choosing the projects for the study:

- Public sector client levels of interest in the study of DI in regard of TCE;
- Projects that engage the services of private-sector contractors are favoured because the assumption is that such transactions are more likely to be driven by cost minimization.

3.1.2 *Sample Size*

It will be necessary to make statistical inferences about the population of construction documentation packages, from the results obtained from the sample. Sample sizes need to be large enough to predict the behaviour of the population as accurately as possible. An effort will be made to choose a dataset with a wide enough variety of projects, to achieve heterogeneity in levels of Uncertainty and Asset Specificity (transaction characteristics) to yield an adequate modelling solution.

3.1.3 *Scope*

Also, it is generally considered good practice to restrict research studies so that adequate depth and rigour of investigation of the topic can be undertaken. Therefore, a number of measures can be adopted to limit the scope of the research to reasonable bounds and ensure a homogenous sample of construction documentation packages (as the unit of analysis) to allow analysis of only DI as the only independent variable (Fellows and Liu, 2008; Creswell, 2009). Some of these measures are outlined below.

3.1.4 *A Specific Time Period*

Because the dataset will be made up of past completed projects, accurate results will depend on memories of those contributing to the study. Therefore, only projects commenced up to a maximum of three years prior to the start date of the research date will be reviewed.

3.1.5 *A Specific Construction Industry Sector*

Preliminary studies will be carried out determine which sectors experienced higher volumes of construction activity. Experienced industry professionals can be consulted to seek their views on construction activity across the different sectors. Also, companies providing procurement services to the construction industry will be approached for assistance in this exercise. Construction economics and statistics data may also provide some insight into construction activity in the recent past.

3.1.6 *A Specific Procurement System (Governance Structure)*

Contract documentation packages chosen will be from projects where the traditional procurement route was used. By this route, design is virtually complete before calling tenders (Clamp et al., 2007). In this way, the degree of *ex ante* DI will be limited. There will be some degree of *ex post* flexibility and provision for this will be made by way of monetary, and presumably, time contingency in schedules. Any such tendencies should be observable in the documentation. Projects procured using JCT forms with Quantities will be reviewed because these are used on the greatest number of projects procured in the UK (RICS, 2010). The entire documentation package (as opposed to only the contract form, which is the case in the guiding studies) will be examined, as this is seen as representing the contract governance agreement. Focusing on this group should ensure a large pool from which to draw a sample of projects for analysis.

3.1.7 *Client Documentation*

This research for the UK construction industry will be developed within the scope of Williamson's (1975) conceptual framework, as presented by Winch (2001), Winch (2010), and Yates and Hardcastle (2003). Only data and information from projects governed in the vertical dimension will be reviewed. Also, although many contributors to the building process should be able to gain insight from using the model, that will be developed for use by public sector clients.

3.1.8 *Project Size*

The right balance must be struck here. Project values cannot be too large nor too small. It is believed that projects of approximately £5mn in value would provide generate the type of issues and consideration that would be of interest for this study and lend them selves to the analytical methods chosen.

3.1.9 *Project Claims*

This study is part of a larger research project that focuses on impacts of DI on management of claims advanced by the main contractor against the client. Therefore, it would also be wise to focus on documentation for projects where claims were a major issue. Loss-and-expense is one of the more frequently-recurring subject of claims advanced by contractors against clients (Kennedy and Milligan, 2005) Also, LEXs can be directly linked to poor quality documentation (Hughes and Barber, 1992; Love, 2005; Love, 2009), which make them a good gauge of documentation performance.

3.2 **Quasi-Experimentation**

Because this is a construction project-based experiment, it will not be possible to allow only the project transaction characteristics (Uncertainty and Asset Specificity) to vary in value. Also, in reality it would not be possible to only restrict effects of any treatments to Uncertainty and Asset Specificity, nor on DI Levels. In other words, many unwanted, external effects are likely to impact on the data and it will be impossible to control them, or, even if identified, allow for them adequately in the analysis. Hence, a comparative study on similar projects, executed at about the same time, by similar firms employing similar organizational arrangements, will be carried out. This is in an effort to ensure that all other variables, except Uncertainty and Asset Specificity, that could impact on DI will be held approximately constant. Then the consequences for DI will be measured.

For this quasi-experiment, the aim is to determine whether varying the levels of Uncertainty and Asset Specificity influences DI choice. Also, the aim is to estimate and/or predict the population mean value of Uncertainty and Asset Specificity in terms of optimal level of DI for the specific type of building project.

Therefore, for each project transaction in the dataset, the variables in this study are:

- Marginal Transaction Costs associated with the level of documentation completeness;
- Marginal Gains accruing from establishing the level of documentation completeness;
- Project Uncertainty;
- Asset Specificity; and
- Level of Documentation Incompleteness.

3.3 Theoretical and Operational Constructs

3.3.1 *Transaction Costs associated with the level of Documentation Incompleteness*

According to Crocker and Reynolds (1993) and Saussier (2000), the primary costs associated with preparing more complete documentation are:

- cost of writing the agreement
- information costs - difficulties of identifying feasible contingencies
- negotiation costs to arrive at mutually-acceptable responses
- potential maladaptation/renegotiation cost of being trapped in a bad contract (costs of resolving claims)

Marginal costs (MC) are defined as: the additional cost of one more unit of service (McPake and Normand, 2008). For this study, they will be measured by subtracting the costs of preparing documentation rated at level ($DCOMP_i$), from the preparation costs of producing documentation that has been rated at the level ($DCOMP_{i+1}$). These documentation ratings will be obtained by using the DI rating instrument described in Walker and Pryke (2009) (WPM). MC can also be computed by plotting total costs for preparing documentation to level $DCOMP_i$ against corresponding completeness levels calculated by the WPM and then taking the slope of the total costs curve at the level of DI in question.

Crocker and Reynolds (1993) and Saussier (2000) also explained the effect of project Uncertainty on costs of preparing complete documentation. While this may be fairly straightforward in simple environments, with proximate dates of contract execution, the task becomes more daunting when possibilities are more numerous and the time to performance remote. This is because remote times to performance means greater uncertainty transaction levels, which translates to more difficulty, expense and riskiness in preparing more complete documentation (particularly because of the costs of being trapped in a bad contract). Moreover, holding environmental complexity constant, parties are likely to face increasing costs of crafting more complete agreements as the additional contingencies become more hypothetical. This is because contingencies are more expensive to add to more complete contracts. Also, more precise documentation lessens the importance of unaccounted-for contingencies and make it harder to evaluate unaccounted-for contingencies, because of the lack of experienced parties. Accordingly, the marginal cost of preparing more precise documentation is assumed to be increasing in the degree of contractual completeness and in the level of environmental complexity, ω .

From the foregoing, we arrive at **Proposition 1**, adapted from Saussier (1990). For this study, we test **Proposition 1**:

The marginal cost of documentation that aims for completeness increases with the level of transaction uncertainty and the level of documentation completeness.

3.3.2 Gains from preparing more complete documentation

The principal gains accruing from preparing more complete documentation are:

- for contractants who have developed specific assets, a reduced exposure to opportunism of the other party; and
- savings on repeated negotiation costs (Crocker and Reynolds, 1993; Saussier, 2000).

Marginal gains (MB) are defined as: the reduction in gain accrued by producing one more unit of service (McPake and Normand, 2008). For this study, they are measured by subtracting the total benefits of using documentation rated at level (DCOMP_i), from the total benefits accruing from using documentation that has been rated at the level (DCOMP_{i+1}), by the WPM. To our knowledge, there is no practical measure of MB for building construction project transactions. Therefore, for a project rated at DCOMP_i, we propose that MB be computed from the difference of the total project costs anticipated by the parties and actual realized costs for a project rated at DCOMP_{i+1}. This is based on the approach used in Bajari et al. (2007), which compares the costs of rigid and more flexible contracts. Similarly to MC, MB can also be theoretically computed by plotting total benefits derived from using documentation prepared to completeness level DCOMP_i against corresponding completeness levels calculated by the WPM and then taking the slope of the MB curve at the level of documentation completeness in question.

Now, tighter agreements reduce parties' abilities to engage in efforts designed to effect privately-favourable distributions of the *ex post* contractual surplus (Williamson, 1985; Goldberg, 1985; Masten, 1988). Such restrictions benefit the parties when the available redistributive tactics consume resources. For example, for contingencies not expressly covered by the documentation, the parties incur the resource cost of bargaining. (denoted by the increasing function $B(SPE)$, where SPE indicates the level of Asset Specificity (SPE) in the transaction. In this case, the expected bargaining costs would be $pB(SPE)$, and the marginal benefit of increased completeness, $B(SPE)$ is the reduction of resource costs associated with bargaining.

In addition to the direct costs of redistributive activities, the possibility of unconstrained renegotiation raised by DI generates potential inefficiencies by discouraging *ex ante* investment (Tirole, 1986; Hart and Moore, 1988). This problem arises because of the parties' inability to precommit to an enforceable division of the contract surplus in all future states (complex allocation rules). Investors recognize that part of the benefits from their efforts may accrue, as a consequence of future redistributive activities, to the other party. In such scenarios, *MB* for a more complete contract is a reduction in the magnitude of investment externality (a side effect of consumption or production which is not traded on the market or taken into account in setting a price (McPake and Normand, 2008)), which is also directly proportional to the Asset Specificity at stake in the transaction. The probability of *ex post* renegotiation is expected to be lower in contracts aiming for completeness (Grossman and Hart (1986), Saussier, 2000). This is very uncomfortable for parties who have developed specific assets.

Therefore, where the transaction involves specific assets, the more the contract specifies the transaction, the smaller the probability that the contract will be renegotiated. This increases *ex ante* investment incentives. Investment incentives are directly linked to the presence of specific assets that generate quasi-rent (Klein et al., 1978, as cited in Saussier, 2000). For more specific assets, the assumption is that the marginal gain from more complete documentation decreases with the level of

completeness because the final contingencies are less probable and exert a lesser incentive effect on investment decisions. Consequently, under the assumption that investment incentives decrease as contracts become more complete, the *MB* from contractual completeness is a decreasing function, as depicted in Figure 1.

From the foregoing, we arrive at **Proposition 2**, adapted from Saussier (1990). For this study, we test **Proposition 2**:

The marginal gain from documentation that aims for completeness increases with the appropriable quasi-rent at stake in the transaction and decreases with documentation completeness.

3.3.3 Project Transaction Uncertainty and Asset Specificity Variables Uncertainty

From descriptions of pre-tender phase activities described in project management texts (Hughes and Barber, 1992; Smith, 2007; Ramus et al., 2008), it is clear that this phase is characterized by high levels of Uncertainty (Mission Uncertainty and Dynamic Uncertainty).

A dataset that is heterogenous in Uncertainty, could include projects that vary in one or more ways, such as and including: times to project construction phase; physical complexities of project sites on which construction will be procured; requirements for new technologies to solve the particular problems posed by the project at hand; extents to which the new building proposed is a copy of existing ones; extents to which standardized components can be used; extent of design ambiguities in documentation; levels of detail in geotechnical surveys; acquisition of regulatory approvals; number of shop drawings provided by component suppliers (Yates and Hardcastle, 2003; Winch, 2010).

3.3.4 Asset Specificity

In keeping with Saussier's (2000) approach, for this study, the level of Asset Specificity measured is a direct assessment of the amount of specific resources that the main contractor supplies to the project. Recalling possible specific assets listed in Section 2.3.2 a dataset that is heterogenous in Asset Specificity could include projects with varying levels of site-specific investments made by the contractor for each project.

Accurate measures of the level of Uncertainty and Asset Specificity in each project transaction in the dataset will be obtained for testing the **TCE Propositions** through, a series of semi-directed interviews with the professionals who contributed to the project in question. During these interviews, guidance will be sought from project personnel in charge of negotiation and monitoring to devise measures of Uncertainty (*UNC*) and Asset Specificity (*SPE*) in each building project transaction. Also, there will be reviews of each contract documentation package in the dataset to determine the variability in *UNC* and *SPE* across the dataset; upper and lower bounds for *UNC* and *SPE*, and gradations of *UNC* and *SPE*.

3.3.5 The Level of Documentation Completeness

The definition of documentation completeness adopted for this research is based on that proposed by Saussier (2000). Basically, one documentation package is more complete than the another if it gives a more precise definition of the transaction, outlines adequate risk-sharing, includes appropriate incentives to motivate expeditious resolution of difficulties, and the means to carry out all of this. Many different types of documentation exist, but the level of DI chosen by the parties reflects are desire to save transaction costs.

In order to econometrically test DI levels for the dataset, the documentation levels must be classified according to their levels of completeness. This is achieved by deploying the methodology, proposed in Walker and Pryke (2009), for evaluating the level of completeness for a documentation package (i.e contract form and all production information (PI)). This approach is based on a TCE view of ICs, which contemplates BR. (*See Section 2.1*). Applying this definition of ICs enables a feasible upper bound for documentation completeness to be determined through a measure that considers all completeness dimensions (complexity, contingency planning, and contractual ambiguity/specificity).

The Walker and Pryke (2009) methodology draws inspiration from best practice recommendations for preparing production information (PI) (Drawings, Specifications, and Bills of Quantities (BoQs)) and contract formulation, combining them into one uniquely comprehensive methodology for assessing construction documentation completeness.

From the literature, contract completeness is made up of multiple dimensions, namely:

- Ambiguity/Specificity of the provisions in the documentation (ASD);
- Complexity of the contractual governance apparatus employed to inflict penalties for uncooperative or violative behaviours (CD); and
- Contingency Planning strategies to achieve efficient adaptation (CP).

To devise the rating instrument for assessing documentation completeness (WPM), those knowledgeable of contract formulation processes, such as lawyers, client and contractor executives and personnel, will be consulted, through interviews, to obtain more detailed knowledge of project conditions; provide greater insight into contract formulation; and confirm and/or supplement literature findings. Also, these individuals will be consulted to assist with dimension coding for the final rating instrument.

Next, using the WPM, the documentation package for each building project (the unit of analysis) in the dataset will be reviewed to assign a DI score, *DCOMP*, that will contemplate: completeness dimensions (complexity, contingency planning, ambiguity/specificity); risk sharing; inclusion of adequate incentives to motivate expeditious resolution of difficulties; and completeness of PI. (*See Walker and Pryke, (2009) for a detailed treatment of the methodologies for operationalizing each completeness dimension.*) Application of the WPM should add content to the chosen construction industry subsector.

3.4 Model Development

Econometric analysis will be used to analyze the problem of determining the optimum level of DI for the chosen building type to be procured by a public sector client.

Econometric analysis will proceed in accordance with the traditional econometric methodology, which involves:

- i) Statement of the hypotheses;
- ii) Specifying the mathematical model;
- iii) Specifying the econometric model;
- iv) Obtaining the data;
- v) Estimating the model;
- vi) Testing the hypotheses;
- vii) Forecasting/predicting the values of DI on the basis of known or future expected values of project transaction Uncertainty and Asset Specificity; and
- viii) Using the model for control or policy formulation purposes.

3.4.1 Statement of the hypotheses

From Sections 3.3.1 and 3.3.2, recall the hypotheses:

Proposition 1:

The marginal cost of documentation that aims for completeness increases with the level of transaction uncertainty and the level of documentation completeness.

Proposition 2:

The marginal gain from documentation that aims for completeness increases with the appropriable quasi-rent at stake in the transaction and decreases with documentation completeness.

3.4.2 Specifying the Econometric Model; and Obtaining the Data

$DCOMP$ represents the DI level decision, with $DCOMP \in (0,1)$. For totally complete documentation, $DCOMP = 1$, and there will be no opportunities for *ex post* bargaining. For totally incomplete documentation, $DCOMP = 0$, and there are no strictures on all the terms under which future trade may be effected. Thus, the choice of DI is illustrated as shown in *Figure 3*.

For each building project transaction, at the time of contract negotiations, parties will have perceptions of the current level of project transaction Uncertainty, ω and the current level of asset specificity, SPE . Parties face a trade-off between costs and gains of $DCOMP$ and will choose a level of incompleteness that equates marginal costs and marginal gain. On one hand, variables leading to increased ω , should shift MC upward, resulting in less exhaustive documentation. On the other hand, variables increasing SPE should shift MB upward, making more structured agreements attractive.

This study proposes that parties, without carrying out complex cost-benefit analyses, are, nevertheless, guided by this calculus approach that is reflected by the trade-off described in the foregoing. This study employs Ordinary Least Squares Regression Analysis (OLSRA) to test these Propositions as hypotheses, using a panel dataset from a sample of construction projects (of the chosen building type), by estimating the structural form relationship (a function of variables that shifts the benefit and gains schedules).

$$DCOMP_{it} = \alpha_i + \beta\omega_{it} + \gamma SPE_{it} + \epsilon_{it}$$

where

$DCOMP$ - is the measured level of documentation incompleteness for the documentation in the contract signed with contractor i , at date t .

ω_{it} - are the variables affecting MC when contracting with contractor i at date t .

SPE_{it} - are the variables affecting MB when contracting with contractor i at date t .

α_i - is a contractor-specific effect

β, γ - are slope coefficients/model parameters

ϵ_{it} - is the independent and identically-distributed error terms (disturbance) with zero mean and constant variance σ^2 .

The error term represents all the factors that affect Documentation Incompleteness, but not taken into account explicitly.

This specification of the Relationship suggests that there is a linear relationship between $DCOMP$ and the project transaction characteristics (Uncertainty and Asset Specificity).

3.4.3 Estimating the Model

Estimating the econometric model involves obtaining the numerical values of β and γ (Gujarati and Porter, 2009). These numerical estimates of the parameters give empirical content to the function defining the relationship. The contract documentation package panel dataset will be used to estimate the parameters of the relationship between Uncertainty and Asset Specificity (independent variables) and Documentation Completeness (dependent variable). OLSRA, a statistical technique is the main tool that will be used to obtain these parameter estimates.

3.4.4 Testing the Hypotheses

Hypothesis testing (statistical inference) will determine the reliability of the parameter estimates obtained. In other words, the hypothesis testing should provide confidence that the numerical results for the parameter estimates were not likely to have occurred by chance (Dougherty, 2007).

Assuming that the fitted model is a reasonably good approximation of the real relationship, suitable criteria must be developed to determine whether the parameter estimates obtained are in accordance with the expectations of the theory being tested. From the theoretical descriptions of the relationships, the expectation is that straightforward estimates of the empirical relationship using OLSRA, will result in β having a negative sign, and γ having a positive sign. This would suggest that: as measured project transaction Uncertainty (*UNC*) increases, this will result in less complete contract documentation (decreased *DCOMP*), while an increase in measured project transaction Asset Specificity (*SPE*) will result in more complete contract documentation (increased *DCOMP*).

3.4.5 Model Verification

Model verification is, essentially, a post-mortem, carried out to find out if the results from the model accord with *a priori* expectations (Creswell, 2009). Because assumptions were made to specify the model, verification checks to see whether the model implements assumptions correctly and does what it was intended to do. In other words, it asks the question: ‘Is the model doing the thing right?’ (Forrester, 1999).

After specifying the econometric model, the project transaction data from projects in the dataset that were not used for model construction will be used to determine whether the model structure is correct. This will be achieved by testing the model through examining the outputs resulting from the model under a given set of inputs. If outputs are inappropriate (i.e. they do not approximate to expectations of what a good model of the reality would yield), then the model will be reassessed through a ‘debugging’ process to see whether variables were coded correctly. The necessary corrections will be made and the model will be re-verified. The process will iterate through these steps of reassessment and re-verification until when outputs are appropriate. When outputs are appropriate, then the model will be deemed verified. Should the model be deemed verified it will then be validated.

3.4.6 Model Validation

Validation checks to see whether the (unbiased, consistency, and efficiency) assumptions made in estimating the model parameters are reasonable, with respect to the real system. The aim of validation is to demonstrate whether the model is a reasonable representation of the actual system. According to Forrester (1999), validation asks the question: ‘Is the model doing the right thing?’

Verification techniques are more general in comparison with validation techniques, which are more specific to the model in question.

Therefore, after constructing the model, its performance will be scrutinized to determine whether it is suitable for the stated objective, which, for this study, is to determine the optimal level of DI for the chosen building type. In validating the model, its output resulting from known inputs will be compared with realizations of the reality, such as an *ex post* forecast of DI by quantity surveyors who are very experienced in authoring documentation packages for the chosen building type. To examine model consistency over a range of conditions (preferably including extremes), an effort will be made to carry out this type of validity testing for several sets of inputs and known outputs of the reality.

3.4.7 *Forecasting/predicting the values of Documentation Completeness on the basis of known or future expected values of project transaction Uncertainty and Asset Specificity*

The OLSRA used during the econometric analysis will produce descriptive statistics calculated for observations and measures made during the experiment. These statistics are means, standard deviations, and ranges, and from these results, the optimal level of DI can be obtained. This optimal level will correspond to the intersection point of *MB* and *MC*. (See Figure 3 for a graphic representation for the theoretical model). The results from this study will also inform a later stage of the ongoing study that examines how deviation from this optimal level of completeness will affect LEXs management. CDPRs will then be able to use this information as a guide to the level of DI to which they would like to aspire for this chosen building type for this client. From the *MB* and *MC* information, not only will the client have a measure of the documentation quality produced, but they can be guided in sanctioning further work to improve documentation quality.

The empirical results can also be used to gain insight in interpreting the chosen client's documentation record for the chosen building type. Once the parameters have been estimated, this will enable statements to be made about how much DI will go up or down as a result of certain changes in the transaction characteristics. This would be akin to conducting sensitivity analyses on contract documentation package incompleteness levels.

3.4.8 *Using the model for control or policy formulation purposes*

Also, it should be possible to use the model as an aid to control or policy formulation procedures in public sector organizations. For example, the chosen public sector client could be guided in their choice of projects (and therefore, project transaction characteristics) in order to guarantee a desired outcome, in terms of DI levels.

The results will have implications for procurement policy, more generally. For example, there is currently much promotion of the NEC as the solution to all contracting problems in the construction industry. The Latham Report (1998), as cited by Morton and Ross (2008), suggested that the NEC came closest to meeting the thirteen ideal principles of a 'modern contract'. While aim for more comprehensive contracts to meet this ideal may suffice to mitigate *ex post* inefficiencies that the documentation was designed to take care of, this type of analysis will indicate whether such benefits may be dwarfed by the costs of preparing truly 'modern contracts' (that may be more 'complete' than the optimal), for projects with particularly complex transaction characteristics. With this tool it would be possible to carry out similar analyses for documentation that includes any contract form. Comparative studies can then be used to settle debates such as that described by BTET (2009) about the most economical choice.

Secondly, Crocker and Reynolds (1993) are of the view that procurement officers should be granted the latitude to craft agreements on a case-by-case basis, where the design of particular documentation packages would depend on the specifics of both the product and the contractor investment. They state that policies attempting to impose homogeneity in contract design, either across contractors, or over time would be misguided, and likely to significantly raise the costs of effecting contractual exchanges for projects. This tool will allow this kind of latitude in crafting agreements and aid the decision-making process on whether to implement homogeneous or heterogeneous contract design policies.

Finally, the potential benefits of arrangements such as Framework agreements can be examined. Proponents suggest that much benefit will accrue to clients through aggregation of demand and delivery through frameworks. Skeptics remain unconvinced and cleave to market economics, providing the driver for production efficiency and suggest that Frameworks are an artificial intervention to yield transactional efficiencies (Morton and Ross, 2008). Now, (Crocker and Reynolds, 1993) have shown that sourcing from multiple suppliers of construction services may generate resource savings for documentation preparation by permitting the implementation of less complete agreements. They suggest that this might be due to the discipline provided by the existence of an alternative supplier, which mitigates the necessity for complete and costly contract documentation to constrain future contractor behaviour. This tool proposed will allow examination of the ability of split awards under arrangements such as Framework agreements to provide their administrative and pricing benefits.

4 Conclusion

There are a number of ICT models in the economics literature that show how contractual completeness may influence parties' incentives to invest in order to realize a transaction. They usually show *ex ante* investments are distorted by such contracts. Also, these ICT models assume that the motivation for signing ICs is because parties are unable to add the contingent clauses that they would like to because states cannot be verified or because they are too expensive to describe *ex ante*. Therefore, in those models, level of documentation completeness is usually given and contractual incompleteness is modelled as an exogenous variable (Crocker and Reynolds, 1993; Saussier, 2000).

In this paper, the less frequently-encountered TCE approach to the problem is reviewed and its application in a construction contracting context is proposed. Special emphasis has been placed on econometric analysis in general, and more specifically, how to test TCE propositions. For the proposed study, Documentation Incompleteness is not taken as a given, nor is it used as a basis for reasoning. Rather, BR is invoked to explain the thought processes involved in documentation preparation for ICs. Also, contractual incompleteness observed in response to different transaction characteristics is explained as resultant combinations of the legal basis for relationships and alternative governance structures devised for governing the transactions. Therefore, the study proposes that incompleteness be modelled endogenously as the result of parties' willingness to minimize the costs of governing transactions. Our literature search for the larger study revealed that there is a paucity of studies covering contractual documentation for inter-firm relationships. Even fewer relate to the parties' choice of the level of contractual completeness, and still fewer are based on sample contracts suitable for econometric testing of the proposals derived from TCE theory. No such studies have been encountered in the UK construction industry context. In this light, we believe that the proposed study described in this paper would represent a contribution to the body of knowledge and understanding of TCE principles and their application to inter-firm contractual

relationships in the UK construction industry. It should also produce valuable documentation preparation information to be fed to the seedbed for claims so that the situation can be improved.

The proposed study also demonstrates deployment of a uniquely comprehensive methodology for measuring Documentation Incompleteness by rating entire documentation packages. This represents simultaneous estimation of all contract provisions, which would enable automatic consideration the possible linkages between all the provisions in the documentation; and allow analysis of linkages between Documentation Incompleteness and other contract dimensions, such as contract duration. Carrying out this study would respond to the research need for empirical studies to estimate complete contract provisions as proposed by Saussier (2000). Therefore, we believe this paper to be a contribution, not only to understanding simultaneous estimation of contract documentation provisions and the law and economics literature, but it also adds specific dimension to analysis of construction industry documentation. This is because studies of this type are generally addressed to the economics audience. By demonstrating application within a construction context, it presents these analytical techniques to an audience that may have been unaware of the potential benefits of alternative approaches to addressing construction documentation concerns.

Notwithstanding the contribution to econometrics and TCE outlined above, we consider that more empirical tests should be developed and refined to permit inclusion of the Frequency transaction characteristic for analysis for general application, and in a construction context. Frequency was not included in the two main studies that were used to guide this proposed study. Review of the literature (Furlotti, 2007) and discussions with other academics revealed that there is still a lack of confidence in interpreting Frequency, as a driver of transaction costs. There have been observations of increases in transaction costs in repeat transactions due increased costs of specifying contingencies in response to histories of friction in past relationships (Argyres et al., 2007). At the same time, observations of decreases in transaction costs due to parties learning about each other and of development of trust, have been reported (Mayer and Bercovitz, 2008). We are optimistic that future research will meet this challenge. Hopefully, this paper will provide a stepping stone in this direction.

5 List of Abbreviations

AEO	- A measure of the Ambiguities, Errors or Omissions in the documentation package
ASD	- Ambiguity/Specificity Dimension
BoQs	- Bills of Quantities
BR	- Bounded Rationality
CC	- Complete Contracts
CCC	- Complete Contingent Contracts
CD	- Complexity Dimension
CDPRS	- Contract Documentation Preparers (such as quantity surveyors).
CP	- Contingency Planning Dimension
DCV	- Documentation Completeness Variable
DI	- Documentation Incompleteness
ICE	- Institution of Civil Engineers
ICs	- Incomplete Contracts
ICT	- Incomplete Contracts Theory
JCT	- Joint Contracts Tribunal
JCTSBCs	- Joint Contracts Tribunal Standard Building Contracts
LEXs	- Loss-and-expense claims
MB	- Marginal Benefits
MC	- Marginal Costs
NEC	- New Engineering Contract
OLSRA	- Ordinary Least Squares Regression Analysis
PI	- Production Information
PRI	- Principal Investigator
SEM	- Structural Equation Modelling
SPE	- Asset Specificity measured from dataset documentation
TCE	- Transaction Cost Economics
UNC	- Uncertainty measured from dataset documentation
WPM	- Documentation completeness rating methodology described in Walker and Pryke (2009).

6 References

- Abdul-Malek, M.A.U., El-Saadi, M.M.H., Abou-Zeid, M.G. (2002). 'Process Model for Administrating Construction Claims.' *Journal of Management in Engineering*, **18**(2), 84-94.
- Al-Najjar, N.I. (1995). 'Incomplete Contracts and the Governance of Complex Contractual Relationships.' *American Economic Review*, **85**(2), 432-436.
- Argyres, N.S., Bercovitz, J., Mayer, K.J. (2007). 'Complementarity and Evolution of Contractual Provisions: An Empirical Study of IT Services Contracts.' *Organization Science*, **18**(1), 3-19.
- Bajari, P., Houghton, S. and Tadelis, S. (2007). 'Bidding for Incomplete Contracts: An Empirical Analysis of Adaptation Costs.' National Bureau of Economic Research Working Paper No. 12051
- Ball, M. (1988). *Rebuilding Construction. Economic Change and The British Construction Industry*, Routledge, London, England, UK.
- Baum, J.A.C. and Oliver, C. (1991) 'Institutional Linkages and Organizational Mortality.' *Administrative Science Quarterly*, **36**(2), 187-218.
- Bingham, T. (2009). 'The One and Only: Using the NEC for Public Projects.' *Building*, **274**(25), 47.
- Bolton, P. and Dewatripont, M. (2005). *Contract Theory*, MIT Press, London
- Boyd, D. and Wild, A. (2003). 'The Tavistock Studies into the Building Industry: Communications in the Building Industry (1965) and Interdependence and Uncertainty (1996)', in Murray, Mike and Langford, David (eds.) (2003) *Construction Reports, 1944-1998*, Blackwell Science, Oxford.
- Building Talk Editorial Team (BTET) (2009). 'Government Curbing Contract Choice' *Building Talk*, May 25 [html]
Available from: <http://www.buildingtalk.com/news/joi/joi125.html>, [Accessed 02/07/10]
- Chan, E.H.W. and Henry, C.H.S. (2005). 'Dispute Resolution Management for International Construction Projects in China.' *Management Decision*, **43**(4), 589-601.
- Clamp, H., Cox, S., Lupton, S. (2007). *Which Contract? Choosing the Appropriate Building Contract*, RIBA Publishing, London, 4th Ed.
- Creswell, J. (2009). *Research Design. Qualitative, Quantitative, and Mixed Methods*, Sage Publications, Inc., London, 3rd Ed.
- Crocker, K.J. and Masten, S.E. (1991). 'Pretia Ex Machina? Prices and Process in Long Term Contracts' *Journal of Law and Economics*, **34**(1), 69-99.
- Crocker, K.J. and Reynolds, K.J. (1993). 'The Efficiency of Incomplete Contracts: An Empirical Analysis of Air Force Engine Procurement'. *RAND Journal of Economics*, **24**(1), 126-146.
- Diekmann, J.E. and Kruppenbacher, T. A. (1984). Claims Analysis and Computer Reasoning. *Journal of Construction Engineering and Management*, **110**(4), 391-408.
- Diekmann, J.E. and Nelson, M.C. (1985). 'Construction Claims: Frequency and Severity.' *Journal of Construction Engineering and Management*, **111**(1), 74-81.

- Dougherty, C. (2007). *Introduction to Econometrics*, Oxford University Press, Oxford, 3rd Ed.
- Fellows, R. and Liu, A. (2008). *Research Methods for Construction*, Blackwell Science, London, 3rd Ed.
- Forrester, J. W. (1999) *Industrial Dynamics*, Pegasus Communications, Waltham, Massachusetts.
- Furlotti, M. (2007). 'There is More to Contracts Than Incompleteness: A Review and Assessment of Empirical Research on Inter-firm Contract Design.' *Journal of Management and Governance*, **11**(1), 61-99.
- Goldberg, V.P. (1985). 'Price Adjustment in Long-Term Contracts.' *Wisconsin Law Review*, **1985**(3), 527-543
- Grossman, S.J. and Hart, O. (1986). 'The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration.' *The Journal of Political Economy*, **94**(4), 691-719.
- Gujarati, D. and Porter, D.C. (2009). *Basic Econometrics. International Edition*, McGraw-Hill/Irwin, London, 5th Ed.
- Harris, R.A. (2000). *An Assessment of Construction Contract Claims in the UK*, A Thesis submitted for the degree of Doctor of Philosophy in Engineering at the University of Newcastle-upon-Tyne, Department of Civil Engineering, Newcastle, England.
- Hart, O. and Moore, J. (1988). 'Incomplete Contracts and Renegotiation.' *Econometrica*, **56**(4), 755-785.
- Hart, O. (1995). *Firms, Contracts, and Financial Structure*, Oxford University Press, Oxford.
- Higgin, G. and Jessop, N. (1965). *Communications in the Building Industry: The Report of a Pilot Study*, Tavistock Publications, London.
- Hoxley, Michael (2000). 'Measuring UK Construction Professional Service Quality: The What, How, When and Who!' *International Journal of Quality and Reliability Management*, **17**(4/5), 511-526.
- Hughes, G.A. and Barber, J.N. (1992). *Building and Civil Engineering Claims in Perspective*. Longman Scientific and Technical, Essex, 3rd Ed.
- Itzer, D. and Dikbas, A. (2009). 'An Investigation of the Factors Influencing Dispute Frequency in Construction Projects' in *RICS COBRA Research Conference*, University of Cape Town, 10-11 September, 2009, RICS London.
- Jannadia, M.O., Assaf, S., Bubshait, A.A., Naji, A. (2000). 'Contractual Methods for Dispute Avoidance and Resolution (DAR)'. *International Journal of Project Management*, **18**(1), 41-49.
- Kennedy, P. and Milligan, J.L. (2005). *Research Analysis of the Progress of Adjudication Based on Returned Questionnaires from Adjudicator Nominating Bodies (ANBs) and on Questionnaires Returned by Adjudicators*, The Adjudication Reporting Centre, Glasgow Caledonian University.
- Klein, B., Crawford, R.G., and Alchian, A.A. (1978). 'Vertical Integration, Appropriable Rents, and the Competitive Contracting Process.' *Journal of Law and Economics*, **21**(2), 297-326.
- Klein, Rudi (2009). 'All For One: The OGC's Contract Decision.' *Building*, **274**(29), 45-48
- Kumaraswamy, M.M. (1997). 'Conflicts, Claims and Dispute in Construction'. *Engineering, Construction and Architectural Management*, **4**(2), 95-111.

- Latham, Sir Michael (1994). *Constructing the Team. The Latham Report – Joint Review of Procurement and Contractual Arrangements in the United Kingdom Construction Industry*, Her Majesty Stationery Office (HMSO), London, England.
- Love, P.E.D. (2005). 'Contract Documentation and the Incidence of Rework in Projects.' *Architectural Engineering and Design Management*, **1**(4), 247-259.
- Love, P.E.D. (2008). 'Forensic Project Management: An Exploratory Examination of the Causal Behaviour of Design-Induced Rework.' *IEEE Transactions on Engineering Management*, **55**(2), 234-247.
- Love, P.E.D. (2009). 'Project Pathogens: The Anatomy of Omission Errors in Construction and Resource Engineering Project.' *IEEE Transactions on Engineering Management*, **56**(3), 425-435.
- Masten, S.E. (1988). 'Equity, Opportunism and the Design of Contractual Relations. *Journal of Institutional and Theoretical Economics/Zeitschrift für die gesamte Staatswissenschaft*, **144**(1), 180-195.
- Masten, S.E. and Saussier, S. (2000). 'Econometrics of Contracts: An Assessment of the Developments in the Empirical Literature on Contracting.' *Revue d'Économie Industrielle*, **92**(1), 215-236.
- Masten, S.E., Meehan Jr., J.W., and Snyder, E.A. (1991). 'The Costs of Organization.' *Journal of Law, Economics, and Organization*, **7**(1), 1-25.
- Mayer, K.J. and Bercovitz, J. (2008). *The Influence of Inertia on Contract Design: Contingency Planning in Information Technology Service Contracts*, Managerial and Decision Economics, **29**(2-3), 149 – 163.
- McPake, B. and Normand, C. E.M. (2008). *Health Economics: An International Perspective*, Routledge, London, 2nd Ed.
- Morton, R. subsequently revised by Ross, A. (2008). *Construction UK: An Introduction to the Industry*, Blackwell Publishing, Oxford, England, UK.
- Murdoch, J. and Hughes, W. (2008). *Construction Contracts: Law and Management*, Taylor & Francis, Abingdon, 4th Ed.
- Newman, P. and Whitfield, J. (1994). *Loss and Expense Explained*. RIBA Publications, London.
- Ramus, J., Birchall, S., and Griffiths, P. (2008). *Contract Practice for Surveyors*, Butterworth-Heinmann, Oxford, 4th Ed.
- RICS (2010). *Contracts in Use – A Survey of Building Contracts in Use During 2007*, The Royal Institute of Chartered Surveyors, London, England, UK, [html]
Available from:
http://www.rics.org/site/download_feed.aspx?fileID=5853&fileExtension=PDF
[Accessed 02/07/2010]
- Salanié, B. (2005). *The Economics of Contracts: A Primer*, MIT Press, London, 2nd Ed.
- Saussier, S. (2000). 'Transaction Costs and Contractual Incompleteness: The Case of Électricité de France.' *Journal of Economic Behavior and Organization*, **42**(2), 189-206.
- Scott, S. and Harris, R.A. (2004). 'United Kingdom Construction Claims: Views of Professionals.' *Journal of Construction Management and Engineering*, **130**(5), 734-741.
- Seeley, I.H. (1997). *Quantity Surveying Practice*, Macmillian Press, Ltd., London, 2nd Ed.
- Semple, C., Hartman, F.T., and Jergeas, G. (1994). 'Construction Claims and Disputes: Causes and Cost/Time Overruns.' *Journal of Construction Engineering and Management*, **120**(4), 785-795.

- Smith, N.J. (2007). *Engineering Project Management*, Blackwell Science, Oxford, England, 3rd Ed.
- Tilley, P.A., McFallan, S.L., and Tucker, S.N. (1999). 'Design and Documentation Quality and Its Impact on the Construction Process' CIB Number 234 - Customer Satisfaction: A Focus for Research & Practice. *Proceedings from CIB W55 & W65 Joint Triennial Symposium*, September 5-10, 1999, Cape Town, South Africa, pp. 261-269.
- Vidogah, William and Ndekugri, Issaka (1997). 'Improving Management of Claims: Contractors' Perspective'. *Journal of Management in Engineering*, **13**(5), 37-44.
- Vidogah, William and Ndekugri, Issaka (1998). 'Improving the Management of Claims on Construction Contracts: Consultant's Perspective'. *Construction Management and Economics*, **16**(3), 363-372.
- Walker, F. and Pryke, S. (2008). 'Relationship Between Degree of Incompleteness of Contract Documents and Loss-and-Expense Claims in Construction Projects' *Proceedings of the RICS Annual Construction, Building and Real Estate Research Conference (COBRA)*, September 4-5, 2008, Dublin, Ireland.
- Walker, F. and Pryke, S. (2009). 'Role, Dimensions, and Degrees of Incompleteness in Construction Contract Documentation.' *Proceedings of the RICS Annual Construction, Building and Real Estate Research Conference (COBRA)*, September 10-11, 2009, Cape Town, South Africa, pp. 1258-1275.
- Whitfield, J. (1994). *Conflicts in Construction. Avoiding, Managing, Resolving*. Macmillan, Press, Ltd., London.
- Williamson, O.E. (1985). *The Economic Institutions of Capitalism. Firms, Markets, Relational Contracting*, The Free Press, Collier Macmillan Publishers, London.
- Winch, G.M. (2001). 'Governing the Project Process. A Conceptual Framework'. *Construction Management and Economics*, **19**(8), 799-808.
- Winch, G.M. (2010). *Managing Construction Projects. An Information Processing Approach*, Wiley-Blackwell, Sussex, 3rd Ed.
- Yates, D.J. and Hardcastle, C. (2003) 'The Causes of Conflict and Disputes in the Hong Kong Construction Industry. A Transaction Cost Economics Perspective.' *RICS Foundation Research Papers*, Volume 4, Number 22.
- Yogeswaran, K., Kumaraswamy, M.M., and Miller, D.R.A. (1997). 'Perceived Sources and Causes of Construction Claims'. *Journal of Construction Procurement*, **3**(3), 3-26.

Is There a Need for Public Private Partnership Projects in Singapore?

Asanga Gunawansa¹

¹School of Design and Environment,
National University of Singapore, 4 Architecture Drive, 117566
Singapore

Email: bdgasan@nus.edu.sg

Abstract:

The concept of Public-Private Partnership (PPP) was introduced to Singapore in 2003. Although there was much hype about what it could accomplish in Singapore in the early days, so far only nine projects have been successfully procured as PPPs. Notably, five of these projects are in the area of building physical infrastructure facilities. Two of the projects are in the area of Information Technology (IT) infrastructure. The remaining two projects deal with the construction of defence facilities. In recent times, several projects initially considered for development as PPPs have been procured following the more traditional methods such as design & build. Despite this lukewarm development in PPPs in Singapore, this paper argues that PPPs remains a procurement method that is viable for development of public infrastructure facilities. The paper examines the key barriers that have contributed to the slow progress made thus far, and what policy and institutional changes are necessary to make PPP a success story in Singapore. The analysis presented and the conclusions reached in this paper are based on a study of the existing literature and policy architecture.

Keywords:

Public-Private Partnerships, infrastructure development, public procurement, investment, Singapore

1. Introduction

Historically, public infrastructure has mainly been delivered by the public sector, using traditional procurement methods such as, Design-Bid-Build and Design and Build. Further, public finance has been used to award contracts to private sector contractors. The public sector entities awarding the contracts have been in charge of the actual delivery of services to the public once the projects have been developed. Thus, the private sector role was limited to designing and construction of facilities.

As governments started facing the challenge of stretching the scarce public funds to deal with increasing demand for new and modern infrastructure facilities, whilst, also performing the other fiscal duties expected of modern welfare governments, the need to

find alternative mechanisms to finance and develop public infrastructure facilities was recognized. Public-Private Partnership (PPP) under which the public sector entities could partner with private sector entities in developing, managing and providing services to the people became the popular option for many countries.

The concept of PPP was introduced to Singapore in 2003 when the first PPP contract was awarded by the Public Utilities Board (PUB) for a desalination plant. Subsequently, the Ministry of Finance (MOF) issued a handbook in 2004 as a guide to the public and private sector for successful structuring and management of PPP projects in Singapore. According to the said handbook, the main aims of implementing PPPs in Singapore include: (a) allowing the public sector to get better value for money in the delivery of public services; (b) offering the private sector more business opportunities and more room to innovate and offer efficient solutions for public services; and (c) combining the expertise of the government and the private sector to meet the needs of the public effectively and efficiently (MOF, 2004). However, since the introduction of the PPP concept in Singapore, it has had very limited success, with only eight PPPs successfully implemented to date.

In the circumstances, it is important to consider whether PPP is a feasible procurement method to be used in Singapore for development of public infrastructure facilities. Further, it is important to examine the reasons for the limited success PPPs have had in Singapore and propose measures that should be taken to give PPP a new lease of life if it is a viable procurement option for Singapore.

The rest of this paper is organised as follows: Section two deals with the research objectives and the methodology. Section three deals with the concept of PPP. Section four examines the current status of PPPs in Singapore. Section five looks at the barriers to PPPs in Singapore. Section six considers the future of PPPs in Singapore. Section seven concludes.

2. Research Objectives and Methodology

The limited success of PPPs in Singapore was not envisioned both, public sector, and private sector entities as the introduction of the PPP concept was met with much enthusiasm. As commented by Harris (2008):

“With an apparent championing behind this within the Ministry of Finance, and a strong pipeline of planned projects, the stage seemed set for a promising period for developers, bankers and the usual slew of advisers (financial, legal and technical). Three years later and the common view seems to be that things haven't worked out as planned. That's not to say things have gone badly. More that things have gone slowly and, on a few occasions, with some unexpected hiccups in implementation. Let's face it, in the Singapore context, that's unusual.”

Thus, this paper is based on a study undertaken with the specific objectives of examining the reasons for the limited success of PPPs in Singapore as a procurement method and to identify the measures necessary to give PPPs a new lease of life in Singapore. The study involved an extensive literature review and examination of the relevant policy architecture

to examine the objectives and rationale for using PPP as a procurement method, the benefits associated with PPPs and the barriers to PPPs in Singapore. A review of the status of PPP projects in Singapore was also carried out through a survey of existing literature and relevant information available in the public domain.

The literature review was used as the primary research methodology based on the assumption that knowledge accumulates, and that a researcher could learn from and build on what others have done (Webster and Watson, 2002). The literature review has been thus used, to integrate and summarize existing knowledge in the area relevant to the key issue researched in the study and, based on a critical appraisal of the same, to stimulate new ideas and suggests hypotheses for testing and make recommendations. The literature that has been reviewed has been integrated into the different sections and subsections in the paper in order to maintain an uncomplicated flow of the analysis presented.

3. Understanding PPPs

3.1. The Need for PPPs

Even though private investment in public infrastructure can be traced back to the 18th century in European countries (Kumaraswamy and Morris, 2002), According to Harris (2004), the increasing adoption of PPPs by countries in the late 1990s was due to the success of PPPs in the United Kingdom. According to him, it was the development and refinement of private finance initiative (PFI) by the United Kingdom in 1992, as one of a range of government policies designed to increase private sector involvement in the provision of public services, which led to the renewed international interest in PPPs. Since then, many countries around the world have either embark on or considered a PPP programme of their own (Harris, 2004).

The need for private sector involvement in providing infrastructure services to the people arose due to a variety of reasons. From the perspective of developing countries, the key reasons were the lack of public funds, need for modern technology and efficient management skills. From the perspective of developed nations, allocation of project risks was the key reason.

Two alternative mechanisms considered for engagement of private sector were total privatization of public facilities and PPPs (Ford and Zussman, 1997). The former enables governments to transfer the total responsibility of developing, managing, and providing services to the people to the private sector. The later enables governments to invite private sector entities to finance and develop infrastructure projects without losing the state control over the regulatory aspects of service delivery, including, the pricing of the services provided by the infrastructure facility (Savas, 2000; Gunawansa, 2001; Abdul-Aziz, 2007). The total privatization of public infrastructure facilities that have provided services to the public at prices heavily subsidized by the governments was considered politically controversial. Further, the governments were hesitant to subject certain facilities to total privatization due to reasons such as interest of national security. Thus, PPP became the popular option.

The concept of PPP was also adopted by many countries as it provides an opportunity for efficient allocation of project risks whereby risks could be borne by parties who are best able to manage them. Further, procurement of PPPs became popular as they enable the provision of value for money to the public by tapping on the expertise of, and benefitting from the technology and management skills, of the private sector (Algarni et al., 2007; ADB, 2008).

3.2. Definition of PPP

PPP is a procurement method which involves private sector supply of infrastructure assets and services that have traditionally been provided by the public sector. As commented by Khanom (2009), there is no precise and commonly accepted definition of PPP.

According to the dictionary definition in Webster, a partner is “one of two or more persons contractually associated as joint principals in business.” The same dictionary defines “partnership” as a legal relation existing between two or more persons contractually associated as joint principals in a business. According to Duhaime’s Legal Dictionary “partnership” is an organization in which two or more persons carry on a business together. Thus it is not difficult to conclude that a public-private partnership is a contractual arrangement in which a public sector entity and a private sector entity come together to do a business.

However, the difficulty in defining PPP as noted by Khanom (2009) arises as a result of the diverse interests and objectives the public and private parties have in entering into PPPs. For the public sector, as noted earlier, the need to enter in to PPPs may arise as a result of one or more of the following reasons: lack of finance, need for modern technology, effective and efficient management skills, and the need to transfer risk. For the private sector entities, PPPs offer new investment opportunities, new markets and the opportunity partner with public sector who has in the past enjoyed a monopoly in provision of certain infrastructure facilities (Gunawansa, 2000).

The needs of public sector as well as private sector entities to enter into PPPs can differ from project to project and jurisdiction to jurisdiction. This is another reason for the absence of a common definition of PPP. For example, the need of a cash strapped developing country to enter into a PPP to develop a project to provide clean water or electricity to the citizens will be different from the need a developed country may have in considering a PPP to develop an airport or a highway.

One definition of PPP as embraced by The Canadian Council for PPP is as follows:

“A cooperative venture between the public and private sectors, built on the expertise of each partner, that best meets clearly defined public needs through the appropriate allocation of resources, risks and rewards.”

In UK, Her Majesty’s Treasury (1998) defines PPP as:

“An arrangement between two or more entities that enables them to do public service work cooperatively towards shared or compatible objectives and in which

there is some degree of shared authority and responsibility, joint investment of resources, shared risk taking and mutual benefit”

In Singapore, the MOF (2004) has defined PPP as:

“PPP refers to long-term partnering relationships between the public and private sector to deliver services. It is a new approach that Government is adopting to increase private sector involvement in the delivery of public services.”

In India, the Department of Economic Affairs of the Ministry of Finance (2005) defines PPP as:

“The Public-Private Partnership (PPP) Project means a project based on contract or concession agreement between a Government or statutory entity on the one side and a private sector company on the other side, for delivering an infrastructure service on payment of user charges.”

The Canadian definition focuses on the cooperative venture between the public and private parties and the appropriate allocation of resources and risks. This gives the indication that PPPs are looked at as partnering arrangements between parties with equal bargaining power. Similarly, the UK definition focuses on compatibility between the parties and sharing of responsibilities, risks, resources, and profits.

The Singapore definition focuses on PPPs as a long term relationship between public and private sectors which enables the public sector to involve private sector in providing services to the people. This definition does not give any indication as to the real need for the public sector to enter into PPPs. Further, in Singapore, PPP is also seen as a way to bring in specialist private sector expertise to stimulate an exchange of ideas and bring more international players into the domestic market (KPMG, 2007).

The Indian definition focuses on the fact that the government gives a concession to the private sector to develop a project and provide services in return for payment of user charges. This gives the indication that the public sector engagement in the partnership is limited to the granting of the concession due to financial constraints and lack of modern technology. The private sector is required to finance and develop the project and offer services in return for payments for services provided.

3.3 Main Characteristics of PPP

Under PPP, usually, a public entity would typically specify the outputs or services required, and a private company or consortium would be responsible for the finance, design, construction, operation and maintenance of a facility. The consortium is typically organized by a project developer who brings together financiers, engineering firms, construction companies and facilities management companies to provide individual services. A typical PPP structure is shown in Figure 3.1. below:

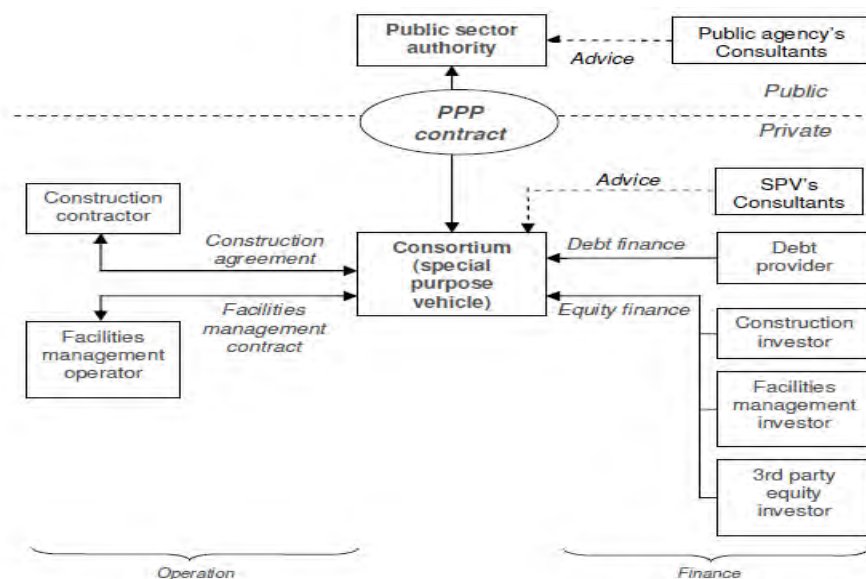


Figure 2.1: PPP Structure (MOF, 2004)

The PPP contract between the public agency and the consortium would usually be for a period of 10 to 30 years, and unlike traditional procurement methods, the public sector does not own the facility during this time period. A PPP thus allows the public sector to move away from directly owning and operating facilities, to purchasing services directly from the private consortium (MOF, 2004). However, the public sector may continue to deliver the core services traditionally associated with a facility (such as teaching in schools and medical services in hospitals) while the consortium may deliver the ancillary services which support the infrastructure (Department of Treasury and Finance, 2002). It is often only after the contract period that the facility returns to public ownership.

Unlike in conventional projects, in a PPP, the consortium would only be able to recover its investment through income earned by operating the facility. As stated by KPMG (2007), the public sector may compensate the consortium with service payments, rights to levy tariffs or fees against the public users, or a combination of these. For a project that produces a public utility service, an off-take contract may be signed between the consortium and the public agency, whereby the public agency agrees to purchase the output of the facility at an agreed price and volume on a long-term basis. This off-take contract serves as the basis for the financing of the project (Yescombe, 2002).

The PPP payment mechanism typically provides the government with the power to withhold or deduct payments if the quality of service provided by the private sector consortium is lower than agreed. The government may also reserve the right to step in and regain control of the asset, in the event of repeated default in service provision by the private sector operator.

3.4 PPP Models

PPPs are implemented for a wide range of social and economic infrastructure projects. However, they are mainly used to build and operate hospitals, schools, prisons, roads,

bridges and tunnels, light rail networks, and water and sanitation plants (IMF, 2004). For its application, the project may take the form of one of the PPP models such as joint-ventures, strategic partnerships between the public and private sector, Design-Build-Operate (DBO), Design-Build-Finance-Operate (DBFO), Build-Operate-Transfer (BOT), Build-Transfer-Operate (BTO), Build-Own-Operate (BOO) and many other variants (MOF, 2004; IMF, 2004). As each PPP project is unique, the different PPP models can be flexibly selected and tailored according to the project (Gunawansa, 2000). The public sector and private sectors have to work closely together to determine the optimal scope of collaboration in each PPP project to the benefit of members of the public who use the services, the government and the private sector (MOF, 2004).

4. Current Status of PPPs in Singapore

4.1 Rationale for Introducing PPP

PPP was first introduced in Singapore in 2003 under the Best Sourcing Framework whereby the public sector will engage private sector providers to deliver non-core government services which the private sector can provide more effectively and efficiently (MOF, 2004). Singapore's interest in PPP was set out in a consultation document and a subsequent "PPP Handbook" that was published in October 2004 by MOF¹. This handbook provides a general guidance on PPP procurement, and dictates that all government infrastructure projects in excess of S\$50 million be actively considered for suitability as PPPs. As stated in the handbook, a number of sectors in Singapore have been identified by MOF as suitable for PPPs. These includes sports facilities, incineration plants, water and sewerage treatment works, large IT infrastructure deals, education and healthcare facilities, expressways and government buildings.

According to the Department of Treasury and Finance (2002), by using the joint skills of the public and private sectors, Singapore will be able to:

- Potentially create new infrastructure which is of a standard beyond that which could be delivered by the public sector alone;
- Support the infrastructure with guaranteed services to ensure its continued usefulness, efficiency and longevity;
- Take advantage of innovative ideas and technology which have traditionally been fostered in commercial environments, for the benefit of users of public infrastructure; and

¹ In addition to the handbook, the MOF created a PPP Advisory Council whose aim is to create awareness of PPP, draft PPP policy and provide guidance on PPP matters. The Council also oversees the progress of major PPP projects and facilitates resolution of inter-agency issues.

- More effectively manage the risks that naturally come with very large and complex infrastructure projects by allocating risks to the party which is best able to manage them.

However, as stated in a report by KPMG (2007):

“Unlike many countries undertaking PPP, the government of Singapore does not need private funds to improve its social and other infrastructure. It has large reserves and typically a budget surplus. Nor would the government necessarily concede that private sector provision of goods and services is more efficient than that of the public sector. The need or ability to raise capital is a less pressing concern in Singapore than it might be in some parts of the region.”

Thus, the rationale for introducing PPP into Singapore is mainly focused on the need to achieve value for money in the delivery of public services as stated in the PPP Handbook (Lam, 2004). This means that PPP was perceived as a mechanism that could be used for projects to allow optimal balance of benefits and costs on the basis of total cost of ownership, even though it may not be at the lowest price. This could be achieved when synergies are generated through the alignment of design, construction, maintenance and operation phases, taking into consideration the whole life cycle of the project (MOF, 2009).

4.2 PPP Projects Implemented in Singapore

Since the introduction of PPP in 2003, the Singapore government has explored various projects to be implemented as PPPs. Some have met with success, whilst the others have been abandoned, awarded under traditional procurement methods, or still under consideration. A summary of these projects and their status are presented in the Table below:

PPP Projects in Singapore as at 27 October 2009.

(Sources: MOF; PUB, 2007; CNA, 2009; Lazauskaite, 2009; Loh, 2007; Wang, 2009; Dalal, 2009; IFR Asia, 2009)

No	Project Description	Launched	Public Agency in Charge	Project Status
1	Desalination Plant - Design-Build-Own-Operate (DBOO) facility to supply desalinated water to the Public Utilities Board (PUB) for 20 years	September 2001	PUB	Tender awarded on 19 Jan 2003. Project completed and in operation since September 2005.
2	Ulu Pandan NEWater Plant – DBOO project for recycling and supply of waste water to PUB for 20 years	May 2004	PUB	Tender awarded on 15 Dec 2004. Project completed and in operation since March 2007.
3	Fifth Incineration Plant - DBOO project to incinerate 800 tones of refuse per day for a period of 25 years	May 2005	National Environment Agency (NEA)	Tender awarded – 14 Nov 2005. Project completed and in operation since the beginning of

				2010.
4	TradeXchange – IT PPP project to develop the software for one-stop integrated logistics information port, including the maintenance and operation of the system for 10 years.	Dec 2005	Singapore Customs	Contract awarded on 8 Dec 2005. Tenure of 10 years between 2007-2017
5	Basic Wings Course – a 20 year PPP to acquire and maintain a fleet of trainer aircraft and ground-based training systems to meet the hardware requirements for the Republic of Singapore Air Force's basic wings flying training.	Aug 2005	Ministry of Defence (MINDEF) — RSAF	Tender awarded / Financial close achieved on 3 Nov 2006. Tenure of 20 years from 2008-2028.
6	Rotary Winged Course – a 20 year PPP to leverage on commercially available platforms and maintenance operations to optimise the resource and risk allocation between public sector and private sector in flying training operations.	Nov 2005	MINDEF –RSAF	Tender awarded in November 2005 for duration of 20 years from 2006-2026
7	Next Generation National Broadband Network - A PPP project for the design and construction of the passive infrastructure of the Next Generation National Infocomm Infrastructure which seeks to transform Singapore into an intelligent nation and global city, powered by infocomm. The project also involves the sale of services to an operating company.	February 2006	Info-comm Development Authority (IDA)	Passive Infrastructure Tender awarded in October 2008. Active Infrastructure: Tender awarded in April 2009. Estimated completion date in 2012.
8	ITE College West - Design-Build-Finance-Operate (DBFO) project for a period of 27 years Under the PPP, the consortium would which includes the management and coordination of all building-related and estate management matters, allowing the ITE to focus its resources and attention on the delivery vocational and technical education.	July 2006	Institute of Technical Education (ITE)	Tender awarded / Financial close achieved on 11 Aug 2008. Construction phase of the project is expected to be completed in 2010.
9	Changi NEWater Plant - A DBOO project for the supply of 50 million gallons of NEWater per day to PUB over a period of 25 years.	Aug 2007	PUB	Tender awarded in early 2008. In first phase of commercial operation since 2009. To be fully completed in

				mid-2010.
10	Sports Hub – A 25 DBFO concession for a sports hub.	December 2005	Singapore Sports Council (SSC)	At preferred bidder stage since 19 th January 2008.
11	NUS University Town – PPP for the development of a 6,200-bed student housing with ancillary facilities.	June 2007	National University of Singapore (NUS)	Consortiums were shortlisted in June 2007. Decision to launch the project as a PPP was subsequently terminated in September 2007.
12	Advanced Fighter Winged Course - PPP for providing classroom teaching and aircraft simulators for Advance Fighter Winged Course.	December 2008	Defence, Science & Technology Agency (DSTA)	Bids submitted for project in August 2009. Current status unknown.
13	Rifles Range Training Facility – A 20 year PPP for the construction, maintenance and management of a training ground for small arms.	Unknown	DSTA	Bids submitted for projects in late August 2009. Currently in tender evaluation stage.

As can be seen from Table 2, the Singapore government has considered 13 PPP projects in various asset classes, including water, solid waste, education, defence and IT sector, since the introduction of the concept in 2004. Nine of these projects have been successfully implemented (items 1 – 9 of Table 2). Five of these projects (Desalination Plant, NEWater Plants, Incineration Plant, ITE College West) involve building infrastructure. The projects involving the water and solid waste sector are utility related PPPs. These have been procured by two agencies, National Environment Agency and Public Utilities Board. These projects consist of a single demand stream model with little political sensitivity (KMPG, 2007), thus their procurement should not be considered complicated.

Singapore has also sought to procure large-scale social infrastructure projects such as the ITE College West, Sports Hub, and the University Town @ Warren. Such projects involve multiple revenue streams and numerous stakeholders, with greater scrutiny of business plans (KMPG, 2007), thus making them more complicated. The implementations of these projects except for the ITE College West project have not been successful. It should be noted that even in the case of the ITE College West, the procurement process had met with delays in financing as the banks which initially backed the private sector consortium had dropped out as a result of the subprime crisis (Lee & Rathbone, 2008; Menon, 2008).

The University Town project was initially earmarked to be financed, designed, constructed and operated by private sector entity for a period of 25 to 30 years. However, according to a government press release in January 2008 (MCYS, 2008), it is now owned and funded by the National University of Singapore (NUS), through government grants. This US\$423-million (SGD \$500-\$600 million) project is expected to be completed before the end of 2010. No clear reason has been disclosed for the shift. However, it is likely that being a social infrastructure development project, the project proponents found it not viable for

development as a PPP, given that administration of the university facilities require active participation of the public sector, whilst, the revenue for the private sector developer has to be sourced from student fees and other sources of revenue to be generated by the facilities within the University Town. Being an educational institution controlled by the public sector, the private developer's freedom to price the services would have to be highly regulated, a condition that the private sector developer may not have found favourable.

As far as the Sports Hub project which was launched in 2005 is concerned, to date it has not advanced from the preferred bidder stage. The completion date for the project has repeatedly been pushed back. The current expectation is that the project would be completed by the end of 2013 or early 2014, provided that there are no further delays (Wang, 2009). The first delay of the project had occurred due to the addition of a public water sports centre into the original bid requirements, resulting in the process of the submission of the bid proposal and the evaluation period being postponed by about a year (Lim, 2007). The second delay had occurred as a result of increased construction cost and the recent global financial crisis which made it difficult to raise funds from financial institutions (Straits Times, 2009).

5. Barriers to PPPs in Singapore

Even though there has been considerable interest in the use of PPP as a procurement method, there are a wide range of barriers to using PPPs. As found by Zhang (2005) in a questionnaire survey to different organizations in different countries such as United Kingdom, Australia, and China, the barriers to PPPs can be broadly classified into six categories:

1. Social, Political, and Legal Risks
2. Unfavorable Economic and Commercial Conditions
3. Inefficient Public Procurement Framework
4. Lack of Mature Financial Engineering Techniques
5. Problems related to the Public Sector
6. Problems related to the Private Sector

It is important to note that the above barriers are not generic to every country. In the case of Singapore, it is not difficult to argue that social, political and the legal risks which usually discourage investors from investing in countries are not present in Singapore (IMF, 2004). Since independence in 1965, Singapore has steadily developed its social, political and legal infrastructure to offer one of the politically steadiest, corruption free and investment friendly environments in the world (Lim and Lloyd, 1986; Peebles and Wilson; 2002). Furthermore, according to some sources, currently Singapore is ranked as the fourth

richest country in the world², behind Qatar, Luxemburg and Norway (Global Finance, 2010; IMF 2010a). Moreover, according to the World Economic Forum's Global Competitiveness Report for 2009 – 2010 (WEF, 2010), Singapore is ranked the third most competitive economy in the world. It provides inter alia that:

“Its institutions are viewed as the best in the world, while business confidence in the government remains strong despite the global recession. That, along with a highly skilled workforce and sophisticated financial markets, helped Singapore jump two places in the 2009 rankings from last year's result.”

In the circumstances, it is also difficult to argue that the categories 1, 2 and 4 listed by Zhang (2005) apply in the case of Singapore as barriers to PPPs.

As far as the public procurement framework is concerned, it should be stated that Singapore has a well knit legal framework that provides a sound architecture for efficient and corruption free public procurement. For example, the Singapore Government is bound by certain commitments in the various international agreements on public procurement which the country has entered into, specifically, the 1994 Agreement on Government Procurement under the World Trade Organisation ("GPA"), the Agreement between New Zealand and Singapore on a Closer Economic Partnership (ANZSCEP), the Agreement between Japan and Singapore for a New Age Economic Partnership (JSEPA), EFTA-Singapore Free Trade Agreement (ESFTA) and Singapore-Australia Free Trade Agreement (SAFTA), and the United States-Singapore Free Trade Agreement (USSFTA), to name some.

In addition, as far as national legislation is concerned, Singapore has enacted the Government Procurement Act No. 14 of 1997 as amended (Chapter 120) which deals with public procurement. There are also four relevant subsidiary legislations, namely Government Procurement (Challenge Proceedings) Regulations, Government Procurement Regulations, Government Procurement (Application) Order and Government Procurement Act (commencement) Notification 2002. In addition, there is a Government Instructions Manual on procurement procedures and this is publicly available³.

All public contracts awarded by the Singapore government are published on the Government e-Business website on the internet. The award notice contains the name of the successful tenderer, contract sum, and a description of the contract along with the name and address of the awarding government procuring entity. A similar disclosure is made by the Ministry of Finance concerning PPP projects⁴. According to the self assessment made by Singapore in response to the ADB/OECD Anti-Corruption Initiative for Asia and the

² There are two standard methods of defining the richest countries in the world. One takes into account the economies that are the largest, as measured by total gross domestic product (GDP). However, the most commonly accepted definition of the wealthiest countries is to determine how rich the average resident of a country is. For this reason, the best method is to use GDP data per capita. The ranking of Singapore is based on this method.

³ See: <http://www.gebiz.gov.sg>.

⁴ See: <http://app.mof.gov.sg/ppp.aspx>

Pacific (ADB/OECD, 2006), normally, the government procuring entity will at the request of unsuccessful renderers, promptly explain reasons why their bids have not been accepted. Thus, transparency in the process is maintained.

As commented by Mr. Kamran Khan, head of the World Bank Office Singapore (East Asia Infrastructure Finance Practice Group), in an interview by The Business Times Singapore in November 2009:

“The public entities in Singapore are one of the most efficient in the world. They are professional and act almost like the private sector, but they protect the interest of the public sector, which is a very unique thing.”

In the circumstances, corruption in public procurement as a barrier to PPPs in Singapore could be ruled out.

This leaves us to consider whether there are any problems with the public or private sector that works as barriers to implementing PPPs in Singapore. Typically, problems with public sector in successfully implementing PPP projects would arise due to unfamiliarity with the PPP mechanism, reluctance to share responsibilities with the private sector, problems relating to regulations, lack of commitment towards investor protection, and hesitancy to share risks (Gunawansa, 2000; Zhang, 2005). On the other hand, problems with the private sector in implementing PPPs would arise due to lack of financial and technological capacity, reluctance to work with the public sector, reluctance to deal directly with the end-users of utilities and other infrastructure services, inability to be competitive in delivering services to the public, and incapacity to work in a consortium (Gunawansa, 2000; Zhang, 2005).

However, given the success story of Singapore’s economic development and market competitiveness as well as the efficient and transparent public procurement in place which have been summarized above, it is not difficult to argue that there are no obvious problems with the public sector or the private sector in Singapore that prevents PPPs. If so, then what reasons, if any, are there for the limited success of the concept in Singapore since its introduction in 2004? In answering this question, the following hypotheses are proposed:

1. The public sector entities in Singapore such as the PUB, which is in charge of power and water utilities, Land Transport Authority, which is in charge of the road network, and Housing Development Board, which is in charge of the public housing, to name a few, have been providing efficient services to the public. They have the regulatory and management capacity to take charge of the facilities. They have the necessary funds for development of infrastructure through direct contracting. Thus, the need for PPPs is limited.
2. Singapore government is committed to finance and encourage research and development in various spheres, including in the area of physical and social infrastructure development in the country. As a result, technologies relevant for sustainable development of the nation are financed by the public sector and developed largely in public sector entities such as state universities and

research centres set up within universities. For example, the Singapore Economic Development Board (EDB.) has announced that it will spend about S\$680 million (US\$483.3 million) to build a clean technology ecosystem over the next five years as part of the country's plan to become a global research and development hub. EDB has already allocated S\$350 million Singapore dollars of the total investment to develop the country's clean energy sector, with a focus on solar energy. The remaining S\$330 million will go toward developing water and waste management solutions. This initiative is part of the government's plan announced in September 2009 to invest S\$1 billion to build a greener and more energy efficient country by 2015 (EcoSeed, 2010). In addition, the Inter-Ministerial Committee for Sustainable Development in Singapore has earmarked about S\$700 million to develop the rest of the clean technology sector and S\$3.4 billion to boost the country's economic output along with the creation of 18,000 green jobs by 2015 (EcoSeed, 2010). Thus, public entities in Singapore which have the money and invests in the advancement of technology might find direct contracting with contractors to build public facilities more appropriate than PPPs.

3. The absence of a centralized body within the MOF or outside of the MOF to champion the cause of PPP has also worked as a barrier. The PPP Advisory Council that was set up at the time PPP was introduced into Singapore has functioned largely as an entity to promote PPP awareness, help draft relevant policy and provide guidance on PPP matters. What is required is to give it more teeth or replace it with an agency that could act as the "one stop shop" between public and private sector entities to facilitate PPPs.
4. Inadequate focus on project identification and feasibility studies is an apparent weakness in the PPP procurement machinery in Singapore. The number of projects that have been unsuccessfully procured which were discussed above supports this point.
5. The failure to maintain a published pipeline of projects has also worked as a barrier, limiting the scope for public-private engagement to develop projects. If a centralized public sector entity could work with the other government agencies and identify sectors for PPPs and lineup possible future projects, then it would provide a more efficient platform for interested private sector entities to engage with the public sector concerning the development of such projects.
6. However efficient and transparent the current procurement mechanism is, use of open tenders may not always be the best procurement practice for PPP. This is because such procurement practice limits the scope of private sector participation to projects specifically identified by the government. If sectors are identified and a pipeline of possible projects are maintained, that would enable the interested private sector entities to submit early proposals to the

government to develop the pipelined projects. This would enable the government to take a decision whether to engage in direct negotiations with the private party that has shown the early interest or to call for competitive tenders.

6. Future of PPP in Singapore

As noted earlier, Singapore offers one of the most reliable and competitive environments for investment. The country has an efficient and corruption free public sector. Further, the people in Singapore have the money and the desire to pay for efficient public services. Furthermore, the local private sector has the capacity to provide the necessary local partnership to foreign investors or, to take the lead in investment projects in the country. In the circumstances, there is no reason for the future of PPP in Singapore to be bleak, if the issues identified above are adequately addressed.

The PPP Advisory Council that has been already established should be empowered to coordinate with the various government agencies to identify sectors in which PPPs could be introduced. It should also take a more active role in promoting guidelines and framework on the implementation of PPP projects as well as promoting PPPs to the different agencies. In addition, the PPP Advisory Council should also act as a repository of knowledge for all projects that have been either considered or implemented as PPPs. In the alternative, the government should consider establishing a new agency that could perform the functions discussed above.

Establishment of a “one stop shop” that could facilitate the engagement of public and private sector entities should be considered. Such an agency should, in addition to giving publicity to the pipeline of projects and sectors identified for PPP, have the capacity to facilitate investor engagement by cutting red tape and enabling the acquisition of relevant approvals and permits provided that such process remains legal.

As far as identification of suitable sectors for PPP is concerned, Singapore should concentrate on sectors where there is still room for its financially strong, technologically savvy, and efficient public sector entities to partner with the private sector. Construction of sustainable cities, an area in which Singapore has shown a keen interest and has a good reputation, is an area that has much scope for such partnering. This is because building environmentally sustainable cities require not only the financial capacity to build, but also the technological knowhow in many spheres such as roads, water, electricity, telecommunication, public housing, industrial and office complexes, and social infrastructure facilities. Such projects require expertise in designing and engineering and participation of manufacturers of environmentally friendly construction materials and equipment. Piecing together a sustainable city by entering into thousands of individual contracts would be impractical and difficult compared to a well knitted project development structure in which a few PPPs can work towards developing the project.

7. Conclusion

PPP is a feasible procurement method to be implemented in Singapore. As has been pointed out in this paper, Singapore has a very conducive environment for PPP projects. However, due to the barriers that have been identified in this paper, too few projects have been implemented in Singapore since the introduction of the PPP concept in 2004. This said, there is room for improvement. If the government addresses the concerns discussed and implements measures to deal with them, including the recommendations that have been made here, PPPs will have an important role to play in the continuing development of Singapore. However, if the government fails to take adequate measures, there is the risk of flight of capital, experts and expertise away from Singapore to other countries in the region. Countries such as China, India and Vietnam offers larger markets to investors, although they may not yet have the efficient and reliable investment environment Singapore offers.

References

- Abdel-Aziz, A.M. (2007), 'Successful Delivery of Public-Private Partnerships for Infrastructure Development' *Journal of Construction Engineering and Management*, 133(12), pp 918-931.
- Algarni, A. M., Arditi, D., & Polat, G. (2007), 'Build-Operate-Transfer in Infrastructure Projects in the United States', *Journal of Construction Engineering and Management*, 133(10), pp 728-735.
- Asian Development Bank (ADB). (2008), *Public-Private Partnership Handbook*, ADB, Manila, pp 12-15.
- Asian Development Bank (ADB)/Organisation for Economic Co-operation and Development (OECD). (2006), *Anti-Corruption Policies in Asia and the Pacific: Progress in Legal and Institutional Reform in 25 Countries*, ADB, Manila.
- ChannelNewsAsia (CNA). (2009), 'Singapore's Largest NEWater Plant Begins Commercial Operations', CNA, www.channelnewsasia.com/stories/singaporelocalnews/view/446564/1/.html, viewed: 17/01/2010.
- Dalal, M. (2009), 'Sports Hub PPP makes Second Attempt to Cross Finish Line', *International Financing Review Asia*, 627, 28 November to 11 January 2009.
- Department of Treasury and Finance. (2002), 'Partnerships for Growth: Policies and Guidelines for Public Private Partnerships in Western Australia', *Government of Western Australia*, www.dtf.wa.gov.au/cms/uploadedFiles/pub_priv_partners.pdf, viewed: 29/01/2010.

- Department of Economic Affairs, Ministry of Finance, India. (2005), 'Scheme for Support to Public Private Partnership in Infrastructure', <http://www.pppinindia.com/pdf/PPPGuidelines.pdf>, viewed: 22/05/2010.
- EcoSeed (2010), *Singapore Ramps Up Clean Energy R&D Investment*, <http://www.ecoseed.org/en/general-green-news/green-business-news/green-business-news/7244-Singapore-ramps-up-clean-energy-R-D-investment>, viewed: 30/05/2010.
- Ford, R., and Zussman, D. (1997), *Alternative Service Delivery: Sharing Governance in Canada*, Institutes of Public Administration of Canada (IPAC), Toronto.
- Global Finance. (2010), 'The Richest Countries in the World', <http://www.gfmag.com/tools/global-database/economic-data/10049-the-richest-countries-in-the-world.html>, viewed: 01/06/2010.
- Gunawansa, A. (2000), 'Legal Implications Concerning Project Financing Initiatives in Developing Countries', *Attorney General's Law Review*, July 2000.
- Harris, J., (2008), 'Project Finance 2008: Singapore Overview', *Expert Guides*, www.expertguides.com/default.asp?Page=10&GuideID=178&CountryID=21, viewed: 11/05/2010.
- Harris, S. (2004), 'Public Private Partnerships: Delivering Better Infrastructure Services', *Working Paper*, Inter-American Development Bank, Washington, D.C.
- HM Treasury, United Kingdom. (1998), *Partnerships for Prosperity: The Private Finance Initiative*, HM Treasury, London.
- Menon, M. (2008), 'Singapore Braces for PPP Financing Boom', *International Financing Review Asia*, 565, 23 August to 5 September.
- International Monetary Fund (IMF). (2004), 'Public-Private Partnerships', *IMF*, <http://www.imf.org/external/np/fad/2004/pifp/eng/031204.htm>, viewed: 23/05/2010.
- International Monetary Fund (IMF). (2004a), 'Singapore: Financial System Stability Assessment Including Reports on the Observance of Standards and Codes on the following Topics: Banking Supervision, Insurance Regulation, Securities Regulation, Payment and Settlement Systems, Monetary and Financial Policy Transparency, and Anti-Money Laundering', *IMF Country Report No. 04/104*, IMF, Washington D.C.
- IMF. (2010), *World Economic Outlook Database*, April 2010, <http://www.imf.org/external/pubs/ft/weo/2010/01/weodata/index.aspx>, viewed: 01/06/2010.
- Khanom, N.A. (2009), 'Conceptual Issues in Defining Public Private Partnership', *Paper presented at the Asian Business Research Conference 2009*, Dhaka, Bangladesh.

- KPMG. (2007), 'Building for Prosperity: Exploring the Prospects for Public Private Partnerships in Asia Pacific', *KPMG*, www.kpmg.com.sg/publications/Industries_PPPinAsia2007.pdf, viewed: 02/06/2010.
- Kumaraswamy, M.M., & Morris, D.A. (2002), 'Build-Operate-Transfer-Type Procurement in Asian Megaprojects', *Journal of Construction Engineering and Management*, 128(2), pp 93-102.
- Lazauskaite, V. (2009), 'Developments of Next Generation Networks (NGN): Country Case Studies', *International Telecommunication Union (ITU)*, www.itu.int/ITU-D/treg/Documentation/ITU-NGN09.pdf, viewed: 29/05/2010.
- Lee, K., and Rathbone, M. (2008), 'PPP Goes to Singapore', *Project Finance International*, Issue 382, 2 April 2008.
- Lim, M. (2007), 'Bigger Sports Hub to Include Water Sports Centre', *The Straits Times*, 14/06/2010.
- Lim, C., and Lloyd, P. (1986), *Singapore's Resources and Growth*, Oxford University Press, pp 183-219.
- Loh, C.K. (2007), 'The Public-Private Way for ITE West Campus'. *TODAY*, 27/11/2007.
- Ministry of Finance (MOF), Singapore. (2004), *Public Private Partnership Handbook Version 1*, MOF, <http://app.mof.gov.sg/ppp.aspx>, viewed: 01/06/2010.
- Ministry of Community Development, Youth and Sports, Singapore (MCYS). (2008), 'Singapore's Prime Minister Kicks Off Construction Work For University Town Where Youth Olympic Village Will Be Located', Press Release of 31/01/2008, MCYS, <http://app1.mcys.gov.sg/PressRoom/SingaporesPrimeMinisterKicksOffConstruction.aspx> viewed: 20/05/2010
- MOF, Singapore. (2009), 'Government Procurement', *MOF*, http://app.mof.gov.sg/government_procurement.aspx, viewed: 01/06/2010.
- Peebles, G., and Wilson, P., (2002), *Economic Growth and Development in Singapore: Past and Future*, Cheltenham, UK: Edward Elgar
- Public Utilities Board (PUB). (2007), 'PUB Calls Pubic Tender for Singapore's Fifth NEWater Plant', press Release, PUB, www.pub.gov.sg/mpublications/Pages/PressReleases.aspx?ItemId=58, viewed: 11/05/2010.
- Savas, E. (2000): *Privatisation and Public-Private Partnerships*, Chatham House Publishers, New York.
- The Business Times. (2009), '9 Singapore Government Agencies Tapped for Urban Hub: Bigger Role Seen for Singapore in Growth of Developing Countries', *The Business Times*, 13/11/2009.

- The Straits Times. (2008), 'Singapore Largest NEWater Plant to Start Construction in April', *The Straits Times*, 28 February 2008.
- Wang, J. (2009), 'Sports Hub Project Hit by Lack of Money', *The Straits Times*, 24 April 2009.
- Wang, M.M. (2009), 'Sports Hub: Vivian asks for patience', *The Sunday Times*, , 22 November 2009.
- Webster, J. and Watson, R.T. (2002), 'Analyzing the Past to Prepare for the Future: Writing a Literature Review', *MIS Quarterly*, 26(2), pp 13-23.
- World Economic Forum (WEF). (2010), *Global Competitiveness Report for 2009 – 2010*, <http://www.weforum.org/pdf/GCR09/GCR20092010fullreport.pdf>, viewed:30/05/2010.
- Yescombe, E.R. (2002), *Principles of Project Finance*, Academic Press, USA, pp 8 – 15.
- Zhang, X.Q. (2005), 'Paving the Way for Public-Private Partnerships in Infrastructure Development', *Journal of Construction Engineering and Management*, 131(1), pp 71-80.

A brief comparative investigation into the regulatory requirements for production of design information in the UK and USA in the context of building information modelling

Brodie McAdam

School of the Built Environment,
University of Salford, Salford, M5 4WT,
United Kingdom

Email: w.b.mcadam@salford.ac.uk

Abstract:

This paper seeks an academic basis for the postulation that in the UK, unlike in the USA, there is no public authority requirement to have one design professional take responsibility for the design of a project, and hence that particular regulatory problem for BIM enabled procurement did not exist in the UK. Aspects of the regulatory systems in the USA and UK are explored and compared by means of a critical literature review. It is concluded that there is an academic basis for the postulation, but that much further research needs to be done into the regulatory aspects of BIM enabled procurement.

Keywords:

Building information modelling; construction law; building regulations.

1 Introduction

The purpose of this brief paper is to address a particular gap in knowledge identified from an earlier paper (McAdam, 2010). That paper constituted a preliminary literature review which identified that procuring structures by exploiting the full collaborative potential of building information modelling (“BIM”) software raises a number of legal, commercial and technical challenges which are not associated at all, or to such a large extent, with other types of design approach.

The literature (e.g. Sieminski, 2007 and Hurtado and O’Connor, 2009) indicated that one feature of BIM procurement which exercises most USA commentators is the tension between, on the one hand, the potential disaggregation of design liability and, on the other, the public authority requirement to have one design professional take responsibility for a project (e.g. Sieminski, 2007 and Hurtado and O’Connor, 2009). It was postulated that it was not clear that this aspect would be a regulatory problem in the UK since there was no such statutory requirement in that jurisdiction (McAdam, 2010:6). However, that postulation lacked academic underpinning, and the purpose of this paper is to provide that underpinning.

2 The problem

It is a feature of idealised BIM implementation that there is a collaborative approach to design and delivery embraced by key stakeholders – contractors, engineers, architects, employer (Van Nederveen, et al, 2009). Accordingly “the design” is not whatever the latest issue drawings say, but whatever the BIM currently says, and what it actually says may depend not just on input from the “designers” but also from contractors and/or the employer. So, on the one hand, there is scope for great collaborative endeavour. Equally, however, the regulatory regimes require minimum standards in order to ensure the health and safety of the occupants of the putative building. It is at this interface that the USA commentators have identified a problem with BIM implementation.

3 Regulatory Issues in the US

In the USA, building design and construction is subject to a degree of legislative and bureaucratic oversight, which historically had the primary aim of ensuring the safety and well-being of end users of the buildings (Bruner and O’Connor, 2010, section 16.2). The system operates at State level, with each State (and even City) setting out its own precise requirements (e.g. see CSLB, 2010). Though these differ in the detail, the general approach appears common; codes of building practice are designed (increasingly these are derived from elements of the modestly monikered “International Code”, ICC, 2010a), and a system of pre-construction design submission is imposed, to facilitate quality control.

Integral to this process is the role of the Architect or Engineer. Each State has legislation which determines the characteristics of an Architect or an Engineer. Subject to exceptions for smaller scale works, only an Architect or Engineer (as defined by relevant legislation) is permitted to undertake construction design work and to sign off the designs which are required to be submitted in order to obtain a building permit. For these purposes, state legislation tends to widen the scope of who may sign off designs to include individuals who are under the “Responsible Control” of the Architect or Engineer¹.

The precise meaning of the term may vary from state to state, but at least so far as Architects are concerned, the National Council of Architectural Registration Boards (NCARB) has set down a definition which may be adopted at state level. The NCARB is the umbrella body for each state’s Architects Registration Board, and periodically it defines Legislative Guidelines and Model Law which can be used as a basis by states for how best to regulate the architectural profession. The most recent iteration of the guidelines (NCARB, 2009b) were issued in 2009 and have been changed as a result of discussions regarding Integrated Project Delivery (“IPD”) and BIM (NCARB, 2009a) specifically to widen the scope of “Responsible Control”.

Under the previous version definition, the review of technical submissions [a defined term meaning any document required by the authorities, e.g. in pursuit of an application for a building permit] prepared by another were expressly excluded from the concept of

¹ For example, § 5500.1 California Architects Practice Act 2009

Responsible Control (Dikis, 2009). The way these definitions were worded were considered to be inimical to the collaborative nature of IPD and BIM, yet at the same time concerns were expressed over the risks which could occur as a result of the Architect (in this case) relinquishing elements of control from the design process (Dikis, 2009). The revised version of the definition moderates the situation, by permitting Architects to still exercise responsible control over activities to which others have contributed, provided those others can be trusted and their contributions are simply “incidental”, hence “Responsible Control” is now defined as:

“That amount of control over and detailed professional knowledge of the content of technical submissions during their preparation as is ordinarily exercised by a registered architect applying the required professional standard of care, including but not limited to an architect’s integration of information from manufacturers, suppliers, installers, the architect’s consultants, owners, contractors, or other sources the architect reasonably trusts that is incidental to and intended to be incorporated into the architect’s technical submissions if the architect has coordinated and reviewed such information. Other review, or review and correction, of technical submissions after they have been prepared by others does not constitute the exercise of responsible control because the reviewer has neither control over nor detailed professional knowledge of the content of such submissions throughout their preparation.” (NCARB, 2009b:14)

Though this amendment is merely a change in the recommended rules for architects, it appears nonetheless to be relatively influential. For example, the definition of “responsible control” in the 2009 Californian Architects Practice Act is *“that amount of control over the content of all architectural instruments of service during their preparation that is ordinarily exercised by architects applying the required professional standard of care”*². The meaning of such a referential definition would arguably change as the standards for architectural professional practice changed.

Of course, it is to be noted that the change in rules still places the Architect squarely at the centre of the design process. The NCARB hearings prior to the implementation of the change confirmed the view that *“responsible control by architects over building design was critical to the public’s safety”* (NCARB, 2009a). Gordon Mills, President of the NCARB, is, however, aware that the potential change towards BIM and IPD facilitated procurement could ultimately oust the Architect from that key role (Mills, 2009).

It is specifically this interplay between building control legislation and the required professional oversight of submitted technical documentation which commentators had in mind when highlighting the potential problems with BIM procurement.

As far as can be judged from the scope of this literature review, there has not been such discussion and debate in the USA engineering fraternity regarding this issue, nor indeed does the engineering fraternity appear to have such an overarching umbrella organisation as NCARB, rather each State’s Board of Professional Engineers appears to operate independently. It is outwith the scope of this paper to consider the different

²§ 5535.1 2009 Californian Architects Practice Act

approaches taken by different states to this issue, but a quick review of the legislation and regulation controlling engineering practice in one state, Texas, reveals that the term “responsible control” seems absent whereas “responsible charge” and “responsible supervision” are stated to be synonymous with “direct supervision” which is defined as:

*“The control over and detailed professional knowledge of the work prepared under the engineer's supervision. The degree of control should be such that the engineer personally makes engineering decisions or personally reviews and approves proposed decisions prior to their implementation.”*³

It is this level of control and oversight which is required in relation to the affixing of a seal to a technical submission⁴. Hence it appears that where projects are lead by Engineers, rather than Architects, in the USA there may even greater problems with satisfying regulatory requirements if a fully BIM enabled procurement route is desired.

So to summarise, in the USA there remain problems in reconciling the potential for multi-party design collaboration, with the regulatory requirement for technical submissions to be certified by an Architect or Engineer who has some degree of overall control of the design.

4 Regulatory Issues in the UK

The UK regulatory landscape shares some general features with the USA but has taken a different detailed approach to solutions.

The quality of buildings in the UK is controlled primarily by the Building Act 1984. This defines the aspects of buildings which the UK Parliament has decided shall fall within statutory control, and empowers the UK Secretary of State to make Regulations to detail how controls shall be exercised⁵. The Act also enables additional documents which provide ‘practical guidance’ to be ‘approved’⁶. Compliance with an Approved Document creates a rebuttable presumption that the works in question comply with the requirements of the Building Act 1984, as transmitted through Building Regulations (and vice versa⁷).

The Building Regulations 2000⁸ (as supported by the Approved Documents) set out in detail the minimum standards for a range of different aspects (Parts) of building construction. They also set out the process by which those seeking to carry out construction works should proceed in order to obtain approval. For example, Section 14 of the Building Regulations 2000 defines what amounts to “full plans” for a proposed development. However, what the regulations do not state is who is to submit the plans. The plans need not be submitted by an architect or an engineer. It is the content of the plans which is focussed upon, not their provenance.

³ §131.81(10) Texas Engineering Practice Act and Rules

⁴ §137.33 Texas Engineering Practice Act and Rules

⁵ (ss 1, 2, 2A and Schedule 1). Building Act 1984

⁶ s. 6 Building Act 1984

⁷ s.7 Building Act 1984

⁸ Since amended on many occasions, most recently by SI 2010/719 Building and Approved Inspectors (Amendment) Regulations.

Likewise, there are limited commercial fetters upon the availability of relevant information. All of the regulations, and the detailed Approved Documents can be downloaded for free from a government funded website (Planning Portal, 2010). This differs markedly from the position in the USA where it is necessary to purchase the base International Code documents, albeit the City and State amendments to it are available for free (ICC, 2010b).

In relation to construction professionals, only the term “Architect” benefits from any sort of statutory protection⁹ akin to what exists in the USA. In the UK it is illegal to practise as an “Architect” unless one is registered with the Register of Architects¹⁰. This is administered by the Architects Registration Board¹¹ (“UK ARB”), who also set out minimum ethical standards for Architects via the Architects Code (ARB, 2010) as well as stipulating what qualifications are necessary to be an Architect.

It can therefore be seen that the general functions of the UK ARB are very similar to those of the NCARB and the state level ARBs, but the context differs markedly. The focus of the UK ARB revolves around maintaining the professional standards of those who hold themselves out to be Architects. However, in the UK, it is perfectly legal for any individual to carry out any activity which an Architect carries out, provided only that they must not claim to be an Architect, unless they are registered with the UK ARB. Accordingly there is no corresponding problem in the UK in relation to “responsible control”.

As for UK Engineers, their title has no statutory protection at all, though there are of course a variety of professional bodies which regulate member behaviour, for example the Engineering Council (Engineering Council, 2010).

Health and Safety in construction is a critical issue for regulators in the UK, and integral to the approach for minimising risk are the Construction (Design and Management) Regulations 2007¹². These do require designers to avoid foreseeable risks both to those who construct a building, and those who will ultimately use it¹³. One might have thought that this type of regulation could again interfere with the concept of BIM enabled collaborative design, however once again, the UK approach to regulation focusses on activities rather than titles. So, a “designer” is:

any person (including a client, contractor or other person referred to in these Regulations) who in the course or furtherance of a business—

(a) prepares or modifies a design; or

(b) arranges for or instructs any person under his control to do so,

⁹ Architects Act 1997

¹⁰ ss. 3 and 20 Architects Act 1997

¹¹ s. 1 Architects Act 1997

¹² SI 2007:320

¹³ s.11(3) Construction (Design and Management) Regulations 2007

*relating to a structure or to a product or mechanical or electrical system intended for a particular structure, and a person is deemed to prepare a design where a design is prepared by a person under his control.*¹⁴

Hence, if a BIM collaborator engages in design, they are required to do so in accordance with this regulation.

In short, it can be seen from the above that in the UK, the regulatory problem affecting fully BIM enabled procurement which is caused in the USA by restricting the category of person who is permitted to deposit technical submissions to either an Architect or Engineer, does not seem to exist in the UK. The UK adopts a much more permissive regulatory regime and does not stipulate that building control documents must be completed by any particular category of person.

This examination is, however, only of one discrete aspect of regulatory control in the UK and USA. There has been no consideration at all of some of the other regulatory problems which are highlighted by the commentators, including, for example how the regulatory need for paper based documentation can be squared with flexibility of BIM based design.

5 Conclusion

The purpose of this paper was to seek an academic basis for the postulation that in the UK, unlike in the USA, there was no public authority requirement to have one design professional take responsibility for the design of a project, and hence that particular regulatory problem for BIM enabled procurement did not exist in the UK. It has been demonstrated, that though there are similarities between the building control systems in the UK and the USA, a key difference is that the UK adopts a much more permissive approach to regulatory functions, focussing on the necessary content of design and regulatory documents, without also requiring that the content be prepared by a particular type of specialist.

It therefore appears that the specific assertion in the earlier paper is correct. It is, however, acknowledged that this paper has addressed only one small aspect of the regulatory issues between two geographic areas, and much more research remains to be done in this area.

6 References

Architects Registration Board. (2010), 'Architects Code', http://www.arb.org.uk/professional_standards/regulating_architects/architects_code_2010/code.php, viewed: 01/07/2010

Bruner, P., L. and O'Connor, P. J. (2010), 'Bruner and O'Connor on Construction Law', Thomson Reuters, New York.

¹⁴ s.2 Construction (Design and Management) Regulations 2007

Contractors State License Board. (2010), 'Do I need a building permit for my project?', California State,
<http://www.cslb.ca.gov/consumers/HireAContractor/BuildingPermitRequirements.asp#City> viewed: 01/07/2010

Dikis, W. (2009) 'Responsible Control in a Changing Practice & Technology Environment, Nebraska Board of Engineers and Architects,
http://www.ea.ne.gov/PDFs/presentations/CE_Control_Dikis10-09.pdf, viewed 01/07/2010

Engineering Council. (2010) 'About Us' <http://www.engc.org.uk/>, viewed: 01/07/2010

Hurtado, K. A. and O' Connor, P. J. (2008), 'Contract Issues in the use of Building Information Modelling' *International Construction Law Review*, pp 262-272.

International Code Council. (2010a) 'International Codes: Adoption by State', ICC, Washington DC, <http://www.iccsafe.org/gr/Documents/stateadoptions.pdf> viewed: 01/07/2010

International Code Council. (2010b) 'Store', ICC, Washington DC, <http://www.iccsafe.org/Store/Pages/default.aspx> viewed: 01/07/2010

McAdam, B. (2010), 'The UK Legal Context for Building Information Modelling', CIB World Congress, Salford Quays, May 2010

Mills, G. (2010), 'Emergent Issues for Architects' <http://www.cadlore.com/specialfeatures/softwarearticles/revit/2437-emergent-issues-f>, viewed: 01/07/2010

National Council of Architectural Registration Boards (NCARB). (2009a), "NCARB to Propose Changes to Model Law to Include IPD and BIM", Washington DC, <http://www.ncarb.com/en/News-and-Events/News/2009/NCARB-to-Propose-Changes-to-Model-Law-to-Include-IPD-and-BIM.aspx>, viewed: 01/07/2010

National Council of Architectural Registration Boards (NCARB). (2009b), "Legislative Guidelines and Model Law", Washington DC, http://www.ncarb.org/~media/Files/PDF/Special-Paper/LegislativeGuidelines_10.pdf, viewed: 01/07/2010

Planning Portal. (2010) 'Approved Documents List', <http://www.planningportal.gov.uk/england/professionals/buildingregs/technicalguidance/bcapproveddocumentslist/>, viewed: 01/07/2010

Sieminski, J. (2007), 'Liability and BIM'. *AIA Best Practices BP 13.01.08*. Available from: http://www.aia.org/aiaucmp/groups/ek_members/documents/pdf/aiap037060.pdf, viewed: 27 January 2010

Van Nederveen, S., Beheshti, R. and Gielingh, W. (2009). 'Modelling Concepts for BIM.' In: Underwood, J., and Isikdag, U., eds., *Handbook of Research on Building Information Modeling and Construction Informatics: Concepts and Technologies*, Hershey, PA: Information Science Reference, 1-18.

Defect Liability Period: Employer's Right and Contractor's Liabilities Examined

Ayodeji S. Ojo
Ministry of National Development
Mahe, Seychelles
sundayayoojo@yahoo.co.uk

ABSTRACT

Almost all standard forms of contract contain a provision that deals with defective works. Defective works during construction is easily admissible to contractor and those that occur during defect liability period are within the tune of contractors' normal music. After this defect liability period, most contractors believe that the ball has been passed to the employer or the buyer. Can this always be true? It is a supposed truth that the final certificate will discharge the contractor's liability for the defective works and the cost of remedying same. This has been proved wrong times without number by the court. The court generally does not regard a certificate as being final except where very clear words are used in the contract. This paper examines the legal position of the construction contract in relation to their rights and whose liabilities are the defects after issuance of final certificates. A careful examination and documentary analysis of the law journals and law reports was carried out with final consideration of the position of Housing Finance Management on large low-cost housing projects. The result shows the circumstances to be considered when determining the liability of defects after issuance of final certificate, these are the conclusiveness evidence, consequential extent of the damages/loss, patent defects, fraud and concealment of defects.

INTRODUCTION

Construction cut across a wide spectrum of activities stretching from simple renovation works for private homes to massive construction projects and all the various building activity creates its own unique set of requirements and circumstance. The unique and sometime complex nature of the construction sector with different professionals and group with their different interests pose a sort of different requirements which compete with one another. Most formal building or engineering contracts contain an initial express obligation of the contractor to carry out and complete the works in accordance with the contract. This is, in fact a dual obligation that is, both to carry out and to complete the works (Wallace, 1995). The contractor's basic obligation, so far as the standard of work is concerned, is to comply with the terms of the contract.

Ficken (2006) acknowledge that the contractor is required to perform construction fully in accordance with the contract documents, usually consisting of at least plans, specifications and the building code within required time. Similarly according to Chan (2002), contractor's obligation in a traditional contract to carry out and complete the works would require him to provide the workmanship and materials as required by the specifications given by the architect and engineers. The contractor's basic obligation in any construction projects, so far as the standard of work is concerned is to comply with the terms of the contract. This includes both express terms (such as the requirement of contract that work shall be of the standards described in the bills) and implied terms (such as the principle that all materials shall be of 'satisfactory quality') (Murdoch and Hughes, 2000).

In a construction contract, a contractor undertaking to do work and supply materials impliedly undertakes

- a) to do the work undertaken with care and skill or, as sometimes expressed, in a workmanlike manner;
- b) to use materials of good quality. In the case of materials described expressly this will mean good of their expressed kind and free from defects. (In the case of goods not described, or not described in sufficient detail, there will be reliance on the contractor to that extent, and the warranty (c) below will apply);
- c) that both the workmanship and materials will be reasonably fit for the purpose for which they are required, unless the circumstances of the contract are such as to exclude any such obligation (this obligation is additional to that in (a) and (b), and will only become relevant, for practical purposes in any dispute, if the contractor has fulfilled his obligations under (a) and (b)). (Wallace, 1995)

Issuance of Certificate of Practical Completion brings the contractor's obligation to an end and only defects due to workmanship and materials not in accordance with the contract are required to be made good at the contractor's cost.

At various times, contractors often believe that liability is limited to what is written in the contract. This is a false impression that does not stand the test of contractual obligation. Simon (1979) noted that there are many areas of contractual liability which are implied and not expressed. This implied contractual liability might be the contractor's obligation to perform its work in a good workmanlike manner.

Frank and James (1988) show that liability, obligations and responsibilities do not stop with the contract they are broader and more inclusive. Frankel (2005) noted that the construction defects can arise from improper soil analysis / preparation, site selection and planning, architectural design, civil

and structural engineering, negligent construction or defective building materials. Frankel (2005) further stated that the recent explosion in new construction has spawned, increased construction defect litigation. Construction defect litigation involves all types and sizes of building projects, but homes are its current focus, with the intensity of the concern growing rapidly. The study focus on the increase number of litigation as a result of unresolved defects and its effect in construction industry particularly housing projects.

CONSTRUCTION DEFECTS

In general terms defects or defective works is where the standard and quality of workmanship and materials as specified in the contract is deficient. Defects can be classified into two main categories, Patent Defects and Latent Defects. Patent defects are defects that can be discovered by normal examination or testing whereas Latent Defects are defects that are not discoverable by normal examination or testing which manifests itself after a period of time (Anon 1, 2007).

Atkinson defines defects as breaches of terms and conditions of contract by the contractors (Atkinson, 1999). Defects can take place in any part of a construction project and at any stage of construction. Cama (2004) defines defect in the context of a building contract as “a failure of the completed project to satisfy the express or implied quality or quantity obligations of the construction contract.” Sweet (1993) acknowledge that the construction defects is defined by the law as failure of the building or any building component to be erected in a reasonably workmanlike manner and Marianne (2005), define construction defect as a failure of a building component to be erected in the appropriate manner. Summerlin and Ogborn acknowledge that construction defects can be the result of design error by the architect, a manufacturing flaw, defective materials, improper use or installation of materials, lack of adherence to the blueprint by the contractor, or any combination thereof (Summerlin and Ogborn, 2006).

Construction defects are inevitable in construction projects delivery and are usually contentious between the employer and the contractor or subcontractors. Construction defects are the unacceptable quality of a project which can be identified and remedied. Defective construction works can be defined as works which fell short of complying with the express descriptions or requirements of the contract, especially any drawings or specifications, together with any implied terms and conditions as to its quality, workmanship, durability, aesthetic, performance or design. Most modern buildings and civil structures are complex undertakings and involve the use of a great variety of engineering methods and processes. Therefore most projects face the possibility of defects and defective work, which generally result in structures that cannot perform their originally intended roles (Atkinson, 1999; Cho et al, 2006)

Types and Causes of Construction Defects

Defects in construction projects particularly building projects are attributable to various reasons for example, improper fixing of water pipe, poor quality of materials supplied by building merchants or by combination of poor quality of materials and poor workmanship. Irrespective the causes of the defects, they diminish the quality of construction works and sometime the value.

It is generally believe that there are two major types of defects in building projects. According to Cama (2004) defects are often referred to as patent defects and the latent defects. Where the Latent defects are the opposite of patent defects.

- i. Patent defects are discoverable upon examination or shortcoming in a structure that is apparent to reasonable inspection for example a roof leak or a foundation crack. Normally, defects are readily apparent to the naked eye and are therefore capable of being assessed and measured relatively easily and then, if necessary, rectified.
- ii. Latent defects are those hidden or concealed defects that would not be discovered in the course of a reasonable inspection. A latent defect is by definition something that is not easily discoverable. Normally, defects only become apparent at some later date or upon an investigation of some consequential effects caused by the defect.

Patent defects are plain to see, or at least, that is the theory. Whether the engineer could or should have seen defects on site during site visits has exercised more than one judicial mind. (Nigel, 1996)

As to the second category of defects, i.e. 'Latent' defects, the same two references ascribe the following definitions/meanings (Harbans, 2003): "A defect which is not discoverable during the course of ordinary and reasonable examination but which manifests itself after a period of time. In building and civil engineering work the most common application is defects becoming apparent after the maintenance period expired." Robinson and Lavers describe a 'latent' defect in the following words (Nigel, 1996): "... a defect that cannot be discovered by normal examination and testing..."

Most often to clearly differentiate the patent from latent defects depend on the direction of examination and the expertise of the judge in the court of law. According to Cama the decision to classify a defect as a patent defect or latent defect is up to the judge and the outcome sometimes surprising (Cama, 2004). However, as stated by Chan (2002) that by its nature, a latent defect cannot be discovered until it becomes patent and yet it may not be discovered immediately since there may be no immediately apparent signs to indicate the presence of the defects.

According to Kenneth (2002), common types of construction defects include: structural defects resulting in cracks or collapse; defective or faulty electrical wiring and/or lighting; defective or faulty plumbing; inadequate or faulty drainage systems; inadequate or faulty ventilation, cooling or heating systems; inadequate insulation or sound proofing; and inadequate fire protection/suppression systems.

According to Marianne (2005) all types of defects can typically be grouped into the following four major categories: Design deficiencies, material deficiencies, construction deficiencies and subsurface/geotechnical Problems.

According to the records of Housing Finance Corporation the defects reported from 1997 to 2006 has been related to design deficiencies, construction deficiencies and subsurface problems. Table 1 shows the percentage records of reported defects within the period.

Table 1: Defects Reported by Occupiers – 1997 to 2006

Types of Defects reported	% Reported on yearly basis									
	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006
Design Deficiencies	16%	13%	22%	18%	21%	15%	28%	22%	32%	16%
Material Deficiencies	32%	19%	22%	37%	28%	24%	12%	8%	9%	6%
Construction Deficiencies	45%	36%	44%	41%	39%	48%	51%	28%	12%	26%
Subsurface/Geotechnical Problems	7%	32%	12%	4%	12%	13%	9%	42%	47%	52%

Table 1 above shows the types of defects experienced in most of the housing projects managed by HFC from 1997 to 2006. On the average, the report shows design deficiencies, material deficiencies, construction deficiencies and geotechnical problems to be 20.3%, 19.7%, 37% and 23% respectively over the period.

Liability for Defects

Liability as stated in Osborne Concise Law Dictionary is ‘an amount owed; or subject to legal obligation; or the obligation itself, he who commits a wrong or break on a contract or trust is said to be liable or responsible for it’. Similarly as indicated in Dictionary of English Law, liability is potentially subject to obligation, either generally as including every kind of obligation or in more special sense to denote inchoate, future unascertained or imperfect obligation, as opposed to debt, to essence of which is that they are ascertained and certain. While Burton’s Legal Thesaurus defined liability as accountability, accountable, amenability and answerability. According to Borja and Stevens (2002) a liability is a present obligation of the enterprise arising from past events, the settlement of which is expected to result in an outflow from the enterprise of resources embodying economic benefits, while, Frankel (2005) in any legal responsibility, duty or obligation, the state of one who is bound in law and justice to do something which may be enforced by action.

Frank and James (1988) noted that in law, a person may owe a duty to another person by his own free will in a Contract or by the operation of common law of Tort. The failure to perform or negligently perform these duties or responsibilities constitute a breach, therefore he or she will be answerable or accountable to the other party who may have suffered as a result of his/her wrongful act. In ***Greaves &***

*Co. v Bayham Meikle*¹, Lord Denning M.R. stated: “Apply this to the employment of a professional man. The law does not usually imply a warranty that he will achieve the desired result, but only a term that he will use reasonable care and skill. The surgeon does not warrant that he will cure the patient. Nor does the solicitor warrant that he will win case.”

Liability must, of course, be established on balance of probabilities. In *Bater v Bater [1951] P. 35*² Denning L.J. said: “So also in civil cases, the case may be proved by a preponderance of probability, but there may be degrees of probability within that standard. The degree depends on the subject matter... The degree of probability which a reasonable and just man would require to come to a conclusion”

As Oliver J. pointed out in *Midland Bank v Hett, Stubs & Kemp*³, the obligation to exercise reasonable skill and care is not the only contractual term which ought to be considered in a professional negligence action, there are implied terms that he will draw up the option agreement and effects registration.

The contractor is expected to be liable for any defects in building project according to standard form of contract, these standard forms usually contain detailed provisions in respect of the employer’s remedies in respect of defective works for example according to Ong(2005):

- i. Defective work to be remedied by contractor
- ii. Defective work to be remedied by employer if contractor fails to do so
- iii. Employer may agree to a reduction of contract price instead of remedying the defect
- iv. Employer may deduct the cost of remedial works from the contract price until the remedial works are carried out
- v. Employer to withhold retention monies, to be released upon issuance of the Certificate of Practical Completion and/or Certificate of Making Good Defects.

The houses managed by HFC are being constructed by another sector of government saddled with the responsibilities of housing provision. Ministry of National Development initiate the project, contracted it out, monitored and supervised both the financial and construction issues before handing over to HFC for allocation and managing the mortgaging processes. As the defects are reported to HFC, during the defect liability period, the Employer (MND) attends to all reported cases and makes good same defects.

¹ [1975] 1 WLR 1095

² [1951] P. 35

³ [1979] Ch. 384

HFC as the mortgagee on behalf of government sold each flat on a long lease under the corporation lease agreement for a period of 25 years. During the 25 years, any defects both latent and patent are taking care by HFC but the litigation has been between HFC and the contractor and this has lingered for a long time due to claims that HFC and the occupant of the constructed houses are categorised as third party, thus contractor has no obligation to them. The house of Lords in *Linden Gardens Trust Ltd v Lenesta Sludge Disposals Ltd*⁴ the court held that the recovery of damages for breach on a contract was not dependent or conditional on the plaintiff having a proprietary interest in the subject matter of the contract at the date of the breach. It was stated that the present owner could recover damages for defective work even though the owner suffered no actual damage as the building had been sold for full value before the damage was discovered. A similar case in the Supreme Court of Queensland, Sir Harry Gibbs in the case of *Director of War Service Home v Harris*⁵ said “If the owner subsequently sold the building, or gave it away, to a third person, that would not affect his accrued right against the builder of damages.” In this case the defective works carried out by the defendant for the plaintiff were not discovered until after the houses were sold. By the verdict of learned judge and agreed with by Stable and Hart JJ, that the owner is entitled to recover damages for the cost of rectification of the defective works.

Liability of Contractor to Employer

The contractor is expected to carry out the construction work in workmanlike manner so as to meet the requirement and specification for the project. The condition of the contract may be such that the period of making good defects in construction can be divided into three during construction, during defects liability period and post defects liability period.

During the construction, it is generally believed that the contractor is entitled or has a contractual right to remedy any patent defect or latent defect becoming patent, at anytime up to the date of handing over of the works to the employer. He is expected to be informed of any defective works by the employer’s representative of the defects and make good at contractor’s cost. Should he fail to rectify such defects either on his own or upon instruction of the contract administrator, he is culpable of breach of contract.

Sometimes contractor seem to capitalize on non-receiving such directives, notwithstanding the failure of the employer’s representative to give notice or investigate defects is not sufficient ground to construe the engineer or the contract administrator has accepted the works. In *Miller v Krupp, NSW*⁶, Giles J. said: “Krupp’s [defendant contractor] submission requires that Miller’s [plaintiff employer] agent for the purpose of supervising Krupp’s performance of the contract owed a duty to take care to prevent Krupp from falling properly to perform: a duty to save it from breach of the very contract it

⁴ [1993] 3. All E.R. 417

⁵ [1968] Qd R 275.

⁶ [1992] 11 B.C.L.74

had to perform to Miller's satisfaction. In my view there was not the requisite proximity, a view confirmed by notions of what is fair and reasonable."

In *William Tomkinson & Sons v Parochial Church Council of St Michael*⁷, it was held that the employer was entitled to recover damages for the defective work as it was still a breach of contract, despite occurring during the construction period.

During the defects liability period, which commences on the completion of the works, standard forms of contract generally give the contractor a licence to return to the site for the purpose of remedying defects. In effect, such condition of contract confers upon the contractor a right to repair or make good its defective works, which can (usually) be carried out more cheaply and (possibly) more efficiently than by some outside contractor bought in by the employer (Harbans, 2003)

Lord Diplock, commenting on RIBA/JCT defects liability clause in the case of *P&M Kaye Ltd v Hosier & Dickson Ltd*⁸, said: "Condition 15 imposes upon the contractor a liability to mitigate the damage caused by his breach by making good the defects of construction at his own expense. It confers upon him the corresponding right to do so. It's a necessary implication from this that the employer cannot, as he otherwise could, recover from the contractor the difference between the value of the works if they had been constructed in conformity with the contract and their value in their defective condition, without first giving the contractor the opportunity of making good the defects."

Defects after defect liability period can still be treated by contractor depending on the standard form of contract. It is the contractor's obligation under the contract to rectify the defects that appears during DLP. According to Lord Diplock in *P&M Kaye Ltd v Hosier & Dickinson Ltd*⁹, the DLP's clause is included in the contract with an intention of giving opportunity to the contractor to make good the defects appear during that period.

Liability to Third Party

Most of the defects experienced in housing project are being experienced after handing over of the houses to individual either through mortgaging, buying or renting for a long time. Housing Finance Corporation relies mostly on the information from the occupant to know the level of defect in the completed housing projects. The situation is fully third party liability between the contractor and Housing Finance Corporation, because another ministry handling housing provision handover the completed project to HFC who in turn gives out to the individuals. The increasing number of cases coming up due as a result of liability to third party has been either pure economic loss or injury to person as a result of the defects. The general argument within contractors and concerned professionals

⁷ [1990] 6 Const. LJ 319

⁸ [1972] 1 W.L.R. 146, at p.166

⁹ [1972] 1 W.L.R. 146, at p.166

has been: should a contractor be liable to third party or only to the client. But in cases involving third party and contractor according to Jackson and Powell (1987), following the Anns' case *Anns v Merton London Borough Council*¹⁰, the proposition that a professional person owes a duty to no one other than his client is clearly untenable. Also Barros (1989) asserted that a contractor owe a duty of care to an increasingly wide range of persons who are not his clients. As stated by Lord Wilberforce, in the case of *Anns v Merton London Borough Council*¹¹ that the contractor owed a duty of care to subsequent purchaser to ensure that the building was safe for occupation and use, and owed a duty of care could result in liability for damage to the building itself.

Cost Effects of Construction Defect

The increasing number of construction defect cases has been linked to the increasing demand for houses at avoidable prices. As the number of house demand increases, many general contractors both inexperienced and experienced are bidding in a competitive ways. Summerlin and Ogborn supported this fact as stated in their research that many builders respond to the competition with low bids for contracts, then cut corners, and frequently employ unskilled or overworked subcontractors and poorly supervise subcontracted work (Summerlin and Ogborn, 2006).

The cost of rectifying the defects has been one of the long contentious issues between HFC, MND and Contractors. Till date there is no litigation outcome in Seychelles to determine who bear the cost of defects but the High Court's decision of *Bellgrove v Eldridge*¹² is the leading authority on the measure of damages for defective and incomplete work. The High Court affirmed that the general rules was that the measure of damages was the difference between the contract price of the work and the cost of making the work conform to contract.

Problems Associated with Defective Works in Seychelles

Unlike other developing countries, there is no law from government that governs how to treat defectives works and its effects generally, but the Contract Condition has been the only available documents made use to settle any dispute regarding defects. Clause 15 of the Contract Condition (East African Condition of Contract 1998 Edition) stipulates the process of dealing with defects and where the liability lies.

(1) When in the opinion of the Project Officer the Works are practically completed, he shall forthwith issue a certificate to that effect and Practical Completion of the Works shall be

¹⁰ [1978] AC 728

¹¹ [1978] AC 728

¹² [1954] 90 CLR 613

deemed for all the purposes of this Contract to have taken place on the day named in such certificate.

- (2) *Any defects, shrinkage's or other faults which shall appear within the Defects liability period stated in the appendix to these Conditions and which are due to materials or workmanship not in accordance with the Contract shall be specified by the Project Officer in a Schedule of defects which he shall deliver to the Contractor not later than 14 days after the expiration of the said Defects Liability period, and within a reasonable time after receipt of such Schedule, the defects, shrinkage's and other faults therein specified shall be made good by the Contractor and (unless the Project Officer shall otherwise instruct, in which case the Contract Sum shall be adjusted accordingly) entirely at his own cost.*

- (3) *Notwithstanding sub-clause (2) of this Condition the Project Officer may whenever he considers it necessary to do so, issue instructions requiring any defect, shrinkage or other fault which shall appear within the Defects Liability Period named in the appendix to these Conditions and which is due to materials or workmanship not in accordance with this Contract to be made good, and the Contractor shall within a reasonable time after receipt of such instructions comply with the same and unless the Project Officer shall otherwise instruct, in which case the Contract Sum shall be adjusted accordingly) entirely at his own cost. Provided that no such instruction shall be issued after delivery of a Schedule of defects or after 14 days from the expiration of the said defects Liability period.*

- (4) *When in the opinion of the Project Officer any defects, shrinkage's or other faults which he may have required to be made good under sub-clauses (2) and (3) of this Condition shall have been made good he shall issue a certificate to that effect, and completion of making good defects shall be deemed for all the purposes of this Contract to have taken place on the day named in such certificate.*

Clause 15(1) deals with issuing of certificate of practical completion opening the starting of Defect Liability Period for the project, but the problem has been implementation of Clause 15(2) which involves informing the contractor of any defects on the project so as to enable the contractor to make good of the defects within the reasonable time.

Despite this clause, the problem of responsibility is still there and the effects of defects are clearly seeing to be scaring. The occupant sometimes not aware of the condition of contract tends to go ahead of HFC and called another contractor to do the rectification without minding who bear the cost. Most

often, occupants want HFC to refund the cost of making good any defects carried out by another contractor; this has always been a contentious issue till date.

It is to be noted also that non availability of legislative rule on matters related to building defects affects both the buyer of the house and the sellers. However, due to non-issuance of notice or schedule of defects according to Clause 15(2) sometimes brings confusion. Looking at the decision of the judge in one of the cases whereby the Employer did not send the notice to Contractor to make good the defects discovered. The decision had been held in the Court of Appeal in the case of *Pearce & High Limited v Baxter*¹³ Evans LJ pointed out clearly that “*In my judgment, the contractor is not liable for the full cost of repairs in those circumstances. The employer cannot recover more than the amount which it would have cost the contractor himself to remedy the defects. Thus, the employer’s failure to comply with clause 2.5 (the clause relating to rectification of defects), whether by refusing to allow the contractor to carry out the repair or by failing to give notice of defects, limits the amount of damages which he is entitled to recover. The result is achieved as a matter of legal analysis by permitting the contractor to set off against the employer’s damages the amount by which he, the contractor, has been disadvantaged by not being able or permitted to carry out the repairs himself, or more simply, by reference to the employer’s duty to mitigate his loss.*”

The rectification costs by the occupant sometimes were high contrary to the expectation of HFC and the main contractor refuses to pay back the charges by the other second contractor and this eventually leads to litigation.

FINAL CERTIFICATE

Approved final certificate issued to a contractor signifies both the satisfaction of the employer with the works according to the terms and condition of the contract and discharge of the contractor from any obligation for any other works. Completion of the contract is clearly evidenced by the issuing of final certificate except for concealment, fraud, dishonesty, and after the issuing of certificate and thereafter the employer is unlikely to have any recourse against the contractor for any breach of contract and the employer may only proceed to rectify the works at its own cost. (Willis, 2006). According to Harbans Singh (2003) the final certificate signifies the contract administrator’s satisfaction. The final certificate at the end of the construction signifies also completion of the works but issuing of final certificate has been a contentious issue over the years for HFC for the period under consideration (1997-2008). Most of the contractors are reluctant to engage in any reparation of defective works due to ambiguity on the

¹³ [1999] 66 Con LR 110

problem associated with issuing of final certificate. A cause of action for rectification of defects is generating a huge cost both to the contractor and HFC. The cost effect of litigation between HFC and the contractor has been estimated to be in the range of 200,000.00 – 250,000.00SR (18,000.00 – 20,000.00USD) per annum which is a huge cost.

Finality of Final Certificate

The question has been asked by various researchers, how final is the final certificate in construction projects. It is evidenced from many courts cases that issuing the final certificate relieve the contractor from any obligation from employer and transfer all other responsibility to the employer to make good any ‘discoverable’ defects on the works. Various court case has proved that final certificate issued signify end of liability of the contractor such as in the case of ***Shen Yuan Pai v Dato Wee Hood Teck & Ors***¹⁴ It was held that as the architect in this case had issued his final certificate, thereby showing his satisfaction with the works carried out by the plaintiff, the plaintiff was entitled to the amount claimed. The certificate issued by the architect in this case was conclusive. On conclusiveness of a final certificate, Lee J said that "*The first question is whether, on the true construction of the conditions, a final certificate issued by the architect is conclusive evidence as to the sufficiency of the works subject only to the exceptions mentioned in clause 24(f) of the contract. Melford Stevenson, J., came to the conclusion that it was not, but thought that it was only conclusive evidence as between the building owner and the builder until power to re-open it vested in an arbitrator appointed pursuant to clause 27. In the Court of Appeal, reversing Melford Stevenson, J., on this point, it was unanimously decided that a final certificate was final and conclusive save as to the exceptions stated in the sub-clause.*" The Final Certificate shall be conclusive evidence as to the sufficiency of the said works and materials. It prevented any further legal action, including legal proceedings started before the certificate was issued. The certificate is not reviewable by the arbitrator because it is ‘conclusive evidence that any necessary effect has been given to all the terms of this Contract which require an adjustment to the Contract Sum’.

Many more cases had proved that final certificates has an effect of finality to a contract as shown in the cases of ***James Png Construction Pte Ltd v Tsu Chin Kwan Peter***¹⁵, ***Sa Shee (Sarawak) Sdn Bhd v Sejadu Sdn Bhd***¹⁶, ***University Fixed Assets Limited v Architects Design Partnership***¹⁷, and ***Crown Estate Commissioners v John Mowlem and Co Ltd***¹⁸¹⁶⁹

The Way Forward

¹⁴ [1976] 1 MLJ 16

¹⁵ [1991] 1 MLJ 449

¹⁶ [2000] 5 MLJ 414

¹⁷ [1999] 64 Con LR 12

¹⁸ [1994] 70 BLR 1

As it could be derived from various cases, the Project Officer of MND should not issue the final certificate until after the end of the defects liability period because any defects that occur thereafter becomes the liability of MND and passed over to HFC. It is an assume risk of the conclusiveness of the final certificate both as to completion and quality of the workmanship.

However the employer through the representative should ensure that defects and defective works are properly checked before issuing the final certificate since, final certificate was irrelevant to any claims for consequential damages in respect of defects which had been found after the employer had taken possession. It is the right of the employer to claims for consequential loss.

Logically, contractors take relieve from a final certificate and assume that no new or further claims can be made against them for defective or incomplete works under the final certificate, other than latent defects, if latent defects are discovered during defects liability period, it is hereby advisable that the period be extended until the contractor has made them good and the MND has so certified.

If defective work or workmanship or design have been knowingly covered up or concealed so as to constitute fraud, the commencement of the limitation period may be delayed. The decided period may be delayed until discovery actually occurs; or at least the defect could have been discovered with reasonable diligence, whichever is earlier.

There should be clear demarcation between latent and patent defects in the condition of contract so as to avoid concealment till end of defect liability period which generally takes care of the patent defective works.

Notification is of high importance before taking over any building either by MND or HFC; furthermore, the buyer of the house must notify HFC as early as possible. The owner must notify latent defects to the MND within 14 days of becoming aware of them and MND must ensure that the contractor acknowledge the receipt of such notice since the contractor's acknowledgement of a defect amounts to notification of the same.

HFC must not accept any works without a proper checking irrespective of time constraints, because if the owner accepts the works without reservation in such cases, he will no longer be entitled to avail of the remedies against defects set out in the contract except in the case of malice. Legal doctrine is of the opinion that if the constructor has concealed the defect with malice, no notification is required.

In the case of the existence of defects, the owner must exercise his rights in respect of the same within period stipulated in the contract and part of the rights are but not limited to these ones: the right to demand repair of the defects; the right to the performance of new works; the right to a reduction in price; the right to termination of the contract; and the right to claim an indemnity.

CONCLUSION

It has been observed from the study that where there is a defects clause and at the end of the defects liability period a binding and conclusive final certificate of satisfaction is given by the project officer, the contractor's liability in contract for any defects lapsed. The employer is not entitled contractually to pursue any defect arising and/or reported after the said date. If the final certificate is not conclusive, it does not derogate in any way the contractor's liability for defective work at law. The contractor continues to be liable for such defects for the duration of the applicable statutory period of limitation.

Issuance of final certificate under the contract, prevent the employer from bringing claims for defects in the works arising from a breach of contract and/or tort irrespective of whether those defects may be said to be latent or patent. In the circumstances where defective work or workmanship have been knowingly covered up or concealed so as to constitute fraud, the commencement of the limitation period may be delayed until discovery actually occurs. This research can be said not to be all-inclusive; but the study hopes to provide assistance to both contractor and employers to comprehend their rights and liabilities in relation to defects which appear after issue of the Final Certificate.

REFERENCES

- Anon, 2007: *"What Are The Obligations Of The Contractor During Defect Liability Period?"*
The Entrusty Group, Master Builders, 1st quarter 2007
- Atkinson, D. (1999) Measures of Damages for Defects. ([Link »](#))
- Ball P.L. (1987): The Economics and Assurance of Quality in Construction, Conference Paper, Quality: A Shared Commitment, EOQC, London, 1987.
- Cama, J. (2004). Who Pays to Fix Building Defects? American Systems USA inc. Berrymans Legal Consultants.Chan (2002),
- Cho, Y. J., Hyun, C. T., Lee, S. B. and Diekmann, J. (2006): *Characteristics of contractor's liabilities for defects and defective works in Korean public projects* Journal of Professional Issues in Engineering Education and Practice, Volume 132, Issue 2, pp. 180-186.
- Ficken. B. W. (2006). Legal Consideration and Dispute Resolution: The Water- Related Construction Failure. American Society for Testing and Materials. Race Street, Philadelphia, PA 19103.
- Frank E. and James A. (1988). Building Subsidence: Liability and Insurance. London:
- Frankel E. R. (2005). Insurance Coverage for Construction Defect Claims. Real Estate Finance; ABI/INFORM Global. pg 20.

- Hammarlund Y., Jacobsson, S., Josephson P. (1990): Quality Failure Costs in Building Construction, CIB International Symposium at the University of Technology, Sydney, 14–21 March 1990
- Harbans Singh. (2003). Engineering and Construction Contracts Management – Post Commencement Practice.” (Singapore: LexisNexis, 2003.) pp. 695-696
- Jackson L. (1987): Kvalitetsstyrning inom ett byggföretag, The Swedish Council for Building Research, Report R76:1987, Svenskt Tryckeri, Stockholm, 1987.
- Kenneth. S. Grossbart. (2002). Construction Defects, An analysis of SB 800. Reeves Journal. ABI/INFORM Trade & Industry, pg.8 London: Spon Press.
- Marianne, J. (2005). Building defects spoil homeowners’ dreams. Portland: The Oregonian News. The Aldrich Law Office, P.C. 522 SW 5th Avenue.
- Murdoch, J and Hughes, W. (2000). Construction Contracts: Law and Management.
- Nigel M Robinson (1996): “Construction Law in Singapore and Malaysia.” 2nd Edition. (Butterworths Asia Malaysia, 1996) pp. 160-162
- Ong See Lian. 2005. “Defective Works.” International Conference on Construction Law & Arbitration. (26th – 28th April 2005, Kuala Lumpur.) pp. 1.
- .
- Simon, S. M. (1979), Construction Contracts and Claims. New York: McGraw-Hill Book Company.
- Summerlin & Ogborn. (2006). Construction Defects. Construction Law Attorneys, Thomson Business.
- Sweet, J. J. (1993). Avoiding or Minimizing Construction Litigation. San Jose California: Wiley Law Publication.
- Van den Beukel A. (1989): Quality Cost, Committee on Housing, Building and Planning, Economic Commission for Europe, Economic and Social Council, United Nations, 1989.
- Wallace, D. (1995). Hudson’s Building and Engineering Contracts. 11th ed. London: Willis, 2006

Standard form of contract for decoration, repair and maintenance works for laymen use

Paul H K Ho

Division of Building Science and Technology,
City University of Hong Kong,
Tat Chee Avenue, Kowloon, Hong Kong SAR, PR China

Email: bshkho@cityu.edu.hk

Abstract:

In Hong Kong, there are a vast number of new and old buildings which need regular decorations, repairs and maintenances. For large-scale projects, the owner normally appoints a professional consultant to help them plan, design and administrate the works. However, for small-scale projects, some owners may directly appoint a contractor to carry out the works. Unfortunately, most owners, as lay customers of construction services, do not have knowledge in building construction and contract. Even a contract is signed, it may not be detailed enough or one-sided in favour of the contractor. As a result, owners are often in disadvantageous position in the event of disputes with contractors. In order to give owners a reasonable protection, a 'Standard Form of Contract for Decoration, Repair and Maintenance Works' was published for laymen use. This paper explains the major clauses of this contract, and compares the differences between this contract and other standard form of building contracts administered by building professionals. This paper also critically examines the underlying considerations on the protection of lay owners under this contract.

Keywords:

Standard form of building contract; decoration, repair and maintenance works

1 Introduction

In Hong Kong, there are currently about 17,000 buildings aged 30 years or above and, in ten years, the number will reach 28,000. Many of these buildings are in disrepair conditions and also contain illegal structures and unauthorized building works, which may potentially endanger the occupiers and the public. According to the Development Bureau (2010a), from the years 2001 to 2009, 379,510 unauthorized building works were removed by the Government; 5,339 illegal rooftop structures were cleared; 20,345 abandoned signboards installed on external walls were removed; 273,946 removal orders, 10,330 repair orders and 35,213 warning notices were issued. Recently, following the tragic collapse of one old building in January 2010, the Government immediately inspected 4,011 buildings aged 50 or above. The results of the inspection show that the conditions of 2 building (i.e. 0.05 %) would require emergency remedial works; 1,030 building (i.e. 25.7%) were found to exhibit different extents of defects which warrant the issue of a statutory repair/investigation order; 1,270 building (i.e.

31.7%) were found with minor defects; and the remaining 1,709 buildings (i.e. 42.6%) were found with no apparent defect (Development Bureau, 2010b). To a certain extent, these statistics indicate the massive need for repair and maintenance works. Recognizing the importance of building maintenance and public safety, the Government in recent years has actively promoted proper building maintenance, has assisted owners in carrying out maintenance works, and has enforced the relevant legislation against unauthorized building works. In particular, the Government has launched in May 2009 the HK\$2.5 billion Operation Building Bright to provide subsidies to help owners of old and dilapidated buildings to carry out repair and maintenance works. Besides, the Government has also established the HK\$1 billion Building Maintenance Grant Scheme for Elderly Owner to enhance financial assistance to elderly owner-occupiers to repair and maintain their self-occupied properties. In addition to existing buildings, there are about 150 – 250 new buildings completed each year which also need certain decoration or fitting-out works before its occupation or use. Therefore, these, all together, have created a great amount of decoration, repair and maintenance works.

Some of these decoration, repair and maintenance projects are large in scale and may include, for instance, re-laying of the whole roofing system, re-titling of external walls, re-decoration of lobbies, staircases and other common areas, replacement of fire services, electrical and lift installations, etc. On the other hand, there are also many small projects involving, for instance, repair of spalled concrete, fixing of water pipes, repair of windows, replacement of fire doors, decoration of individual apartments, etc. For large-scale projects, the appropriate professional consultant (such as the architect, engineer and surveyor) should be employed to implement the works. Basically, they should be responsible for the preparation of the design and contract documents. For small-scale projects, owners may not be affordable or justifiable to employ a professional consultant in view of the small contract sum. Under such circumstances, they may directly appoint a contractor to carry out the works without signing a proper contract.

Owners generally do not have adequate knowledge about construction and law. For many small projects, there may be no formal or comprehensive contract signed between the owner and contractor. Even if there is a formal contract, the contract provisions may be one-sided without reasonable protection to owners. As such, owners, as laymen, are always in an unfavourable position, particularly when there are claims for an extension of time and/or additional cost by contractors. In order to overcome this problem, the Hong Kong Institute of Surveyors published a 'Standard Form of Contract for Decoration, Repair and Maintenance Works' for use by lay owners. The author, as the Chairman of the Quantity Surveying Division of the Institute at that time, was the principal drafter of this contract.

The consideration for drafting each provision of a contract for use by laymen is entirely different from a contract for use by building professionals. For instance, since the contract is administered by lay owners, it should be easily understood by using plain English and without many technical or legal jargons such as the employer, site, contract documents, contract sum, contract period, variation, and extension of time, which are replaced by such commonly used words as the owner, premises, work details, price, change, working period, and extension of working period, respectively. In addition, while some fundamental duties, responsibilities and obligations of the contracting

parties are well established in common law, it cannot assume that they have knowledge about building law. Thus, many implied terms in law are written as express terms in this contract in order to let both parties clearly understand their duties, responsibilities and obligations. Furthermore, particular consideration is also given to the protection of lay owners as customers of construction services rendered by experienced contractors. The main purposes of this paper are to compare the differences between the main provisions of this contract and other standard forms of building contract, and to critically examine the underlying consideration of main clauses of this contract on the protection of lay owners.

2 Main Provisions of the Contract

This contract is designed for use where the owner deals directly with a contractor making improvements to his premises such as decorations, repairs, maintenance or minor alterations, and the works involved is simple in character with a contract value up to HK\$400,000 (at 2008 prices). Where the works are of a complex nature or involve building works defined under Buildings Ordinance, the owner should consider appointing a professional consultant who will act for them, including choosing the appropriate contract.

While a good contract can provide certain protection to the owner, it is more important to find a good contractor. Thus, the owner should choose reputable contractors for the works. The owners must decide exactly the work details before asking the contractors to quote for the works. This will avoid misunderstandings and disputes later on. The owner should normally get quotations from more than one contractor in order to obtain a competitive price for the works. Three quotations are quite usual. Before signing the contract, both parties must agree all contract details, including the work details, price, working period and hours, etc. The owners must make sure that all the appropriate parts of the contract have been filled in.

As mentioned above, this contract is written in plain language (in both Chinese and English) and is kept as simple as possible. It only consists of 20 clauses in 4 pages plus 2 pages guidance notes which do not form part of the contract. The following paragraphs describe each of these clauses in detail.

2.1 Documents forming parts of the Contract

Modern building contracts typically comprise the articles of agreement, the form of tender, the employer's letter of acceptance of the contractor's tender and any correspondence between the parties expressed to form part of the contract, the special conditions (if any), the conditions, the contract drawings, the specification and the contract bills. Although each of these contract documents is prepared by professional consultants, disputes always arise whether a particular document has been incorporated into, or forms part of, the contract. Disputes also often arise from ambiguity or discrepancy in or divergence between the various contract documents. While the courts normally give more weight to documents such as the contract bills which have been specifically prepared for the particular contract, they have also reversed the ordinary rule in the case of *English Industrial Estates Corporation v. George Wimpey & Co. Ltd.* (1972) that written words prevail over printed words in standard form contracts. Not

infrequently, the contract bills may be found inconsistent with the contract drawings and/or the specification. Under such a circumstance, whether the contract bills will take precedence over the contract drawings and/or the specification will depend on the express contract provision, if any, governing the order of precedence of the various contract documents.

As no professional consultant is employed to produce the drawings and specification, the owner usually makes known his requirements in writing or orally to the tendering contractors. The contractor then produces the design in the form of the drawings and specification together with a quotation for the owner's consideration. It should be aware that if the owner's requirements are very brief, he will have little control over the end product. On the other hand, if the owner's requirements consist of a set of drawings and specification, the contractor will need to do little but to carry out the works and the owner will know exactly what is to be provided. In any event, the owner must crystallise his requirements before executing the contract. To facilitate the formation of a legally binding contract, a simple and easily administrated contractual arrangement is adopted by asking the contracting parties to tick the appropriate boxes provided to confirm whether there are any quotations, drawings, specification or other documents (such as the owner's requirements) which have been agreed to form part of the contract. These are collectively called the "work details" in this contract. Since an offer may be accepted by conduct in law, guidance note is thus provided to remind both parties that they must sign all agreed documents, and that the owner must keep the original documents.

2.2 Lump Sum Contract

There are three main types of building contract: lump sum, measure and value, and cost plus percentage. In order to give a greater certainty on the agreed contract sum, this contract is based on the lump sum contract in which the contractor undertakes to carry out and complete the works in accordance with the submitted quotation, drawings, specification and other agreed documents (i.e. the work details) for an agreed price. Administratively, both parties are simply required to fill in the agreed price in the Contract. As a lump sum contract, the contractor is entitled to nothing until the works has been wholly completed: *Cutter v. Powell (1795)*. This lump sum can be increased or decreased if the building owner has subsequently requested any changes (i.e. variations) to the originally agreed works.

The traditional lump sum building contract works well on the situation that the works itself can be precisely defined in form of either the drawings and specification or the drawings and bills of quantities. As noted above, the various contract documents (i.e. the quotation, drawings and specification) are primarily produced by the contractor. These documents may be unclear, ambiguous or incomplete, frequently leading to the dispute whether a particular work should be included in the original contract. For instance, whether the door is fixed by two or three hinges and its material quality may not be clearly shown in the drawing and/or described in the specification. The incomplete or ambiguous quotation, drawings and specification may create a loophole for the contractor to make claims for both additional cost and extension of working period. For the avoidance of misunderstandings and contractual claims, guidance note is provided to warn the owner that he should decide the work details in sufficient detail to

enable them to be defined and priced by the contractor. More importantly, the contract specifies that 'the price is deemed to include the cost of all works necessary for the timely and satisfactory completion of the works in its entirety, whether it is expressly stated in the work details or not'. One may argue that this clause is not fair to the contractor. However, Wallace (1994) argues that no description will descend to every detail of building processes so that it is a cardinal principle of construction that, in the absence of an expressed contrary intention (for instance, the bills of quantities should be prepared in accordance with the specified standard method of measurement), an obligation to do the described work imports an obligation to do all the necessary ancillary work or processes, whether described or not, which are needed to produce the described works. The contractor who prepares the various contract documents but misses out some obviously necessary works will be responsible to do so at his peril.

2.3 Contractor's Obligations

Same as the traditional standard forms of building contract, this contract imposes on the contractor an obligation to organize, carry out and complete the works diligently and in a good and workmanlike manner in accordance with the work details and the provisions of this contract. This is the most fundamental obligation of the contractor. As an entire contract, the contractor must complete the works in its entirety. If he fails to complete the works, he will be in breach of the contract unless the contract has come to an end for some reason. Within this fundamental obligation, there are two distinct parts which need further considerations. First, this contract contains an express provision which requires the contractor to work diligently. As a matter of business efficacy, it is an implied term that the contractor will proceed with reasonable diligence and maintain reasonable progress. If the contractor does not 'regularly and diligently' carry out the works, the owner is entitled to terminate this contract. Whether this standard is achieved is probably to be judged according to the usage of the industry: *Hounslow Borough Council v. Twickenham Garden Development Ltd. (1970)*. Thus, this contract provides a control on the contractor's progress and completion. Second, this contract also contains an express provision which requires the contractor to complete the works in 'a good and workmanlike manner'. According to *Hancock v. B.W. Brazier (Anerley) Ltd. (1966)*, it is an implied term of all building contracts that the contractor will do his work in a good and workmanlike manner. If the contractor is incompetent so that the works are of an unacceptable standard and not in accordance with work details, the owner is again entitled to terminate this contract.

In the classic standard form of building contract, the owner and his professional consultants assume responsibility for the design of the whole works, whereas the contractor agrees to carry out and complete the works according to the design and thus, does not undertake liability for the design: *Cable (1956) Ltd. v. Hutcherson Brothers Pty. Ltd. (1969)*. As mentioned above, since no professional consultant is employed for minor decoration, repair and maintenance works, the owner has to rely on the contractor's skill and judgment in design. Thus, in addition to the fundamental obligation described above, this contract imposes on the contractor an obligation 'to complete the design for the works using reasonable skill and care, including the selection or specification of the materials and goods and workmanship to be used in the works so far as not stated in the work details'. This is similar to the design-and-build

contract. However, as far as the design element is concerned, the contractor's liability is the same as a professional person.

This contract also expressly places an obligation on the contractor to comply with all statutory requirements applicable to the works. It includes obtaining all approvals and consents from Buildings Department and other authorities, and paying all fees and charges. This obligation covers not only the design, but also the whole of the construction. Based on the case of *Ano. v. Sibbabwe Ltd. and Ano. (1980)*, it is an implied term of all building contracts that the contractor would comply with all statutory obligations, including of course the Building Regulations. This obligation overrides the contractor's obligation under the contract to comply with the work details and the owner's requirements. If there is such a conflict, then the implied term must prevail. It is the contractor who commits a criminal offence if the statutory regulations are breached.

For proper management of a building with multi-owners, there are often a Deed of Mutual Covenant and Management Agreement and House Rules of the Premises in respect of the working hours, control of air, noise, wastewater and waste, delivery of materials and goods, connection to public services, etc. This contract also expressly places an obligation on the contractor to comply with any of such Deed of Mutual Covenant and Management Agreement and House Rules of the Premises. These restrictions should be made known to the contractor during the tendering stage.

2.4 Materials, Goods and Workmanship

As noted above, the owner may rely on the drawings and specification produced by the contractor. Sometimes, there is no specification at all. The contractor may argue that provided the works comply with the agreed work details (including, for instance, a faulty material), the express obligation has been complied with. In order to set a minimum requirement in respect of the quality of materials and goods and standard of workmanship, this contract requires the contractor to 'supply materials and goods which are of satisfactory types, standards and quality and as set out in the work details'. In addition, this contract also requires the contractor to 'warrant that all materials and goods are reasonably fit for the intended purposes'. Moreover, this contract further requires the contractor to 'carry out the works in a good and workmanlike manner, as set out in the work details and to the owner's satisfaction'. Basically, these express provisions are consistent with the various implied terms previously established in law, whereby the contractor impliedly agrees firstly that he will supply good and proper materials: *Young & Marten v. McManus Childs Ltd*, (1969), secondly that it will be reasonably fit for the purposes required: *Viking Grain Storage Ltd. v. T.H. White Installations Ltd* (1985), and finally that he will do his work in a good and workmanlike manner and with care and skill: *Hancoak and Others v. B.W. Brazier (Anerley) Ltd*. (1966). If an express term of the contract deliberately or unintentionally sets a lower liability standard for the contractor, it will be subject to the requirement of 'reasonableness' under the various statutory regulations.

2.5 Working Period and Hours, and Sufficient Working Areas

The contracting parties must clearly state the agreed commencement and completion dates (i.e. the working period) in the contract. As the works are carried out in existing

premises, they must also clearly state the working hours in the contract. Thus, the owner has an obligation to give the contractor access to the premises during the agreed 'working hours' throughout the 'working period'.

It is an implied term of all building contracts that the owner will give possession of the site to the contractor in sufficient time for him to complete his works by the specified completion date: *Freemen & Sun v. Hensler* (1900). While exclusive possession of the site may not be applicable to the decoration, repair and maintenance projects, a sufficient working area is clearly a necessary pre-condition of the contractor's performance of his obligation: *The Queen in Right of Canada v. Watler Cabott Construction Ltd.* (1975). Thus, the contract contains an express provision which requires the owner to 'keep the working areas sufficiently clear of obstructions to allow the contractor to carry out the works so as to enable him to complete the works on time'. Whether the working area is sufficient is a question of fact to be determined in light of all the circumstances: *London Borough of Hounslow v. Twickenham Garden Developments Ltd.* (1970). If he prevents or obstructs the contractor from carrying out the works for a continuous period of more than 14 days, the contractor is entitled to terminate the contract.

2.6 Changes of the Works

Most standard forms of building contract empower the owner's architect to issue an instruction to change the original work. Otherwise, if the contractor has not done the originally contracted work, he may not be able to recover the contract price and may even be liable in damages for breach of contract (Wallace, 1994). This contract also provides a variation clause; otherwise, the owner will have no authority to order changes of the original works. The scope of changes may include the 'alteration of the design, quality or quantity' of the work details, which should be wider enough to cover the possible changes needed in ordinary decoration, repair and maintenance works. However, there are certain limitations on the power to order changes. The owner cannot take away part of the contract works from the contractor and gives to others: *Commissioner for Main Roads v. Reed & Stuart Pty Ltd.* (1974). In addition, he has also no power to impose restrictions regarding access to the premises, limitation of working space or working hour, or the sequence of carrying out the works. Moreover, he cannot fundamentally change the scope or nature of the works.

The variation clause in most building contracts also contains valuation rules for variations. As long as the instruction is empowered under the contract, the contractor must carry out the variation order as soon as practicable, and the variation will be measured and evaluated by the quantity surveyor in accordance with the valuation rules. Given that no quantity surveyor or other consultant is employed, this contract does not contain such valuation rules. In order to realize the cost and time implications, this contract requires the contractor to quote a price for the change and the time involved for the owner's agreement. The contractor will only proceed with the change after receiving the owner's written agreement. If the owner does not agree with the price and time quoted by the contractor, he can only find another contractor to execute the varied work after the completion of the original works.

2.7 Sub-contracting

Sub-contracting is a common practice in the building industry. In order to protect the owner, this contract expressly states that if the contractor wishes to sub-contract any part of the works, he should firstly obtain the owner's consent. Such consent must not be unreasonably delayed or withheld. There is no requirement that the contractor must inform the owner of the names of sub-contractors. It is merely consent to the fact of sub-contracting which is required.

2.8 Injury, Damage and Insurance

All building works involves a certain degree of risk. An accident may result in personal injuries by workmen, occupier or neighbours, or damage to the works itself, existing property and its contents or adjoining owners' properties. Similar to classic standard forms of building contract, this contract provides an indemnity by the contractor to the owner 'against any expense, liability, loss, claim and damage which may be caused to his employees and his sub-contractor's employees of all tiers, or to the premises and its contents, the works or neighbouring properties'. This indemnity clause is wider enough to cover most (if not all) possible injuries or damages to persons and properties arising from the carrying out of the works. However, there is one important qualification. It is well established that if a person obtains an indemnity against the consequences of certain acts, the indemnity is not to be construed so as to include the consequences of his own negligence unless those consequences are covered either expressly or by necessary implication: *Walters v. Whessoe Ltd. And Shell Refining Co. Ltd. (1960)*. Thus, the indemnity does not cover the consequences of a negligent act committed by those for whom the owner is responsible under the express terms of the contract, and others for whom the owner is in law vicariously responsible (Parris, 1985).

Given the short contract duration and small contract sum for minor decoration, repair and maintenance projects, it is not uncommon that the contractor procures no or inadequate insurances to cover the potential risks. Thus, the contractor is expressly required to take out and maintain, in the joint names of the owner, himself and his sub-contractors of all tiers, three types of insurances; Contractor's All Risks insurance, Third Party Liability insurance and Employees' Compensation insurance. The Contractor's All Risks insurance covers the full reinstatement value of the works; the Third Party Liability insurance covers the liability to the third party in respect of bodily injury or death and damage to properties with a specified minimum amount for any one occurrence or series of occurrences arising out of any one event, and the Employees' Compensation insurance should comply with the Employees Compensation Ordinance to cover claims for bodily injury to or the death of any workers employed in connection with the works. The insurance clause is subject to the opening words of 'without prejudice to his obligation to indemnify the owner', which indicates that the indemnity is to stand independently of the insurance and that the contractor cannot claim to have fulfilled his obligation under the indemnity clause in respect of injury to persons or property by effecting insurance policies: *Gold v. Patman & Fotheringham Ltd. (1958)*. If the contractor has failed to take out or maintain the cover or if he has procured an inadequate cover, he will still be liable for the additional loss and damage under the indemnity clause. For the additional security of the owner, the contractor must insure in the joint names of the owner, himself and his sub-contractors of all tiers.

2.9 Extension of Working Period

As a matter of principle, the extension of time (or working period) clause aims to protect the owner rather than the contractor. If there is no contract provision allowing the owner to grant an extension of time for any delay caused by him, time will be at large. The contractor's obligation is to complete the works within a reasonable time: *Wells v. Army & Navy Co-operative Society Ltd. (1902)*. If the owner wishes to recover any loss and/or damage arising from the contractor's delay, he has to prove that the contractor has not failed to complete the works within a reasonable time. Nevertheless, the extension of time clause in most building contracts is administered by a professional consultant in view of the complicated nature of construction processes. However, given the situation that no professional consultant is employed, this contract adopts a simple mechanism to enable the owner to administer the extension of time clause. Thus, a fair and reasonable extension of time will be granted only if the delay is due to two relevant events. The first relevant event is due to any change requested by the owner. As mentioned above, if there is a change requested by the owner, the contractor is required to provide a quotation for the change and the time involved. If the proposed price and extra time is found to be acceptable, the owner can grant the extension of time accordingly. If not accepted, he must find another contractor to carry out the varied work later. The second relevant event is due to 'any delay caused by the owner'. This simple provision may widely cover any special but justifiable circumstances (such as an act of God), and an act of prevention, a breach of contract or other default by the owner or any person for whom the owner is responsible. In minor decoration, repair and maintenance projects, the second specified event is not expected to be happened frequently.

In modern building practice, the principal function of the extension of time clause is to enable a date to be fixed for the calculation of liquidated damages for delay. Thus, most building contracts also contain a liquidated damages clause, in which the owner's right to deduct liquidated damages from the contractor is dependent upon the issue by the architect of a certificate of non-completion to confirm that the contractor has failed to complete the works by the original or extended completion date. This is a condition precedent, without which the owner has no right to make any deduction or recover any liquidated damages: *Algrey Contractors Ltd. V. Tenth Moat Housing society Ltd. (1972)*. However, the building owner will not be able to administer this complicated provision. Thus, this contract does not contain a liquidated damages clause. Nevertheless, it does not mean that there is no protection to the owner in the event of delay caused by the contractor. As mentioned above, if the contractor does not regularly and diligently carry out the works, the building owner is entitled to terminate the contract.

2.10 Payment

Since payment is the lifeblood of the contractor, most standard forms of building contract provide interim payments to the contractor on the basis of the payment evaluation and certificate by the quantity surveyor and architect, respectively. Given that there is neither quantity surveyor nor architect, this contract provides two simple payment methods: single payment and stage payment. Where the project duration is short (for instance, less than or about one month), the owner can choose to make a 'single payment' when the works are all completed by the contractor. The contractor is

not entitled to any payment until he has completed the whole works. Thus, this simple payment method gives a greater protection to the owner. On the other hand, where the project duration is relatively long, the owner can choose to make a 'stage payment' when each of the specified stages of the work is completed by the contractor. Under such a method, the owner must insert the details of each work stage, together with the corresponding amount of payment, in the contract.

When all of the works are properly completed or when each stage of the work is properly completed, the contractor is required to give the owner an itemised invoice for the amount due, taking account of any price increase or decrease due to changes. The itemised invoice is to facilitate the owner to check the actual works done. Since there is no architect to certify whether the works have been 'properly' completed in accordance with the work details, the contract expressly states that 'the works shall be considered as being properly completed only when they are free from obvious defects'. If there is any obvious defect, the owner has the right to withhold payment to the contractor. On the other hand, if no obvious defect is found, the owner must pay within 7 days after receiving the contractor's invoice unless a greater sum has already been paid (i.e. overpayment in the previous stage payment). Failure to pay is a ground entitling the contractor to terminate the contract.

The payment clause in most building contracts also contains provisions governing the retention and defect liability period which are to protect the owner. This contract also provides a simple retention clause, in which the owner is only required to pay 90% (or other specified percentage) of the amount of the invoice. In effect, 10% retention money is withheld from each payment to the contractor. While no defect liability period clause is provided, this contract contains an express provision, whereby the owner is only required to pay the remaining portion of the invoices after the contractor has made good all defects in the works which may be or become apparent at any time within 3 months (or other specified period) after its completion. In effect, there are 3 months (or other specified period) defect liability period. The owner has the right not to release the retention money unless and until the contractor has made good all defects within this defect liability period.

2.11 Occupation and Security of the Premises

The owner must specify in the contract whether or not the premises will be occupied during the working period. If the premises are indicated to be vacant, the implication is that possession of the whole premises will be given to the contractor throughout the working period, including any extension of time granted. During this period of time, the contractor has an obligation to take practical precautions to deter intruders entering into the premises.

If the premises are indicated to be occupied, the building owner still has an obligation to provide sufficient working areas for the contractor to carry out the works so as to enable him to complete the works on time as described above.

2.12 Use of Facilities on the Premises

While the duration of most decoration, repair and maintenance projects is relatively short, the contractor also takes time to set up the various site facilities such as the

temporary electricity, toilet, telephone and water before starting the construction works. This period can be significantly shortened if the owner allows the contractor to use his existing facilities. Thus, the owner is required to state in the contract whether he will allow the contractor to use the electricity, toilet, telephone and water free of charge.

2.13 Health and Safety

Health and safety are an important issue, particularly when the works are carried out within existing premises. This contract imposes on the contractor an obligation to take all practical steps to prevent health and safety risks to the building owner and other people occupying or visiting the premises, and minimize environmental pollution, nuisance, disturbance or inconvenience to the existing and neighbouring occupants. The contractor is required to provide suitable safety and precautionary measures for carrying out the works, particularly on external walls of the premises. In view of the importance of this issue, the owner is entitled to terminate this contract if the contractor does not comply with his health and safety and environmental responsibilities.

On the other hand, this contract also imposes on the owner an obligation to take notice of all warnings the contractor gives about any health and safety or environmental risks which he is taking measures to prevent or minimize.

2.14 Protection and Cleaning of the Works

Since the works are carried out within existing premises, the contractor is required to stow away his tools, equipment and ladders at the end of each working day in a place agreed to be used for such storage, and regularly dispose of any rubbish from the premises. The contractor is also required to protect the finished works and the premises during the progress of the works, and bear all costs incurred in making good any damage caused. Upon completion of the works, the contractor is required to remove all plant, tools and surplus materials, make good and reinstate all damages, clean the works and leave the premises in a clean and tidy condition to the building owner's satisfaction. These are typical preliminaries clauses in the bills of quantities or specification of most building contracts.

2.15 Owner's Right to Terminate the Contract

Most building contracts usually contain clauses entitling one party to bring the contract to an end in certain circumstances. This contract sets out the following three grounds entitling the owner to determine the contract:

- The contractor does not regularly and diligently carry out the works; or
- The contractor does not comply with his health and safety and environmental responsibilities; or
- The contractor is incompetent so that the works are of an unacceptable standard and not in accordance with the work details.

If the owner intends to terminate the contract based on one of the defaults described above, he must firstly give the contractor a written warning notice which must specify the default. It is necessary that the contractor has no doubt about the default alleged, and the owner must give sufficient detail to identify the incident if there is any danger of confusion. If the contractor continues the default for 7 days after he receives such

notice, the owner may terminate this contract by serving the contractor another written notice of termination, which is to take immediate effect.

Upon termination of this contract, the owner will not be bound to make any further payment to the contractor until after completion of the works by another contractor. The owner is entitled to recover from the contractor any direct loss and/or damage caused to him by the termination of this contract, including his loss of liquidated damages, making good defects, employing a new contractor at a higher price and such consequential losses which are reasonably foreseeable as arising out of what is to be treated as equivalent to the contractor's breach of contract: *Wraith Ltd. V. P.H. & T. (Holdings) Ltd. (1968)*. There can rarely be any sum due to be paid to the terminated contractor.

2.16 Contractor's Right to Terminate the Contract

This contract sets out the following two grounds on which the contractor may determine the contract:

- The owner does not pay an amount properly due to the contractor without having good reason; or
- The owner prevents or obstructs the contractor from carrying out the works for a continuous period of at least 14 days.

Adopting the same procedure prescribed for the owner, the contractor must serve a written notice specifying the default. If the owner does not remedy the specified default within 7 days, the contractor may terminate this contract by giving the owner another written notice of termination, which is to take immediate effect.

Upon termination of this contract, the owner must pay the contractor within 14 days for the value of work properly executed and unfixed materials and goods ordered for the works.

2.17 Settlement of Disputes

Standard forms of building contract typically contain express provisions on dispute resolution methods, which may include negotiation, adjudication and arbitration. In view of the small amount of contract sum in decoration, repair and maintenance projects, the dispute resolution methods must be simple, fast and cost effective. This contract provides two alternative methods of dispute resolution: mediation and legal proceedings. If the contracting parties intend to resolve a dispute through mediation, they should agree a person to act as the mediator. If, for whatever reason, the parties cannot agree a mediator within 14 days of one party requesting the other to agree on a mediator, then either party can request the President of the Hong Kong Institute of Surveyors to appoint a mediator for them.

In addition to mediation, this contract also allows the contracting parties to take legal proceedings to settle their disputes. If the sum in dispute between the parties is less than HK\$50,000, the dispute can be resolved through the Small Claims Tribunal, whereby the parties can present their views and arguments themselves.

3 Conclusion

Historical development in building contract indicates that when the architect, engineer and surveyor face any claims or loopholes in a contract, they will amend the original condition or to add a special condition so as to avoid the same to be happened in the future. With this continued improvement, the relevant parties can have a better understanding of their duties, responsibilities and obligations under the contract. However, the contract has become more and more complicated. Apparently, this approach is not applicable to the building contract solely used by lay owners.

For a standard form of building contract used by lay owners, we have to consider the contractual provisions in an entirely different manner; for instance, how the parties incorporate the relevant documents to create a valid contract, how the traditional lump sum contract includes those works not expressed stated in the work details, how the owner relies on the contractor's skill and judgment in design, how the cost of a change can be agreed without valuation rules, how an extension of time is assessed and granted, how payment is evaluated and certified, what grounds the parties can terminate the contract, and how the parties resolve the dispute in a cost-effective manner, etc. All of these issues have been satisfactorily resolved in this contract. Finally, while this contract is written for use in Hong Kong, the various legal principles and underlying considerations are also applicable to other countries.

4 References

- Chappell, D. and Powell-Smith, V. (1999), *The JCT Design and Build Contract*, Blackwell, Oxford.
- Development Bureau (2010a), *Building Safety*, A paper submitted by the Development Bureau of the SAR Government to the Legislative Council for discussion on 23 February 2010.
- Development Bureau (2010b), *Building Safety Concerns and Enhancement Measures Related to the Ma Tau Wai Collapse Incident*, A paper submitted by the Development Bureau of the SAR Government to the Legislative Council for discussion on 27 April 2010.
- Eggleston, B. (1997), *Liquidated Damages and Extension of Time*, Blackwell, Oxford.
- Keating, D. (1978), *Building Contracts*, Sweet & Maxwell, London.
- Furmston, M. (2006), *Powell-Smith & Furmston's Building Contract Casebook*, Blackwell, London.
- Parris, J. (1985), *The Standard Form of Building Contract JCT 80*, Collins, London.
- Wallace, I. N.D. (1994), *Hudson's Building and Engineering Contracts*, Sweet & Maxwell, London.

Inhibiting and enhancing impingements on second year, undergraduate construction law education

Tim McLernon¹ and Robert Eadie¹

¹*School of the Built Environment,
University of Ulster, Newtownabbey, BT37 0QB
United Kingdom*

Email: t.mclernon@ulster.ac.uk; r.eadie@ulster.ac.uk

Abstract:

Higher education in the UK continues to progress through a changing operating environment. In particular, the economic environment of the past two years is likely to affect higher education. It has already impacted significantly on the construction industry and associated professions and indicators suggest a decline in the number of applications to built environment programmes. This paper reports on a specific element of a longitudinal study which aims to find out more about how best to facilitate the learning of undergraduate students in the operating environment that is likely to subsist over the next five years. The specific element is a study of second year, undergraduate students on a construction law module and the study will focus on those inhibiting and enhancing factors that impinge on the teaching, learning and assessment regime associated with this module. The study draws on data obtained from participant observation, minutes of meetings, discussions with colleagues and a dedicated survey of immediately past and present students of the module. The findings will be used to inform the design of a model for a programme structure and a teaching, learning and assessment regime that engenders graduates with individual, embedded knowledge, skills, abilities and attributes that allow them to better contribute to the built environment disciplines, to the growth of the economy and to societal needs.

Keywords:

Impingements, undergraduate, construction law, education.

1 Introduction

This paper reports on an element of a larger longitudinal study of higher education that will bring together a critical examination and analysis of current teaching, learning and assessment regimes in the built environment disciplines with an evaluation of learning traits, behaviours and attitudes of students in the built environment disciplines. The purpose of the overall study is to determine whether or not educational methods currently in use produce graduates with the individual, embedded knowledge, skills, abilities and attributes that allow them to excel in their contribution to the built environment disciplines in which they will ultimately make their career. In this paper, we shall use the term ‘students’ to refer to observed, common, generic traits, behaviours and attitudes of students. However, this is not to discount the fact that students have their own individual traits behaviours and attitudes. This paper is specifically concerned with

construction law education of second year, (level 5) students studying a module that has learning outcomes associated with:

- The legal systems of the UK;
- The law of Tort;
- The law of Contract; and
- Construction contracts.

The study was conducted to inform curriculum development of the subject in a rapidly changing operating environment. The design of the teaching, learning and assessment regime associated with this module is, as with the greater number of modules in higher education, a satisfactory compromise that meets the requirements of policy, strategies and stakeholders.

The two key research questions that are posed by this study are:

- (i) What do students think about the teaching, learning and assessment regime associated with this module?; and
- (ii) Using the outcomes of (i), what improvements to the teaching, learning and assessment regime associated with this module might be proposed?

2 The Operating Environment

UK Higher Education has changed significantly over, particularly, the last 25 years. The transformation from an elite to a mass system of higher education in the UK brought with it many demands of policy, practice and management that continue to challenge the system. Fifteen years ago, Eriksen (1995) discussed the changes needed in degree course management to create a system which would be sufficiently flexible and efficient to cope with more students without loss of quality. He advocated the application of the theory of total quality management to higher education, indicating the need to treat degree courses ‘as continuous production systems, rather than batch runs with standardized components and early detection of faults (unlearned work) to reduce quality costs (student failures and exam resits)’. He proposed, at that time, that the US higher education system should be used as a model for the UK.

Higher education in the UK continues to progress through a changing operating environment. The past decade has seen particular challenges arising for higher education in the UK that impact on the student experience. These include:

- Funding and top-up fees;
- Attracting overseas students to contribute to income;
- Global competition;
- Advances in ICT;
- On-line provision and competition;
- Blended learning delivery and assessment; virtual learning environments;
- The teaching-research nexus;
- Teaching quality assessments;
- The National Student Survey;
- Research assessment exercises and the current Research Excellence Framework;

- Criticisms of standards of UK higher education; and
- Student numbers.

In addition, there are two particular facets of this changing and challenging operating environment that have changed the nature of student learning in recent years. The first is that students now have to make a significant financial contribution to their studies. In order to fund this financial contribution, the greater proportion of undergraduate students works in employment on a regular basis. The second relates to the advances in information and communications technology which change the ways in which students socialise, source information and communicate with one another.

The policy of the recently replaced Labour government, that 50% of young people should enter higher education by 2010 has been, arguably, the key challenge over the last decade. It has brought with it a parallel challenge on how to fund higher education. Williams (2010) reports that the Association of Graduate Recruiters, which represents 750 employers, many of them blue-chip companies, are calling for the abolition of the Labour target of sending 50% of young people to university, justifying the call with the statement that: '[t]he government's strategy has driven down standards and devalued degrees'. Whether or not standards have, in fact, fallen is difficult to judge on the evidence. However, there is a prevailing perception and some concern amongst some academics that the UK degree is no longer the gold standard. This perception was recently expressed anonymously in an article in The Times Higher Education Supplement of 25th February 2010 in which a university lecturer suggested that 'poor student attendance' and 'generous exam marking' are 'making a mockery of higher education'. The thrust of the arguments of this lecturer is as follows:

'It is obvious that the current policy to increase participation in higher education has led to too many students attending who are not suited to studying at university level. And, as a consequence, standards fall. ... The crisis in university funding could be solved overnight by reducing participation to 20 per cent (from the current level of 43 per cent). Standards would rise, staff morale would rocket, and engaged students would get the education they deserve. ... Degrees would regain their value... .' (Anon, 2010).

In particular, the economic environment of the past two years is likely to affect higher education. It has already impacted significantly on the construction industry and associated professions and indicators suggest a decline in the number of applications to built environment programmes. The new Conservative and Liberal Democrat government alliance is expected to make significant changes and impact on higher education over the next five to ten years. Nevertheless, demand for places in higher education, and student numbers in higher education continue to rise. The media release from UCAS on 21st January 2010 reported that in 1999 there were 334,594 accepted applicants into university or college. Ten years later in 2009 there were 481,854 - an increase of 44%. This represents an increase of 25,227 (5.5%) on entry for 2008. There was an overall increase in the number of applicants of 8.7%. Applicants aged 20 and under showed a 6.9% increase. In this UCAS media release, Mary Curnock Cook, UCAS Chief Executive was quoted as saying:

'2009 saw an unprecedented demand for places at university or college, but significantly more students have been accepted into higher education than ever before. Whilst there have been increases across the board, our figures show that there has been a particularly large increase in applicants aged 25 years and over - 89,133 in 2009, compared to 77,286 in 2008 - a 15.3% increase (see Table [1]). Males aged 25 and over have seen the biggest rise in acceptances - up 10.8% to 20,963.'

Table 1. Applicants and acceptances by gender (all domiciles)
(Source: The media release from UCAS on 21st January 2010).

Year	Applicants			Accepted applicants		
	Male	Female	Total	Male	Female	Total
2009	284,757	355,103	639,860	218,185	263,669	481,854
2008	259,878	328,811	588,689	204,695	251,932	456,627
% change	9.6%	8.0%	8.7%	6.6%	4.7%	5.5%

She went on to report that:

'last year's cycle was record breaking for UCAS in terms of acceptances and applications. There are many factors that drive people to consider higher education, including the current economic situation, with more people looking to long term retraining in the traditionally more secure or transferable careers, such as nursing (+19.9%) and combinations within business and administrative studies (+11.4%).'

Table 2 below illustrates the increase in number of accepted applicants of UK domiciled students by age and by gender. It is noticeable that the greatest increase is in the category of male students of 25 years and over.

Table 2. Accepted applicants (UK domiciled) by age and gender
(Source: The media release from UCAS on 21st January 2010).

Year		Female	Male	Total
20 years and under	2009	177,831	148,943	326,774
	2008	172,925	141,643	314,568
	% change	+2.8%	+5.2%	+3.9%
21 to 24 years	2009	24,565	22,017	46,582
	2008	22,851	20,389	43,240
	% change	+7.5%	+8.0%	+7.7%
25 years and over	2009	33,248	18,459	51,707
	2008	30,893	16,323	47,216
	% change	+7.6%	+13.1%	+9.5%
Total	2009	235,644	189,419	425,063
	2008	226,669	178,355	405,024
	% change	+4.0%	+6.2%	+4.9%

There has been a significant, general increase in students from overseas over the past year illustrated in Table 3 below. Overseas students are attractive to universities because of the substantial fee that can be levied. However, overseas students can present challenges for the classroom and the overall student experience as well as for the pastoral care of those overseas students.

Table 3. Accepted applicants from overseas domiciled applicants by country of domicile - countries with 600 or more acceptances in the 2009 cycle

(Source: The media release from UCAS on 21st January 2010).

Country	2009	2008	Percent change
China	6,509	6,120	+6.4%
France	3,194	2,703	+18.2%
Ireland	2,823	2,609	+8.2%
Cyprus	2,610	2,305	+13.2%
Hong Kong	2,575	2,615	-1.5%
Germany	2,415	2,276	+6.1%
Malaysia	2,390	2,016	+18.6%
India	1,843	1,674	+10.1%
Nigeria	1,592	1,681	-5.3%
Greece	1,527	1,652	-7.6%
United States of America	1,330	1,191	+11.7%
Bulgaria	1,283	823	+55.9%
Poland	1,238	1,522	-18.7%
Singapore	1,187	968	+22.6%
Pakistan	1,081	1,307	-17.3%
Romania	976	630	+54.9%
Norway	959	769	+24.7%
Canada	914	791	+15.5%
Lithuania	895	682	+31.2%
Italy	873	691	+26.3%
Spain	869	724	+20.0%
Sweden	808	777	+4.0%
Saudi Arabia	772	530	+45.7%
Korea (South)	618	557	+11.0%
Belgium	605	534	+13.3%
All overseas	56,791	51,603	+10.1%

Table 4 extracts from UCAS those data related to Law and the Built Environment disciplines. It is noticeable that there has been a decrease in those applicants accepted in the Joint Academic Coding System subject of Building.

Table 4. Accepted applicants by subject; Law and Built Environment Programmes.
(Source: adapted from the media release from UCAS on 21st January 2010).

Joint Academic Coding System (JACS 2) Subject Line	Degree 2009	Percent change
M1 - Law by Area	17,527	+2.8%
H2 - Civil Engineering	4,483	+3.5%
K1 - Architecture	4,344	+3.5%
K2 - Building	3,536	-3.4%

In the design and delivery of the teaching, learning and assessment regime for a module or for a programme, attention must be paid to those key assessment elements that can impact on the quality and reputation of the module, programme and the institution. For example, the various teaching quality assessment exercises over the past fifteen years have impacted significantly on policies, strategies, procedures and methods associated with the management and delivery of higher education. The outcomes of these exercises suggest a hierarchical assessment of quality in the disciplines and in the institutions. Similarly, but perhaps more dramatically because of the direct relationship to funding, the various research assessment exercises over the past eighteen years have likewise significantly impacted on policies, strategies, procedures and methods associated with the management and delivery of higher education and which affect the student learning experience. Virtually all of research funding from HEFCE is distributed to higher education institutions according to the quality and amount of research done. The first research assessment exercise to cover the entire higher education sector was undertaken in 1992. More recently, the National Student Survey was introduced. The National Student Survey (NSS) aims to provide information on teaching quality and the student experience at higher education institutions across the UK. The NSS ran for the first time between January and April 2005; it was repeated between January and April 2006. Students in their final semester at a higher education institution are asked to give quantitative responses to 21 statements organised into six sections that deal with:

1. The teaching on this course
2. Assessment and Feedback
3. Academic Support
4. Organisation and management
5. Learning Resources
6. Personal development.

The students are additionally asked to quantitatively respond to the statement: ‘Overall, I am satisfied with the quality of the course’. A consequence of this survey is the production of ‘league tables’ of institutions based on the outcomes of the survey.

The student experience is central to higher education. It is imperative that the design and delivery of the teaching, learning and assessment regime is founded on students enjoying a good, robust learning experience. There is however a view that, probably because students now pay tuition fees, students are to be treated as ‘customers’. Such a view brings with it key challenges for a robust educational experience because knowledge is not a commodity that can be acquired passively. Little *et al* (2009) posit the following view:

‘Institutions view student engagement as central to enhancing the student experience, but *more emphasis seems to be placed on viewing students as consumers and rather less on viewing students as partners in a learning community* [authors’ italics]. For students’ unions, the emphasis tends to be on the latter aspect. Notions of students as ‘partners in a learning community’ seem to be stronger in certain subject areas than others. The study goes on to make a number of recommendations to lead the sector towards working more closely together when developing student representation practices.’

Viewing students as partners in a learning community should create a more robust institutional culture and ideology for advancing learning.

3 Data Collection and Analysis

The research context for the study is founded on qualitative data obtained from participant observation, policy documents, procedural codes of practice, discussions with academics and focus group discussions with students during tutorial sessions on a construction law module. This section of the paper deals with the quantitative and qualitative data obtained from a survey of two groups of students: one group who studied the construction law module this academic year (2009/2010) and the other group who studied the module the previous academic year (2008/2009). This construction law module, on which the survey was carried out, is a level 5 module taken by second year students studying on three different programmes, namely: BSc (Hons) Construction Engineering and Management; BSc Construction Engineering; and BSc (Hons) Architectural Technology and Management. The survey was conducted to elicit from students their views on key areas of the construction law curriculum. The identified areas were constructed from a priority list prepared from issues raised in several fora, including course committee meetings, School Board meetings, Faculty Teaching and Learning Committee meetings and other fora, and including discussions with colleagues internally and externally. These identified areas were considered by us, the authors, to be key, inhibiting and enhancing impingements on second year, undergraduate construction law education. They are as follows:

- Attendance,
- Class size,
- Preferred attendance times,
- Amount of contact time,
- Assessment methods,
- Modes of delivery,
- Relevance to the construction industry,
- Using electronic learning resources and support,
- Likes and dislikes about the module.

3.1 Methodology

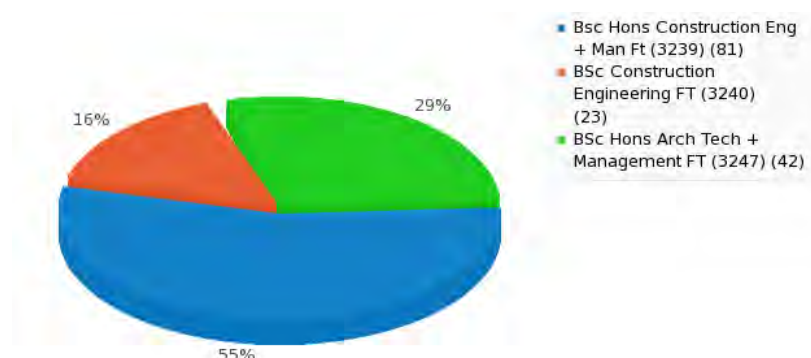
Data collection was carried out in 2010 using Limesurvey™, a system similar to that used by Solomon (2001). This software allows responses to the web-based interface to be collected and stored in an on-line MySQL database. Simple statistical analysis can be carried out internally without the need to retype data. For the purposes of this survey data was exported directly to Microsoft Excel™ for additional in-depth analysis and confirmation purposes.

3.2 Sample size

To be statistically viable Bartlett *et al* (2001) suggest sample numbers for various population sizes. These numbers are based on type of data required and a 65% response rate. For a population size of 300 Bartlett *et al* (2001) suggest that the sample size must be 143 to provide an alpha margin of error of 0.05 (finding a difference that does not exist in a population), a p of 0.5 and a t of 1.65. p is the probability and t is the test statistic. A possible error level of 5% (Krejcie & Morgan, 1970) is statistically acceptable for relationships and differences. The 302 student population, creating a sample size of 143 and requiring a 65% response rate, therefore requires 93 responses to result in a statistically defendable result. In order to ensure that this value was exceeded oversampling took place with questionnaires sent to the complete population. One hundred and forty six (146) responses were obtained well above the statistical requirements of Bartlett (2001). Rubin and Babbie (2004) and OIG (1997) indicate that a 50% response rate is acceptable. The response rate was 48%, minimally below the 50% threshold. However, Holbrook *et al.* (2005) examined surveys with 4 to 70 percent response rates and concluded “*lower response rates do not notably reduce the quality*”.

3.3 Sample Breakdown

Figure 1 identifies the breakdown of the responses. The question asked appears in the top row of each table. It can be seen that the response rates from the various different courses involved in the module range from 41 to 58%. This shows that “buy-in” from the different courses occurred.



On which of the following courses are you enrolled?				
Answer	Count	Percentage	Possible	Percentage of course total
Bsc Hons Construction Eng + Man Ft	81	55.48%	174	47%
BSc Construction Engineering FT	23	15.75%	56	41%
BSc Hons Arch Tech + Management FT	42	28.77%	72	58%
Overall Response Rate	146		302	48%

Figure 1 Breakdown of Sample

3.4 Attendance

Chan and Kumaraswamy, (1996) provide a methodology to determine the importance of each of a number of factors. They call this the Relative Importance Index(RII).

$$RII = \sum w / (A \times N)$$

Where w = Weighting given to each factor, A = Highest Weight (6) and N = Number of Respondents. In the question below w is 1 for 0-2 weeks up to 6 for 11-12 weeks, A is 6 and N is equal to 146. A similar methodology is used throughout the paper wherever RII is calculated.

RII was calculated for the lectures, tutorials and seminars in the current mode of delivery. It can be seen from Table 1 that attendance at lectures was deemed most important with a value of 0.76, closely followed by tutorials at 0.75 and lastly by seminars at 0.71.

Table 1 Attendance Statistics

Out of 12 weeks, how many weeks did you attend	Lectures		Tutorials		Seminars	
Answer C	Count	%	Count	%	Count	%
0-2	3	2.05%	4	2.74%	9	6.16%
3-4	9	6.16%	12	8.22%	9	6.16%
5-6	11	7.53%	17	11.64%	25	17.12%
7-8	40	27.40%	30	20.55%	29	19.86%
9-10	46	31.51%	42	28.77%	36	24.66%
11-12	37	25.34%	41	28.08%	38	26.03%
Relative Importance Index (RII) value	0.76		0.75		0.71	
Mode	9-10		9-10		11-12	

It can be further seen from the statistics 84% attended more than half the lectures, 77% attended more than half the tutorials and 70% attended more than half the seminars. These statistics vary significantly from the tutor recorded attendance that asserts lower attendance figures. The students' perceive that their attendance is more regular than their actual attendance.

3.5 Usefulness of modes of delivery

A similar procedure was used to determine the usefulness of the lectures, tutorials and seminars in terms of delivery mode. It can be seen from Table 2 that tutorials were considered to be the most useful mode of delivery with a value of 0.68, followed by lectures with 0.66 and seminars at 0.61.

Table 2 Usefulness of Delivery mode

How useful did you find:-	Lectures		Tutorials		Seminars	
Answer Scale 1 not useful, 5 very useful. C	Count	%	Count	%	Count	%
1	6	4.11%	12	8.22%	19	13.01%
2	17	11.64%	21	14.38%	22	15.07%
3	56	38.36%	36	24.66%	49	33.56%
4	58	39.73%	52	35.62%	45	30.82%
5	9	6.16%	25	17.12%	11	7.53%
Relative Importance Index (RII) value	0.66		0.68		0.61	
Mode	4		4		3	

Each respondent was asked to state the class size for each of the modes of delivery that most effectively assisted their learning. Modal size preferred for lectures was 21-30. The preferred size of tutorials was divided equally between those who wanted 2-10 and those who considered 11-20 as the optimum size. The optimum size for seminars was considered to be the latter of these two groups at 11-20.

As the actual sizes of classes were much larger than the stated sizes this question was followed by one asking if they considered sizes greater than these to be detrimental to their learning. The majority in each case considered this to be the case: Lectures 79%, Tutorials 84% and Seminars 63%. The extent of the disruption to learning by larger class sizes was then determined from those who considered their learning to be disrupted.

Table 3 Disruption due to class sizes

Disruption due to larger class sizes:-	Lectures		Tutorials		Seminars	
Answer Scale 1 little disruption, 5 large disruption.	Count	%	Count	%	Count	%
1	0	0.00%	1	0.82%	0	0.00%
2	4	3.48%	4	3.28%	3	3.23%
3	28	24.35%	29	23.77%	25	26.88%
4	51	44.35%	50	40.98%	36	38.71%
5	23	20.00%	28	22.95%	23	24.73%
Relative Importance Index (RII) value	0.714		0.714		0.731	
Mode	4		4		3	

Table 3 indicates that students considered class sizes larger than their preferred values would adversely affect the seminars more than the lectures or tutorials. The lectures and tutorials would be affected equally.

3.6 Mix and timing of modes of delivery

An in-depth analysis was then carried out into the way in which the module was being delivered. Responses proposed an average of 4.53 hours per week contact time with a standard deviation of 4.04. Students currently on the module have 4 hours per week so this substantiates the *status quo*. Eighty-four percent (84%) considered that they had enough contact time with teaching staff. Sixteen percent (16%) suggested they required extra contact time. These were further probed as to where this extra contact time should be allocated. It can be seen from Figure 2 that 39% suggested it should be added to lecture allocation, 36% to tutorial allocation and 25% to seminar allocation.

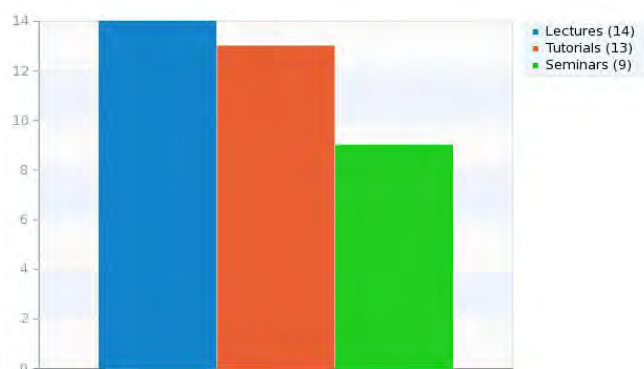
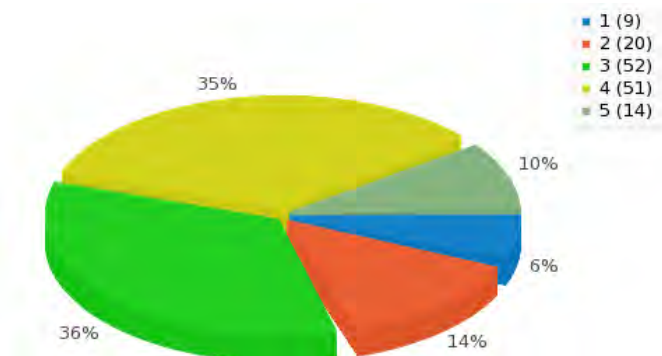


Figure 2 Allocation of extra contact time.

Currently contact time is spread throughout the week. A further question resulted in 50.68% suggesting that they would like it condensed into a single day with the remainder happy with the current provision.

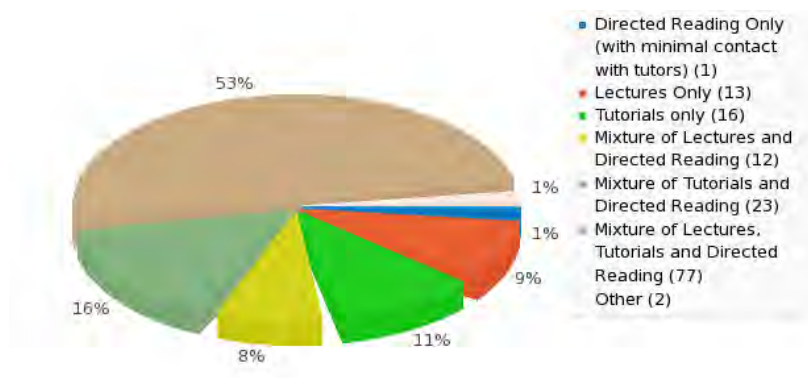
It can be seen from Figure 3 verification that the mix of lectures, seminars and tutorials suits 45% of the cohort, with 35% remaining neutral and only 20% feeling it did not help them learn effectively.



Do you consider the mix of lectures, tutorials and seminars to be a format that helps you learn effectively?		
Answer 1-Not very 5-Substantially	Count	%
1	9	6.16%
2	20	13.70%
3	52	35.62%
4	51	34.93%
5	14	9.59%

Figure 3 Learning effectiveness of lectures, tutorials and seminars

A question followed which probed some other delivery options available and left an option as to state alternatives. Two students suggested that a mixture of lectures and tutorials would be the best option. The rest suggested the options indicated in Figure 4. The current system was already established as the preferred option and Figure 4 indicates the second preference to be a 'Mixture of Lectures, Tutorials and Directed Reading'. Only a minority held other views.



Would you prefer any of the following formats?		
Answer	Count	Percentage
Directed Reading Only (with minimal contact with tutors)	1	0.69%
Lectures Only	13	9.03%

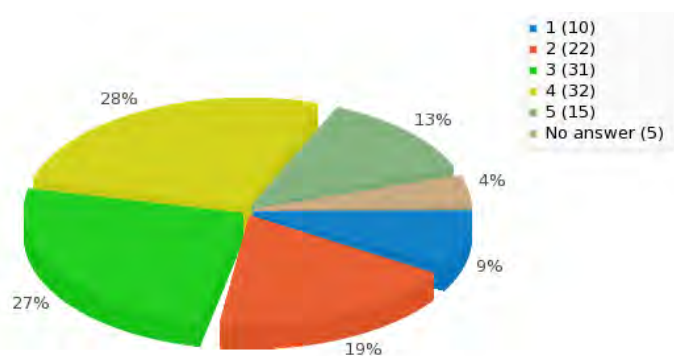
Tutorials only	16	11.11%
Mixture of Lectures and Directed Reading	12	8.33%
Mixture of Tutorials and Directed Reading	23	15.97%
Mixture of Lectures, Tutorials and Directed Reading	77	53.47%
Other	2	1.39%

Figure 4 Format Preferences

3.7 Relevance to the construction industry

The next section was limited to responses from those who had worked in the construction industry. This may have been through summer jobs, placement or prior to entry to university. This equated to 115 of the 146 (79%) who completed the sample. Of these 115, (83%) considered that the content of the module met the needs of those forming a career in the construction industry. They were also asked to suggest other topics which may be helpful to those going into the industry. A variety of answers were received. Most were already covered during the module, possibly in weeks those suggesting them had missed. It was noted that some considered that it was too general and should have been more specific. Two suggested that JCT conditions of contract should have been taught alongside the NEC conditions to give a broader perspective.

Figure 5 confirms the module's relevance to the construction industry. It indicates that 41% of students considered it relevant to their experience in the construction industry, with 28% considering it lacked relevance and the remainder remaining neutral or refusing to answer. This was not a mandatory question and five chose not to respond.

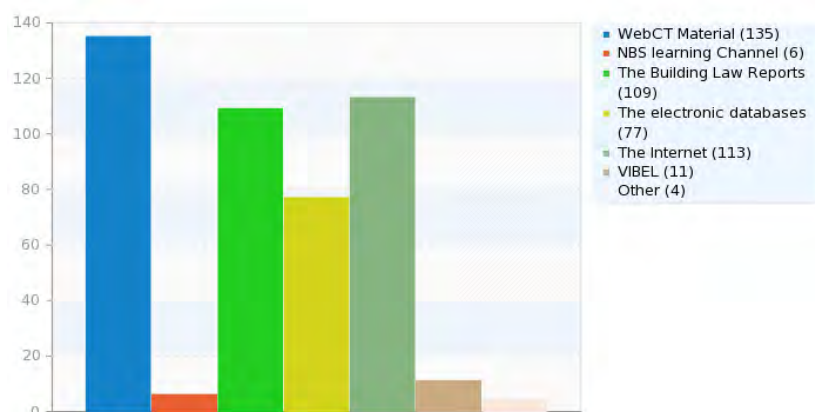


On a scale of 1-5 rate how the curriculum of this module was of relevance to your work in the construction industry?		
Answer 1-Little relevant 5- very relevant	Count	%
1	10	8.70%
2	22	19.13%
3	31	26.96%
4	32	27.83%
5	15	13.04%
No answer	5	4.35%

Figure 5 Module content relevance to the construction industry

3.8 Use of electronic resources

Although electronic resources were widely used throughout the module, 83% would not prefer these as a replacement to classroom teaching. This may emphasise the importance that students attach to personal contact between lecturers and students. It can be seen from Figure 6 that the most widely used resource was the lecture notes and other material provided through WebCT, the online learning environment. This was followed by the internet and the building law reports.



What electronic resources did you use to help your learning during this module?	Count	Percentage
Answer		
WebCT Material	135	92.47%
NBS learning Channel	6	4.11%
The Building Law Reports	109	74.66%
The electronic databases	77	52.74%
The Internet	113	77.40%
VIBEL	11	7.53%
Other	4	2.74%

Figure 6 Use of electronic resources during the module.

Students were asked to rank the usefulness of the electronic resources provided on the module. The ranking initially mirrored the use of the resources with WebCT ranked in first place. However, the Internet and the building law reports swapped places, with building law reports ranked second and the Internet third. Little use was made of the resources on the NBS learning channel and the Virtual Built Environment Library (VIBEL).

Table 4 Ranking of usefulness of electronic resources on the module

Please place the electronic resources listed below in rank order from "most helpful to your learning" to "least helpful to your learning" during the module?	RII	Rank
Answer		
WebCT Material	0.91	1
The Building Law Reports	0.72	2
The Internet	0.69	3
The electronic databases	0.61	4
NBS learning Channel	0.29	5
VIBEL	0.28	6

3.9 Analysis of assessment methods used

Similar to the electronic methods of teaching, 63% considered electronic methods of assessment less preferable than traditional forms. The remainder of the sample provides reasons for the preference for electronic assessment, set out in Table 5. No definite conclusions could be drawn from these responses.

Table 5 Responses as to why electronic assessment is preferred [extracted verbatim from students' responses]

4	less pressure
6	It is alot more relaxing and does not have same stresses involved as an exam has.
20	I felt that when we had webCT test there wasnt as much pressure on people thereofre they may preform better under less stress
37	I thought they were a less pressurising method of assessment.
40	less writing involved gives more time to answer the questions set
43	The WebCT class test was quite enjoyable to do because 1. it gave me a good idea of what I did and didn't know 2. the results were back fairly quickly 3. I enjoy doing computer assessments and quizzes... except for this one that is!!
47	MULTIPLE CHOICE IS GOOD
49	They're alot less stressful and seem more organised. Also less chance of having a bad marker or loosing your script.
60	Used to working more with electronic methods
67	They are a quick and easy means of testing a students capability.
68	I believe they are easier to understand and carryout. I liked the idea of the assessment being on WebCT as i feel there wasnt as much pressure so I could concentrate easier rather than be under strict exam conditions.
82	Tests are completed faster and submitted for marking with limited effort.
84	more independent learning
85	I feel pressured in written tests and they are never a true reflection of my true ability.
86	I feel that some people may be disadvantaged when it comes to speed of completing assessments (written)
89	Impartiality.
100	Carry out assessment in a quiet environment
101	more relaxed approach, felt under no pressure for webct exam for heather!
102	Maybe just as effective and saves on travel to campus to complete a single assessment
103	Its esier slightly less demanding, but you cant waffle
104	more relaxed atomsphere, think clearer.
106	can work anywhere
113	Instant results
116	No need to come to Jordanstown
118	Easier to look over where you have gone wrong and quicker feedback.
120	there is not as much pressure

121	it makes it easier to change parts that have already been written, and if part of a webtest the opportunity to breakdown modules into questions that can be more easily understood
122	Easier to complete and a lot less time consuming considering work for other modules.
123	it means you can access the assessment from your own home instead of travelling for 2 hours to do a 2 hour exam and then drive home again!
133	i feel underpressure by paper assessments
147	.
152	if they have multiple choice?
158	Find it more beneficial
161	Even though the presentations etc. was useful they were very time consuming and affected the time i had for other modules.
167	Doesnt seem to be as much pressure as in written assessments
168	helps save the enviroment from unneeded paperwork
171	The Assessments in relation to examinations, I feel, adds more pressure by carrying it out within an exam hall than using a electronic method.

The continual assessment process within the module was the preferred assessment method. This was substantially the highest ranked method of assessment with a value of 0.83. The next highest value was 0.65 for 'Continual Assessment during the module and end Examination'. The least preferred method was 'no assessment'. This may indicate that the students understand the value of assessment as a learning tool which adds value to their qualification.

Table 6 Ranking of the most effective methods of assessment to promote learning

Please rank the methods of assessment that would help you learn most effectively in general if used in this module?	RII	Rank
Answer		
Continual Assessment during the module	0.83	1
Continual Assessment during the module and end Examination	0.65	2
End examination of open book format	0.64	3
End of module case-study	0.61	4
Unseen, written, end of module Examination	0.41	5
No Assessment	0.36	6

Six methods of assessment are currently being used within the module. These are indicated in Table 7 below. It can be seen that the final case study assessment is most successful in promoting learning. The first five assessment methods below are quite closely grouped in order of importance with the sixth method – the reflective log – lagging far behind.

Table 7 Specific to the module, effectiveness of assessment in promoting learning

Specifically in relation to the module, how effective were each of the following assessment methods in helping you to learn	RII	Rank
Answer		
The final Case Study Assessment	0.74	1
The on-line WebCT Law assessment	0.71	2
The Individual Seminar on a building law case	0.67	3
The Duo Seminar (carried out in pairs and focusing on Contract or Tort)	0.65	4
Composing a contract exercise	0.63	5
The Reflective Log	0.43	6

3.10 Section of module liked and disliked most

This section of the survey produced a lot of contradictory answers. What some students enjoyed, others disliked. Some of the responses to this section were illogical and lacked focus. Some took the opportunity to provide feedback on other elements of the course, the institution and other matters of concern to them but not relevant to, or associated with this module. For this reason these two questions did not provide substantive feedback on the module and consensus of opinion could not be achieved from the responses.

4 Conclusions

From the data analysis, the students' responses relating to the identified inhibiting and enhancing 'impingements' were considered within the context of the two research questions, i.e.:

- (i) What do students think about the teaching, learning and assessment regime associated with this module?; and
- (ii) Using the outcomes of (i), what improvements to the teaching, learning and assessment regime associated with this module might be proposed?

The conclusions for this paper focus on two unexpected aspects of the study:

- (i) The first relates to a specific area of the survey, namely, student attendance.
- (ii) The second is more general in nature. The expectation, prior to the survey, was that student responses would suggest radical changes to the teaching, learning and assessment regime of this construction law module. However, with some exceptions, the students' responses suggest that the status quo should prevail.

4.1 Student Attendance

There is a general concern across higher education that students are not attending classes as they should be. There is a concern that standards may be detrimentally affected as a consequence of general bad attendance. Construction law is a textual subject that requires explanation, debate, discussion and real life scenarios to help understanding. The classroom environment is best for these learning support methods. The survey suggests that students agree that the various

classroom mechanisms used to support and enhance their learning do, in fact, support and enhance their learning. The attendance requirement is four hours per week for 12 weeks; equating to 48 hours contact time. The notional learning time requirement for such a module is 200 hours; thus leaving 152 notional hours for private study, research and completing assessments. The survey indicates that only, approximately, 25% of the students stated that they attended 11 or 12 weeks. Students generally perceive their attendance at classes to be sufficient. The statistics gathered in the survey in respect of attendance vary significantly from the tutor recorded attendance that asserts lower attendance figures. The students' perceive that their attendance is more regular than their actual attendance.

4.2 Students are satisfied with the status quo?

A generic observation and interpretation of the data collected from this student survey is that students don't offer any radical criticism or suggest any radical changes to the teaching, learning and assessment regime of this construction law module. There are suggestions of some tinkering around the edges of the module design, but, generally, the suggestion from the data is to leave as is. There are probably two (at least) answers to this. The first is that students are satisfied with the teaching learning and assessment regime and have no radical changes to offer. The second is a more contrived interpretation of the phenomenon. Second year students have limited contextual knowledge of higher education and the system associated with higher education. Educational policy and curriculum design are not of fundamental importance to the student. The key focus of students' attention is on the timetable, the assessments and what free time is available. The ultimate attainment of the degree is important, but possibly only for the exchange value of the degree with learning and understanding secondary to these matters. Consequently, students want to be given these learning parameters and will work within the set parameters.

The answer to the first research question: "What do students think about the teaching, learning and assessment regime associated with this module?" is probably that they accept what has been designed by the 'experts'. The answer to the second research question: "Using the outcomes of (i), what improvements to the teaching, learning and assessment regime associated with this module might be proposed?" is that no significant improvements were proposed and the reason is probably that students do not see themselves as key stakeholders or authorities in curriculum design. Control of inhibiting and enhancing impingements on second year, undergraduate construction law education lies with the academics and the higher education system.

5 References

- Anon, (2010), 'Wise up to the Naked Truth', article in *The Times Higher Education Supplement* (25th February 2010).
- Bartlett, J., Kotrlik, J, and Higgins,C, (2001), 'Organizational Research: Determining Appropriate Sample Size in Survey Research', in *Information Technology, Learning, and Performance Journal*, **Volume 19**, (1), Spring 2001, pp 43-50.
- Eriksen, S,D., (1995), 'TQM and the Transformation from an Elite to a Mass System of Higher Education in the UK', in *Quality Assurance in Education*, **Volume 3**, (1), pp14-29.
- HEFCE, (2010), 'The Higher Education Workforce Framework 2010', HEFCE Issues paper, February 2010/05a, http://www.hefce.ac.uk/pubs/hefce/2010/10_05a/#exec, viewed 12/05/2010.

- Holbrook, A, Krosnick, J, and Pfent, A., (2007), 'The Causes and Consequences of Response Rates in Surveys by the News Media and Government Contractor Survey Research Firms', <http://comm.stanford.edu/faculty/krosnick/docs/TSMII%20chapter%20proof.pdf>, viewed 2/05/2010.
- Krejcie, R, and Morgan, W, (1970), 'Determining sample size for research activities'. In *Educational and Psychological Measurement*, **Volume 30**, pp 607-610.
- Little, B., Locke, W., Scesa, A. and Williams,R., (2009), '*Report to HEFCE on Student Engagement*', Centre for Higher Education Research and Information, February 2009, The Open University, p4.
- Office of the Inspector General(OIG) (1997), 'Audit of the Office of Program and Integrity Reviews' Special Studies',http://www.ssa.gov/oig/ADOBEPDF/audit_https/96-51142.htm, viewed 3/12/2008.
- Rubin, A, and Babbie, E, (2004), '*Research Methods of Social Work*', Thomson Wadsworth, UK, pp 288-289.
- Solomon, D., (2001), 'Conducting Web-based Surveys', *Practical Assessment, Research & Evaluation A peer-reviewed electronic journal*, <http://pareonline.net/getvn.asp?v=7&n=19>, viewed 20/05/2010
- UCAS,(2010), 'Media Release; 21st January 2010'; http://www.ucas.ac.uk/about_us/media_enquiries/media_releases/2010/210110; viewed 11/04/2010.
- Williams, R., (2010), 'Abolish Labour Target of Sending 50% to University, Report Urges', The Guardian,<http://www.guardian.co.uk/education/2010/mar/09/abolish-50percent-target#history-link-box>, viewed 09/04/2010.

Dispute Resolution alternatives: Problems, Preference and Process

Carl van Zyl¹, Prof Basie Verster¹, Stephan Ramabodu¹

¹Department of Quantity Surveying and Construction Management
University of the Free State, Bloemfontein,
South Africa

Email: vZylCH@ufs.ac.za; VersterJ@ufs.ac.za; RamaboduMS@ufs.ac.za

Abstract:

Purpose of the paper: This paper serves to identify some professionals' opinions of various dispute resolution methods, the problems related to project communication which may cause differences and disputes in respect of construction projects and suggested processes.

Methodology/Approach: Two studies were conducted. A case study (Study 1) was undertaken with 10 construction professionals who worked on a large shopping centre development project, which experienced problems with communication and disputes. This study enabled the identification of communication problems that may lead to disputes in answering problems, preference and process questions.

Study 2 was a follow-up to previous studies. Part of this study was also conducted to further test opinions regarding dispute resolution processes and preferences as well as the reasons for selecting Alternate Dispute Resolution (ADR) methods. A captured respondent group of 26 quantity surveyors were requested to respond. The response rate was 100%.

These studies enabled the researchers to identify communication problems, preferences as well as methods and procedures that may assist project teams in resolving disputes in an amicable manner. Various dispute resolution methods were analysed and reviewed to determine problems, processes and preferences related to dispute generating issues and Alternative Dispute Resolution methods.

Results: Results enabled the researchers to identify an approach to the methods that may be used to influence effective dispute resolution between clients and contractors. The studies also identified communication problems and elements of good dispute resolution processes and strategies. Structured communication management and continuing interaction are basic elements to avoid disputes; however, the implementation of a tested dispute resolution method and process is important.

Value: The results of this study may assist professionals and contractors by enabling them to identify problems and to select the most appropriate processes and forms of dispute resolution methods. These processes or forms of dispute resolution methods should be agreed upon in cases where differences and disputes arise and where the settlement of these disputes has to be anticipated.

Key words: Claims, Communication, Construction Industry, Disputes, Professions.

1 Introduction

The beloved South African attorney, politician and writer, C.J. Langenhoven, quipped a century ago “Litigate for a jacket, but keep the trousers ready for the legal costs”, admitting then that the litigation route is a costly process and should be avoided if possible (Translation from Scanell, 1993:49).

It is proposed that a closer look at Alternative Dispute Resolution methods (ADR) should be considered. They are less adversarial, time consuming and costly, and may become more valuable (Finsen, 2005:221-222).

The objectives of this study are to firstly identify some dispute resolution problems and then to analyse ADR methods in use and investigate the preferences in order to understand which processes and methods may be used to resolve construction disputes.

2 The goals of claims and dispute resolution

The goals of claims and dispute resolution are firstly to establish the right of any party to submit a claim, and secondly to enable the other party to consider the claim in terms of its validity, contractual terms and possible outcome (Verster, 2006; Chappell, Powell-Smith & Sims, 2005:3).

Lodging or considering a claim does not mean that a dispute exists, but should the rejection of a claim occur, or agreement not be reached, or a different interpretation of a claim or opinion exist, one has to realise that a dispute may then be lodged. Dispute resolution should then assist the parties in resolving such an impasse in a cost effective, satisfactory and timeous manner (Verster, 2006).

An alternative dispute resolution method, like mediation was developed for various reasons, but mostly because the traditional processes, for instance litigation, were seen to be less favourable for the following reasons:

- Cost – Attorneys, senior council and expert witnesses, all contribute to heavy costs being incurred by both parties;
- Time – long waiting time for court dates where cases were often postponed for on-site visits, calling for expert witnesses, etc.;
- Most magistrates or judges are not specialists in the field of construction; and
- The outcomes often cause more damage.

(Finsen, 2005:221-222; McKenzie & McKenzie, 2009:251-254; Mackie, 1991:66).

3 The processes: Alternate Dispute Resolution Methods

The following are some of the ADR methods used in the construction industry, all having the common goal to resolve disputes voluntarily and initiated by the parties themselves:

- Agent resolution;
- Informal discussion;
- Negotiation;
- Mediation;
- Conciliation;
- The mini trial;
- Engineering expert assessment;
- Adjudication;
- Dispute Review Board;

- Partnering; and
- Dispute Resolution Advisor System (Tiruneh, Verster & Kotzé, 2007; Verster, Ramabodu & Van Zyl, 2010).

3.1 Choice of ADR Methods

For the purpose of this paper, the following five methods were investigated:

- Agent resolution;
- Adjudication;
- Conciliation;
- Mediation; and
- Arbitration.

The various ADR methods are briefly discussed to show the background of the elements and methods covered by the surveys and what was understood by each method under consideration.

3.2 Agent Resolution (expert resolution)

Traditionally, in the South African context, agent resolution (usually the architect) was the first stage towards resolving differences and disputes. The resolution by this agent was final and binding unless disputed within an agreed period (Quail, 1978:11-12, 166; Finsen, 2005:32-33). In terms of the Joint Building Contracts Committee (JBCC) (2004:30), the principle agent shall give a decision, on request by the contractor, should there be any disagreement between the parties. Such a decision shall be final and binding unless timeously disputed. This clause has however been removed in the latest editions of JBCC Series 2000 (JBCC, 2007:30)

3.3 Adjudication

Adjudication is an accelerated form of dispute resolution in which a neutral, impartial and independent third party deals with the dispute as an expert and not as an arbitrator, and whose determination is binding unless and until invalidated or overturned by an arbitration award (Hibberd & Newman, 2001). The adjudicator shall not advise the parties or their representatives regarding any aspect of the agreement in respect of which he has been appointed other than in accordance with stated rules (JBCC, 2005: cl. 1-7)

The adjudicator's written determination of the dispute shall:

- be delivered to the parties, and
- outline reasons for his decisions (JBCC, 2005, cl. 6 -7).
- The decision is final and binding until and unless reversed by an arbitrator (JBCC, 2007:30)

3.4 Conciliation

Conciliation involves a process of bringing disputing parties together in a forum to investigate the problem and assist the parties to formulate their own solutions; the conciliators may also be requested to formulate their own opinion.

The parties decide who the conciliator will be. The conciliator should, however, be a person with good communication skills and relevant knowledge (Loots, 1991:8-13).

3.5 Mediation

Mediation means different things to different people, but in the construction industry, it usually denotes a procedure in which a neutral third party seeks to resolve a dispute between contracting parties through mutual agreement, by conducting an enquiry, similar to arbitration, but less formal and by giving a non-binding opinion. The parties represent themselves without calling in legal professionals. The mediator should know the details of the dispute and should give each party the opportunity to state their case. The mediator should decide on the best procedure, based on circumstances (Eilenberg, 2002; Finsen, 2005:230-232; Hibberd & Newman, 2001; McKenzie & McKenzie, 2009:174; Quail, 1978:11-12).

Quantity surveyors often perform mediation tasks for clients or other parties, be it informal as a quantity surveyor, or a formal mediator by appointment. However, in terms of many contracts, JBCC, the parties shall agree on the appointment of a mediator and meet with the mediator in an effort to reach a settlement. If a settlement is reached, the mediator shall record such an agreement which shall become binding on the parties on the signing thereof (JBCC, 2007:31).

3.6 Arbitration

In some countries, arbitration is a process provided for by an act of law, adopted by parties through mutual agreement, stipulating that they will submit any dispute that may arise between them to the impartial judgement of some third party of their choice and that the award by this impartial person will be final and binding. Arbitration is not a new process; in fact, it was known to the Romans, used by the Dutch and English in the days of colonial expansion and is currently widely used in the construction industry and further afield (Finsen, 1999: 203-204).

Binnington (2005:49-50) suggests approaching an arbitrator rather than a lawyer. He further mentions the importance of securing a competent arbitrator, one well acquainted with the process of arbitration.

In South Africa, arbitration is regulated by an act of parliament (South Africa 1969, Act 42).

Arbitration is a more formal process than other dispute-resolution processes mentioned earlier, but arbitration has many advantages. Some of these are:

- Expert knowledge of a selected arbitrator;
- Possible savings in legal representation costs;
- Flexibility of the process;
- The decision is final and binding;
- Time and money are saved; and
- Arbitration is a private matter (Butler & Finsen, 1999; Eilenberg, 2002; Finsen, 2005:216-217; McKenzie & McKenzie, 2009:235-236).

Arbitration is sometimes also criticised as being only marginally quicker than litigation, especially where FIDIC documents are used (Binnington, 2005:49-50).

4 The studies: Problems and Preferences

Two different studies related to communication, problems and dispute resolution methods were undertaken by the University of the Free State and evaluated to identify problems related to communication and disputes and to identify construction professionals' opinions of various processes and methods available to solve these problems. These studies followed previous surveys in respect of the same problems, but were also aimed at understanding the problems related to communication that lead to relationship disruptions and disputes.

4.1 Study 1: Case study related to communication problems

Study 1 was a case study conducted with 10 quantity surveyors. The case study project (shopping centre) investigated was not unique in respect of problems related to communication, communication instruments, claims and counter claims, and differences of opinion. Due to many communication problems, differences, claims and disputes related to the project, it was however seen as a good example to investigate.

The original contract was concluded between the main contractor, an international company, and a client who operates throughout South Africa. The amount involved was about ZAR500 million. This

amount was based on the provisional bills of quantities method, of which about 50% of costs were for building and structural work and 50% for specialist installations included in the contract sum as provisional sums. A project manager was appointed after the production process commenced and communication systems were largely changed by him. Many sub- and direct contracts had to be concluded with specialists. The complexity of many contracts and changes in communication systems added to the difficulties.

The research project aimed at analysing the opinions and perceived frustrations of the quantity surveyors who, on behalf of the client, had to develop tender and contract documents, cost reports, notifications, payment valuations and general cost and cost planning communications during the design and construction process.

The 10 quantity surveyors who were involved in the project in various capacities were interviewed, and responded to a specific questionnaire related to the main aspects that were identified as quantity surveying elements. These aspects are: communication instruments (13 questions) and the effectiveness of these instruments (8 questions).

The responses to the questions enabled the research group to identify the most serious problems and suggested solutions. For this paper the most serious problems are identified.

Table 1 shows the identified communication instruments, related to this project, that were considered as the most serious problem areas.

Table 1: Results in respect of some identified instruments used in the case study

Source: University of the Free State, 2007. (Ratings: 1= low, 3= intermediate, 5 = high)

OPINIONS ON COMMUNICATION INSTRUMENTS USED	Average Result 1-5	%
• Employer's instruction	1.6	34%
• Packages (identification of work packages by project manager)	2.1	42%
• Claims communication	2.4	48%

The 10 respondents' opinions (100% response) on the communication instruments, used during the project, were tested (on a scale of 1 = poor and 5 = excellent).

It is clear that the employer's instructions and the identification of packages (tenant installations, refuse removal, shop-fronts, etc.) lead to disruption and eventual disputes linked to pricing, delays, uncertainty and frustration. Notification of claims was also, according to the survey, not managed well and timeously.

Table 2 shows the results in respect of general communication elements. It is also clear that the following communication processes performed below average in respect of expectations of quality procedures, and thus also influences effective management and claims communication.

Table 2: Results in respect of all general communication elements

Source: University of the Free State, 2007. (Ratings: 1= low, 3= intermediate, 5 = high)

COMMUNICATION ELEMENT STATEMENTS	Average Result %
• Drawing distribution was managed well and on time	46%
• Professional team had an effective relationship within the team	48%
• The Project manager's communication was effective and well managed	44%
• Owner/professional team relationship was managed effectively	32%
• Time and programming was managed well	34%
• Budget was clear and met the owner's strategy	44%
• Professional agents empowered to do their work well	40%

The respondents were clear on their opinions related to the effectiveness and quality of general communication in respect of the specific project and the problems identified. (Tables 1 and 2). The overall opinion was that these problems were not identified early and managed well, leading to disputes.

Clear problems were thus identified. The elements that stand out in respect of communication disruptions are:

- Site, employers and contract instructions;
- Identification of work packages by project managers;
- Drawing distribution and design availability;
- Professional teams' relationships;
- Project Manager's project communication;
- Owner / professional team relationships and empowerment;

- Time and programme management was not done well; and
- Strategy of budget and scope management.

The current situation regarding the project is that the final account cannot be settled and many claims and resolutions are outstanding. Many claims for delays are not resolved. The above-mentioned communication disruptions and problems influenced communications negatively. A clear communication strategy was absent and this contributed further towards the unacceptable state of affairs. These results show not only the importance and need for a structured communication process to avoid disputes, but also a model to deal with disputes once disputes are evident. This model should enable stakeholders to identify problem areas early, manage these problems through a tested process and taking the preference of the parties into account.

4.3 Study 2: Quantity Surveyors and ADR reference

Study 2, conducted among 26 quantity surveyors, aimed at establishing the preference rating of various dispute resolution methods (processes), their opinions in respect of the reason why a specific method is selected, and the success rate of the methods actually used. All responded, reflecting a 100 % response.

Table 3 shows the preference rating of respondents related to the tested dispute resolution methods.

Table 3: Preferred dispute resolution methods (1 = Not preferred, 5 = most preferable)

Preferred dispute resolution methods	Quantity Surveyors average	Previous study (2006)
• Resolution by Principal agent	4.39	4.3
• Conciliation	3.88	3.4
• Mediation	3.86	4.0
• Adjudication	3.82	3.2
• Arbitration	3.10	3.0
• Litigation	2.00	-

The responses indicate that the processes allowing for communication and early solutions are preferred, and that it should be done by people who are closely linked to the project. Resolution in terms of negotiations by Quantity Surveyor / Principal Agent, are seen as the most preferred method. This study is fairly consistent with previous studies, although previously mediation was ranked higher and also received a better average score.

4.4 Study 2: Reasons for selecting an ADR method

This study was also conducted to show the reasons why an ADR method was selected. The reason for this study was also to identify the reasons linked to the preference of methods and by doing so, to enable parties to select the best claims and dispute resolution communication procedure.

Table 4 shows the reasons for choosing a specific ADR method

Table 4: Reasons why method is preferred: percentage average allocated

Reasons	Percentage
Control	86%
Cost effective	85%
Time effective	84%
Consensus	84%
Satisfactory end result	82%
Continuity	82%
Confidentiality	82%

Percentages allocated to the reasons why a specific ADR method should be used are important and consistent. Although control and cost effectiveness received the highest percentage from the respondents, the results show little difference of opinion, and all the reasons should be seen as important, leaving the selection of a most beneficial method to the parties concerned.

5 Findings

The results show the importance of quality and effective communication instruments, and also that these should be well managed to avoid communication problems, disruptions and disputes.

The most serious problems in respect of communication that may lead to disputes were identified. In the study the following problems were seen as very serious.

- Employer's instructions;
- Employer/professional team relationship;
- Mismanagement of time and programming; and
- Professional agents' non-empowerment.

It is evident from this study that the professionals prefer the following ADR methods. The studies related to preferences show a consistent result.

- Agent resolution/negotiation type of process;
- Conciliation;
- Mediation; and
- Adjudication.

The above- mentioned results suggest that dispute resolution methods and the selection of a method could possibly address the above-mentioned reasons why some ADR methods are preferred, or at least a combination of the reasons, to ensure the risk and interests of the parties are taken into account and managed well. Resolution methods that do not address these issues should be avoided.

6 Conclusion

The studies had the objective of identifying the most serious problems related to communications that may lead to disputes and the extent to which various dispute resolution methods are preferred and used in South Africa, as well as the perceived reasons why an ADR method should be used. According to the majority of the respondents, ADR methods are preferred to traditional litigation to solve disputes. Of the various types of ADR methods that respondents were involved in, mediation, negotiation type methods and expert (agent) resolution stood out as the most preferred methods. Respondents also identified the reasons why they prefer ADR to traditional methods. The selection was mostly motivated by control by the parties and time- and cost-effectiveness. The case study showed that bad communication and communication instruments may lead to disruption and disputes. The management of communication is therefore fundamental to avoiding disputes.

7. Recommendations

7.1 Problems

Important problems were identified and it is thus recommended that the possible occurrence of these problem areas should receive special attention. This will lead to better communication, less disruption, lower costs and a decline in the number of disputes. A dispute resolution with strong supporting communication enhancing procedure is recommended.

7.2 Preferences

If the strong preferences for specific ADR methods are considered, it is clear that these methods should form part of the contractual and communication arrangements. It is thus specifically recommended that contracts should include clauses that allow continuous communication to avoid disputes and include methods that will enhance the settlement of differences rather than an early formalised dispute resolution process in a court-type situation.

7.3 Processes

The reasons why a specific ADR method is selected should be considered to lead parties to a contract to design a communication, claims and dispute resolution procedure that best suit their needs and that will be effectively communicated and managed by their professional teams.

The following organogram (Figure 1) reflects a provisional proposed model that shows the relationship and link between the three P's; namely problems, preferences and processes.

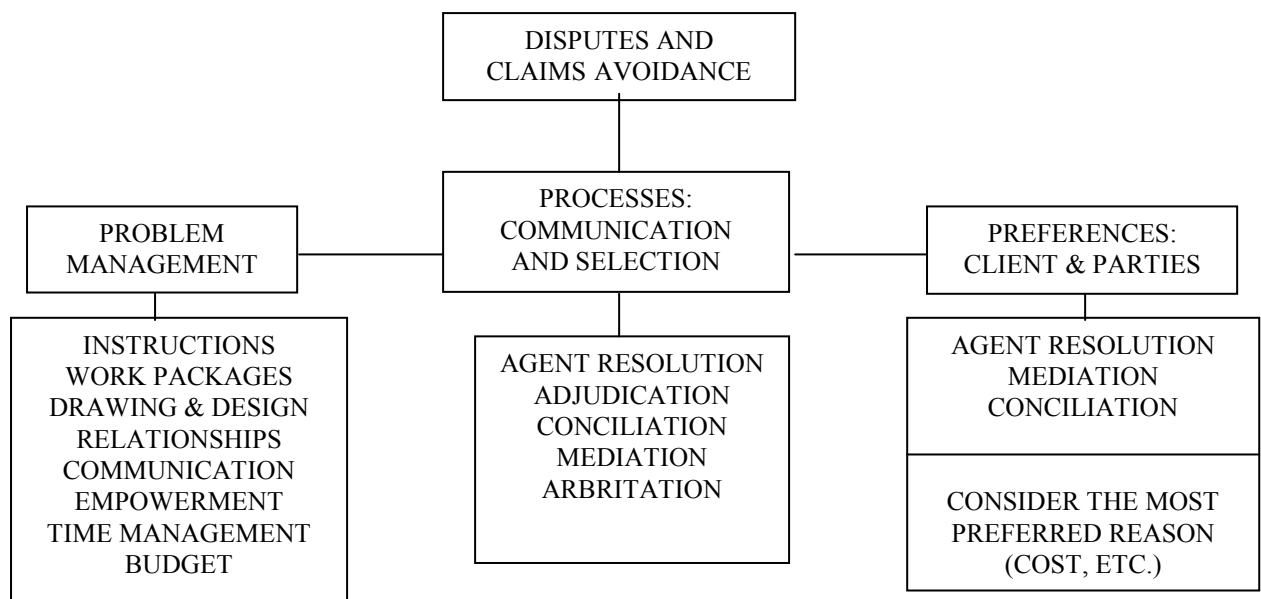


Figure 1: Proposed Model

The aim of the proposed model is to assist parties to a contract to seek a procedure that best suit their needs, but that will also take the problems that may become evident during the construction process, based on past experience, into account and to select the best ADR method applicable to them.

An approach to enhance communication and a management and contractual system that identifies problems early and that helps with the selection of an applicable process based on business principles preferred by the parties is thus proposed. The above-mentioned model therefore shows a recommendation in respect of these considerations.

8 References

- Binnington, C. 2005. Arbitration – a cost effective mechanism? Legal matters. *Civil Engineering Contractor*, 39(11), p. 49-50, September.
- Butler, D.B. and Finsen, E. 1993. *Arbitration in South Africa: law and practice*. Cape Town: Juta.
- Chappell, D., Powell-Smith, V. and Sims, J. 2005. *Building contract claims*. 4th ed. Oxford: Blackwell Publishing.
- Eilenberg, I. 2002. *Dispute Resolution in Construction Management*. Sydney: UNSW.
- Finsen, E. 1999. *The building contract: A commentary on the JBCC Agreements*. Cape Town: Juta.
- Finsen, E. 2005. *The building contract: A commentary on the JBCC Agreements*. 2nd ed. Cape Town: Juta.
- Hibberd, P. and Newman, P. 2001. *ADR and Adjudication in Construction Disputes*. Oxford: Blackwell Science.
- Loots, P. 1991. Alternative methods of dispute resolution in the construction industry. *SA Builder*, p. 8-13, May.
- Mackie, K.J. 1991. *A handbook of dispute resolution: ADR in action*. London & New York: Routledge and Sweet & Maxwell.
- McKenzie, H.S. and McKenzie, S.D. 2009. *The law of building and engineering contracts and arbitration*. 5th Ed. South Africa: Juta.
- Quail, P. 1978. *Die boukontrak*. Pretoria: Boupublikasies.
- Scanell, J. 1993. *Spreuke van Langenhoven*. 6th Ed. Cape Town: Tafelberg.
- Tiruneh, G.G., Verster, J.J.P. and Kotzé, B.G. 2007. A study on the application of alternative dispute resolution (ADR) methods in South Africa. *Proceedings of Pacific Association of Quantity Surveyors (PAQS) 2007 Congress: Construction from a different perspective*. Auckland, New Zealand, June.
- University of the Free State (UFS). Department of Quantity Surveying and Construction Management 2006. *Survey on Procurement, communication and dispute resolution*. Bloemfontein: University of the Free State.
- University of the Free State (UFS). Department of Quantity Surveying and Construction Management. 2007. *Survey on communication and communication instruments used in the construction industry*. Bloemfontein: University of the Free State.
- University of the Free State (UFS). Department of Quantity Surveying and Construction Management. 2010. *Survey on dispute resolution and preferences*. Bloemfontein: University of the Free State.
- Verster, J.J.P. 2006. Managing cost, contracts, communication and claims: A Quantity Surveying perspective on future opportunities. In: *Proceedings of 1st ICEC & IPMA Global Congress on Project Management, 5th World congress on Cost Engineering, Project Management and Quantity Surveying*. Ljubljana, Slovenia, 23-26 April.
- Verster, J.J.P., Ramabodu, M.S. and Van Zyl, C.H. 2010. Construction dispute resolution as seen by Quantity Surveyors and Construction Managers. *Proceedings of the 7th International Cost Engineering Council World and 14th Pacific Association of Quantity Surveyors Congress*. Singapore. 23-27 July.

South African Principal Building Agreement Documents and South African Acts

Joint Building Contracts Committee (JBCC). 2004. *Principal building agreement*. 4th ed. Houghton: Joint Building Contracts Committee Inc. (Series 2000 Code 2101), March, p. 30.

Joint Building Contracts Committee (JBCC). 2005. *Adjudication Rules: for use with the JBCC Principal Building Agreement and Nominated / Selected Subcontract Agreement*. Ed. 4.1, cl. 1.1 3.2, 6.0-7.0. Houghton: Joint Building Contracts Committee Inc. (Series 2000 Code 2019), March.

Joint Building Contracts Committee (JBCC). 2007. *Adjudication Rules: for use with the JBCC Principal Building Agreement & Nominated / Selected Subcontract Agreement*. 5th ed. cl 1 – 7. Houghton: Joint Building Contracts Committee Inc. (Series 2000 Code 2101). July. Cl. 1 – 7. p. 30 - 31.

South Africa. 1969. *The Arbitration Act, Act 42 of 1969*. Pretoria: Government printer.

A framework for the application of Dispute Resolution Boards in the Chilean construction industry

Maria Alejandra Medina Oyuela

Department of Construction Engineering and Management, Pontificia Universidad Católica de Chile,
moyuela@uc.cl

Alfredo Serpell Bley

Department of Construction Engineering and Management, Pontificia Universidad Católica de Chile,
aserpell@ing.puc.cl

Abstract

In Chile, the main mechanisms used to resolve contractual conflicts in construction projects are arbitration and litigation. Although the Dispute Review Board (DRB) mechanism has been used internationally with success for more than 30 years, it has been rarely applied in domestic contractual conflicts. A study was carried out to find out the reasons restricting the application of this mechanism in Chile and to propose an implementation framework for it, adapted to the reality of the domestic construction industry. This is deemed necessary due to the fact that both arbitration and litigation are not producing good results for involved parties in contractual conflicts. The study was based on literature research and interviews to professionals with recognized expertise in this topic who provided their opinions about the application of the DRB mechanism in Chile and validated the proposed methodology. The paper concludes that the main barriers for the implementation of the DRB mechanism are the lack of knowledge of it, lack of experience in using it and change resistance. In addition, it was found that there is a high potential for the application of DRB but this is subject to the perceived benefits of the mechanism.

Keywords: Dispute Review Boards- Conflict - Mechanism – Resolution- Framework

1. Introduction

The construction industry is characterised by a highly confrontational nature (Campero and Alarcón, 2003). Therefore it has been one of the first industries that has sought to resolve its contract disputes through means that are alternative to arbitration or court trials (Paredes and Grey, 2008), as the latter two turn out to be lengthy and usually expensive processes.

The relationship between owners and contractors in construction projects has always been a source of concern among practitioners and researchers in the field of construction management (Serpell, 1999). Interest has been shown at the international level for incorporating non-controversial alternative mechanisms of dispute resolution that are proactive rather than reactive to problems that arise during the development of a project (Harmon, 2003a; Jones 2006). At the national level, the need to create awareness and incorporate means of dispute resolution other than arbitration and the ordinary courts has been suggested (Helmlinger and Cruz, 2006). This could reduce the negative impact of contractual disputes and adversarial relationships between clients and contractors (CDT, 2006).

One such mechanism is the Dispute Resolution Board (DRB), the subject of this research study. In Chile there are just a few companies that know the benefits and the operation of the DRB. For this reason, it would be highly convenient to diffuse its characteristics and the benefits of incorporating it in the country, and to promote its application in the different construction sectors. Through this research and taking into account the international experience, a framework for adapting and implementing this mechanism for resolving disputes in the domestic environment is proposed with the hope that Chilean clients and contractors will consider making use of it in the near future. It is hoped that the application of DRB may bring the same benefits that have been reported in the international literature and will help resolve disputes in a non-controversial, practical and objective way within a climate of cooperation between parties.

2. The Dispute Resolution Board

The Dispute Resolution Board is a pre-arbitration stage (Arkin, 1993) as it prevents and resolves conflicts that may occur between the parties in a timely manner, from the beginning and during the development of works (Figueroa, 2009; Harmon, 2003b). It is implemented under a formal written agreement signed by the parties to a contract (Figueroa, 2009). The board's working procedure is oriented to safeguard the progress of the project work and to recommend solutions to disputes that arise. If the DRB fails in solving the disputes, then parties can always resort to an arbitral or judicial body.

One of the strengths of the DRB and what it might be its main difference from other forms of alternative dispute resolution is its nearness to the development of the contract and the project (Harmon 2003b; Gebken, 2006). Since the project is visited regularly by the DRB, its members acquire the knowledge about difficulties that arise between the parties and can provide recommendations to resolve complaints before they become serious disputes.

The success of the DRB is reflected by its extensive use globally since its incorporation in the El Cajon hydroelectric project in Honduras in 1980 (Spurin, 2003; DRBF, 2007). It has been applied in several major construction projects in different countries being the Channel Tunnel France-England one of the cases (DRBF, 2007; Figueroa, 2009). These performances have led large global institutions (e.g. World Bank, FIDIC, American Arbitration Association and the Chamber of International Trade) to incorporate the DRB into their projects (Buchanan and Hassall, 2009; FBD, 2006; Bunni, 2005).

The construction industry has been broadening the range of application of the DRB, from complex projects such as hydroelectric and mining projects, to small and relatively simple projects (Spurin 2003) including highways, process plants and building projects, among others (Buchanan and Hassall, 2009; Harmon, 2009).

2.1 *The DRB in Chile*

In Chile, DRBs have been established by contract and exceptionally in some complex projects in the mining sector. The mechanism has also begun to be incorporated in the Ministry of Public Works via the amendment of the Concessions Act, by replacing a conciliatory commission by a DRB (Figueroa, 2009).

A recent study carried out by the Chamber of Arbitration and Mediation of the Chilean Chamber of Commerce about the construction sectors where the majority of disputes occur, shows that the concessions (16%) and mining (12%) sectors only account for 28% of all the disputes in the industry (CAM, 2009), which means that 72% of the disputes occur in other construction sectors which could make use of this alternative mechanism for disputes resolution.

At the national level, there is not one institution that promotes the use of this device and its main characteristics yet. It should be noted that for the incorporation of DRB in a construction project it is only needed an agreement between the parties and to leave it stated in the contract.

2.2 *Comparative Analysis between Arbitration and the DRB in Chile*

Arbitration is the traditional and most common form of conflict resolution used in Chile, in addition to regular courts (Helmlinger and Cruz, 2006; Arrau, 2007). These methods are reactive and adverse (Harmon, 2003). A culture of negotiation does not exist in Chile as it can be found in the United States

and European countries, where a solution to the dispute is sought and a change from antagonistic attitudes to a cooperation climate between the parties is promoted.

In Chile, in 56% of the cases, dispute resolution through arbitration is initiated when works have already been developed, the information is not available, the parties already have adversarial positions and the problems have been accumulated (CAM, 2009). In contrast, the use of the DRB allows an updated knowledge of the project through its permanent monitoring.

According to Arrau (2007) the average duration of arbitration processes in Chile is 8.3 months and sometimes they have lasted more than a year. The final result of arbitration is a mandatory and ending sentence without appeal. The DRB has the advantage that it is able to issue a decision within a relatively short period of time. This resolution may be a recommendation or a decision. As indicated above, if parties disagree, they can always resort to arbitration or to ordinary courts.

Arbitration has become an appropriate forum for achieving agreements between parties, as reflected by the 60% of conciliations that have taken place during arbitration processes along the period 1992-2009 (Arrau, 2007, CAM, 2009). This number reflects the high interest of the parties to resolve disputes by mutual agreement and to incorporate other mechanisms where attitudes win-win subsist (CAM, 2009). A mechanism that achieves these attitudes is the DRB.

Cost is a very important factor to consider when choosing the mechanism to resolve disputes in construction projects. The cost of holding special meetings to hear disputes and to prepare the recommendations of the DRB itself is small compared with the costs of arbitration or trial (Thompson, 1998; DRBF, 2007). A DRB maintains the dispute-related costs to a minimum (Harmon 2003a). To pay arbitrators, lawyers, consultants and experts who are necessary for arbitration or trial, does not add value to the project but it represents the sunk costs of disputes.

Given the fact that with the arbitration it is often that parties are not satisfied with the results, the DRB can therefore become a convenient alternative for contract disputes resolution in the Chilean construction industry.

3. Research Methodology

The objectives of this research were to: a) identify the main features of the Dispute Resolution Board mechanisms and the way in which it can help reducing the impact of contractual disputes, b) identify critical issues that would constrain the implementation of the mechanism in Chile, and c) diffuse the DRB mechanism in the domestic construction sector and propose a framework for its implementation in Chile as recommended by the experts, helping owners and contractors to incorporate it as a mechanism for dispute resolution in their projects.

The study encompassed a literature review and specialists interviews. The literature review covered the topics of construction disputes and disputes resolution mechanisms, including the DRB. Its results provided the researchers with the fundamentals of the DRB and the identification of the main issues to be included in specialists' interviews.

The interviews were semi-structured and applied to 15 specialists with proven expertise and experience in the area. The interviewed professionals make up almost a census of specialists on this topic in the country. To assure this condition, a thorough search was carried out and also, the same interviewed experts were asked about additional specialists to interview. The interviewed specialists included: a) five professionals who have been members of Dispute Resolution Boards in Chile, b) three experts on construction claims, c) two professionals who participated in the application of the DRB by the only client that has incorporated it in construction projects, and d) five contractors who have participated in domestic projects that included the DRB mechanism.

The information obtained both from literature and interviews was reviewed and analysed. With this input, an implementation approach of the DRB in Chile was developed, giving shape to the implementation framework that is presented later in this work. To establish reliability and credibility on the research, multiple data sources were used as well as data triangulation. Also, the chain of evidence was established and the results were verified with appropriate persons (Hernandez et al, 2006).

The validation of the study was conducted by the review of the DRB and the recommendations included in the proposed framework, by a panel of experts formed by the same interviewees. On the other hand, to validate the instrument for information collection, a pilot questionnaire was used in order to minimize incorrect interpretations by interviewees (Yin, 2003).

4. Results of Interviews

Interviews were designed to answer two basic questions as follows:

1. How and why the Dispute Resolution Board can help reducing the impact of contractual disputes? The variables used to answer this question were:
 - a. Current practices of the application of the DRB mechanism in Chile, and
 - b. Advantages and disadvantages of the system with regards to the duration of the project, contractual relationships and cost.
2. How should the DRB mechanism be adapted to the Chilean/Latin America conditions so that it will be accepted and used? The variables used to answer this question were:
 - a. Barriers to the implementation of the DRB, and
 - b. Required conditions for its implementation.

4.1. *Current Practices of the application of the DRB mechanism in Chile*

Interviewees indicated that the Dispute Resolution Board is a new figure in Chile. An imitation of what has been done internationally for the last 30 years has been attempted in the domestic sector since 2007. The first incorporation of the mechanism to a construction project was done by a mining company in that year. The scheme being applied has consisted of a Dispute Resolution Board which makes recommendations and, if a party requests so and the other party is not opposed to, issues resolutions. This mechanism of dispute resolution has been implemented in high risk contracts between a state mining company and a private construction company.

According to interviews, this pre-arbitration dispute resolution mechanism has demonstrated several advantages, including the timely identification of problems that arise in the project, providing solutions to them in a quick, objective and practical manner. Since the committee is a neutral, independent and experienced third party and it has a comprehensive understanding of the project obtained through continuous monitoring from its start, the prevention of disputes is facilitated and, if they occur, their solution is then carried out in a timely way and avoiding the high costs of arbitration and trials.

However, some disadvantages have also been perceived. Interviewees believe that this mechanism would not be suitable for short-term projects lasting less than one year. On the other hand, some users have seen the DRB mechanism as a nuisance, since it is a third party to which information should be

given continuously.

4.2. *Relevant issues of the DRB in projects*

The views expressed by the interviewees on relevant issues of the implementation of DRB in construction projects, are shown in Table 1.

Table 1: Relevant issues of the DRB on Construction Projects

Relevant Issues		<i>f</i>
Regarding contract relationships	• It helps maintain a balance between the parties	10
	• It helps to have honest and transparent relationships, conversations and discussions.	9
	• If parties disagree with the recommendation, parties may resort to arbitration.	9
	• It limits the potential abuse and arbitrariness from the owner. Today owners can eventually change their actions and decide on their own rather than waiting to go to arbitration or trial.	8
	• Parties have the same responsibilities and obligations, but must explain the decisions made.	5
	• Parties assume the duty of providing all the required information.	5
	• The DRB does not establish any limitation, but now parties are more careful to show data that are not factual, as this would lead the DRB members to have the impression that they are acting in bad faith.	4
Regarding the duration of the project	• Members of the DRB are familiar with the project since the mechanism is incorporated at the beginning of the contract. This helps reducing the time required for solving problems.	15
	• It allows DRB members to be permanently aware of the project through meetings and site visits every two months and monthly internal meetings.	13
	• It helps to resolve conflicts in a timely manner.	9
	• There is a deadline for making recommendations or decisions ranging from 30 to 60 days from the moment all antecedents are available.	8
	• In projects of duration of one-year or less, the DRB meetings and visits are held once a month.	7
Regarding the project's cost	• Benefits of the mechanism outweigh costs.	11
	• The cost of the DRB is irrelevant and negligible compared to the cost of the project.	9
	• DRB fees are paid between the two sides, equally.	8
	• It avoids costly and lengthy arbitration processes that need long and complex administrative processes.	5
	• It has a fixed monthly fee for accompanying parties and a system of hourly rate depending on the hours spent on conflict resolution.	5
	• The cost is around a 3% of the contract amount.	1
Interviews: 15 experts		

Due to its features of the early incorporation in contracts, the continuing familiarity with the project and time constraints for issuing decisions, the DRB generates an impact that is considered positive by interviewees. This mechanism seeks to improve the contractual relationship and create an atmosphere

of cooperation and trust. It also helps maintaining fairness between parties and reducing problems potential.

Interviewees agree with the literature in the sense that benefits of the DRB are higher than its costs and that the mechanism is very profitable. The cost of maintaining a Dispute Resolution Board is very reasonable because it usually avoids falling into arbitration or a trial, averting the damage to contractual relationships and the delay to resolve a dispute through arbitration or trial.

4.3. Barriers to the implementation of Dispute Resolution Boards

The barriers to successful implementation of the DRB in the country as were identified by interviewees are presented in table 2 together with the frequency with which they were mentioned.

Table 2: Barriers to the implementation of the DRB in construction projects in Chile

Barriers	Frequency
Resistance to change.	15
Lack of knowledge about the DRB.	9
Lack of experience in its use.	9
The arbitration or trial are entrenched in the environment	2
Clients prefer to postpone the possible dispute resolution.	2
They believe that is going to reduce their margin of movements and that it will become a burden and an hassle for them.	1
Interviews: 15 experts	

According to interviewees, owners and contractors do not yet know the DRB, its benefits or the way it operates and resolves. Lawyers also don't know how this mechanism works and that is the main reason why they do not incorporate it into the contracts.

Human nature often resists change because people do not know other alternatives that are more efficient and have multiple benefits. For this reason, most industry players have no confidence in the use of the DRB. Moreover, the interviewees mentioned that the figure of arbitration is well entrenched in the country and that everybody seeks to resolve contractual issues when the project is well advanced and parties already have a defensive posture. According to interviewees, both the client and the contractor most often start the contract in a spirit of bad faith and from the beginning they do not trust each other because a particular contract strategy is already in place in each party. This leads the parties to believe that having a third party intervening, i.e. the DRB, to which each party has to

provide information from the beginning of the project, is going to take out some of their action margin in the contractual relationship and thus becoming a drawback for them.

Interviewees also indicated that a lack of experience in the use of the DRB has been observed when minimum conditions of parties' behaviour are not met, which has led to a party slowing down the incorporation of the mechanism, not allowing it to act immediately. Interviewees consider that the DRB has been widely accepted by contractors; however, there has been reluctance in its use by some of the site managers.

In one project where a DRB formed by just one member has been used, a lack of resources for its proper functioning was observed. This was due to the fact that one person was not able to address all the problems that have arisen in the project. It has been seen that when problems grow into major conflicts, the hours devoted to the project and to visits were not sufficient.

4.4. Required conditions for adapting the Dispute Resolution Boards in Chile

Interviewees stated their observations regarding to how to modify the mechanism and the conditions required to implement the Dispute Resolution Board in Chile. Table 3 summarizes the conditions set by the interviewees in this sense.

Table 3: Required conditions for implementing the Dispute Resolution Board in Chile

Conditions of use	Frequency
DRB members must prevent and avoid problems as they occur in the project.	15
Knowledge of the mechanism.	13
Members must have experience in construction and in dispute management.	13
Commitment of the parties to the DRB.	12
Members should be aware of the project and of all its documents, including the work log, among others.	11
The Chilean Construction Chamber should promote the mechanism.	11
Incorporation of the mechanism at the national level: it is sufficient that the parties agree to enter the facility and incorporate it into the contract.	9
Members of the DRB must be of recognized professional prestige.	9
In Chile, during the early years, the DRB should function as an organ that only issues recommendations.	9
DRB members must be independent of the parties so that no conflicts of interest exist.	8
DRB members must keep track of the project, with meetings and visits to the project.	8
The DRB must be implemented in high cost projects.	7
The DRB must be implemented in projects with complex definition, execution of the work and operation.	6
The DRB must be implemented in large construction projects.	6

There must be cooperation of the parties for the DRB to work properly.	6
The Santiago Chamber of Commerce should promote the use of the DRB.	5
The Ministry of Public Works should promote the use of the DRB.	4
The DRB should be used in every project and in every construction sector, regardless of size.	4
For contracts of high budgets and complexity, a 3-member DRB should be used.	2
In small contracts, a one-member DRB should suffice.	2
Interviews: 15 experts.	

5. Framework for adapting the DRB in Chile

By taking into account the information obtained from the interviews and the literature review, an implementation framework is proposed for the DRB. First, actions to overcome the implementation barriers are proposed. Second, the conditions required for the DRB implementation are identified. Third, the factors deemed necessary to create the conditions to apply the DRB are discussed and fourth, an approach for the DRB application is proposed.

5.1. Actions to overcome implementation barriers

Experts mentioned barriers to a proper implementation of the DRB mechanism in Chile. Figure 1 shows actions that would help lower these barriers.

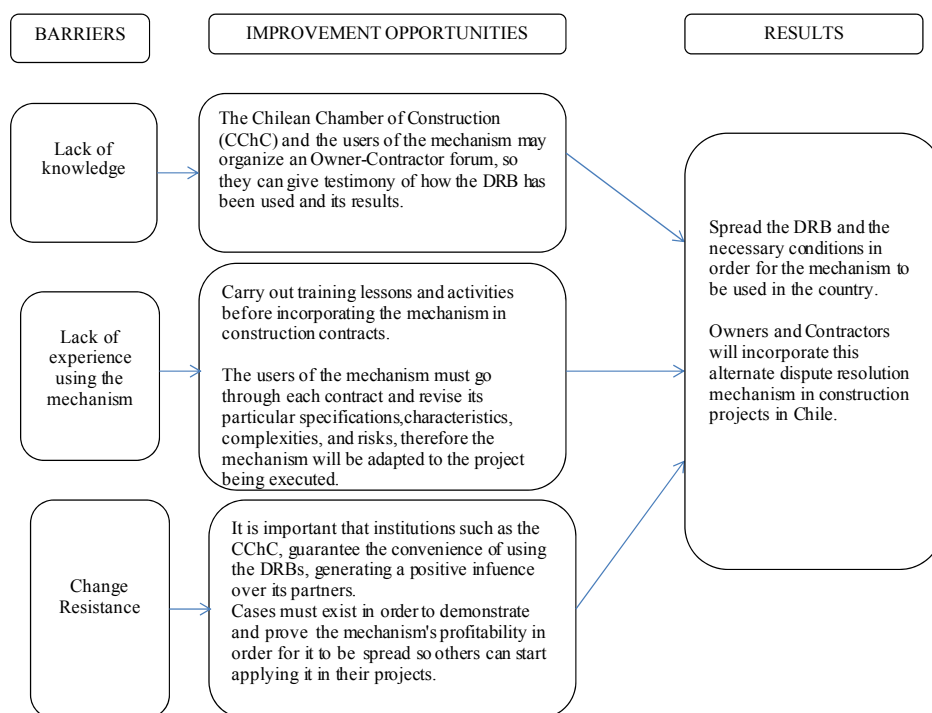


Figure 1 Actions to overcome the DRB implementation barriers

5.2 Conditions required for implementing the DRB

According to the opinions given by the interviewees, a scheme which reflects the conditions necessary for adapting the DRB mechanism in Chile is presented in figure 2. This scheme should help its acceptance and promote its usage.

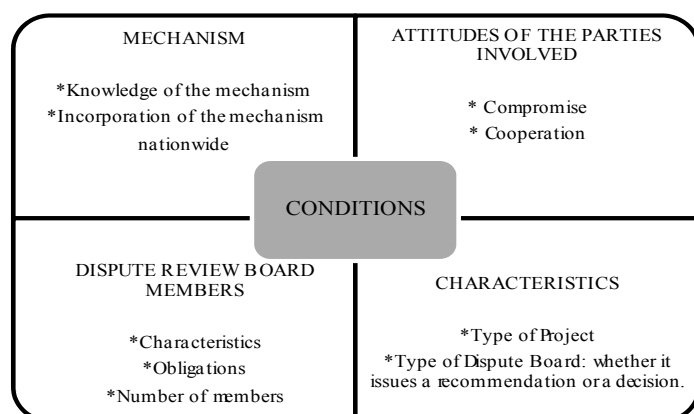


Figure 2 Conditions required for adapting the Dispute Resolution Board in Chile

To meet the conditions displayed above, a framework that will help owners and contractors to incorporate the DRB as a mechanism for dispute resolution is proposed as follows:

- a. Knowledge of the DRB and its incorporation in Chile: for the system to be successful it is essential that the parties know and accept the way in which DRB operates and the critical aspects of its implementation as well as the benefits and disadvantages of its application in contracts. To achieve this it is necessary to educate the parties and users about the DRB and it should be used only if both parties are convinced that they will benefit from it. Within the principle of contractual freedom and autonomy, the parties are free in a construction contract or any other, to agree to the mechanisms of dispute resolution, and within them, the DRB as a pre-arbitration instance. This mechanism is not enshrined as a legal obligation on public works contracts, except in the draft amendment of the Law of Concessions of Public Works. The client may decide to incorporate it into the bidding stage and, once accepted by the contractor, there will be a clause to govern its application for both parties. On the other hand, institutions like the Chilean Construction Chamber, the Santiago Chamber of Commerce, the Mining Council, the Institute of Engineers, the Society of Civil Engineers and the universities should diffuse the DRB and promote its application in the country.
- b. Attitudes of the parties: to generate a commitment from the parties, it is imperative to have contractual documents that collect the obligations of the parties in an equitable manner, so as to avoid an imbalance between the risks that threaten the productivity and the efficiency of projects implementation and the final price. Among the documents that must exist for the incorporation of

the mechanism into the contract, there must be: a clause related to Dispute Resolution Boards; an operation regulation from any arbitral institution, such as the International Chamber of Commerce or otherwise, and agree on the same contract an ad-hoc regulation to which the DRB must abide to; and a contract for the DRB members and the contractual parties. Parties should cooperate fully with the DRB and provide the information in a timely manner. In particular, the parties and the board must, as soon as the board is constituted, work together to ensure that the DRB is fully informed about the contract and its implementation by the parties.

- c. Members of the DRB: the regulation proposed by the International Chamber of Commerce which has been adapted in projects where the DRB has been used in Chile is used as a reference. This ruling is general and gives no indication on the number of members or the type of resolution that the DRB must issue. The parties are the ones who must agree about each issue freely. The number of members will typically depend on the amount of the contract and the complexity of the project. For high cost or high complexity contracts it is advisable a three-member committee, where a multidisciplinary work is carried out between lawyers and engineers. The parties shall agree the first two members. They may be nominated by major industry institutions. The third member, who acts as the board chairman is chosen by the first two members. When the parties have agreed in a one-member board, they will name this person by a common agreement. The DRB requires hiring top-level experienced professionals and who are not involved with any of the parties. They are persons that do not lead to any doubt as to potential disqualifications. It may be that parties agree to members despite having some link to any of the parties. This could result in the case of people with high moral character, impartiality and discretion.
- d. Characteristics of projects and types of determinations to be issued by members: regarding the use of the DRB, this is fully applicable to any contract work with the following characteristics: higher level projects with multiple disciplines and specialties that present complexities. For smaller projects, it should be projects that despite their size, they have many complexities and it is justified to have a DRB member studying and following up the project. In the contract, the parties will also define what type of DRB will be used. The definition will depend in each case of the characteristics of the contract. It is difficult to suggest what kind of determination is desirable, without knowing the details of each contract. It is the will of both parties if they want that the decisions issued by the DRB are to be recommendations or abiding decisions. Experts recommend that at the initial implementation processes the DRB issue recommendations and, as they gain experience, the DRB start issuing decisions.

5.3 Factors deemed necessary to create the conditions to apply the DRB

The factors that are considered necessary to apply the DRB are shown in figure 3.

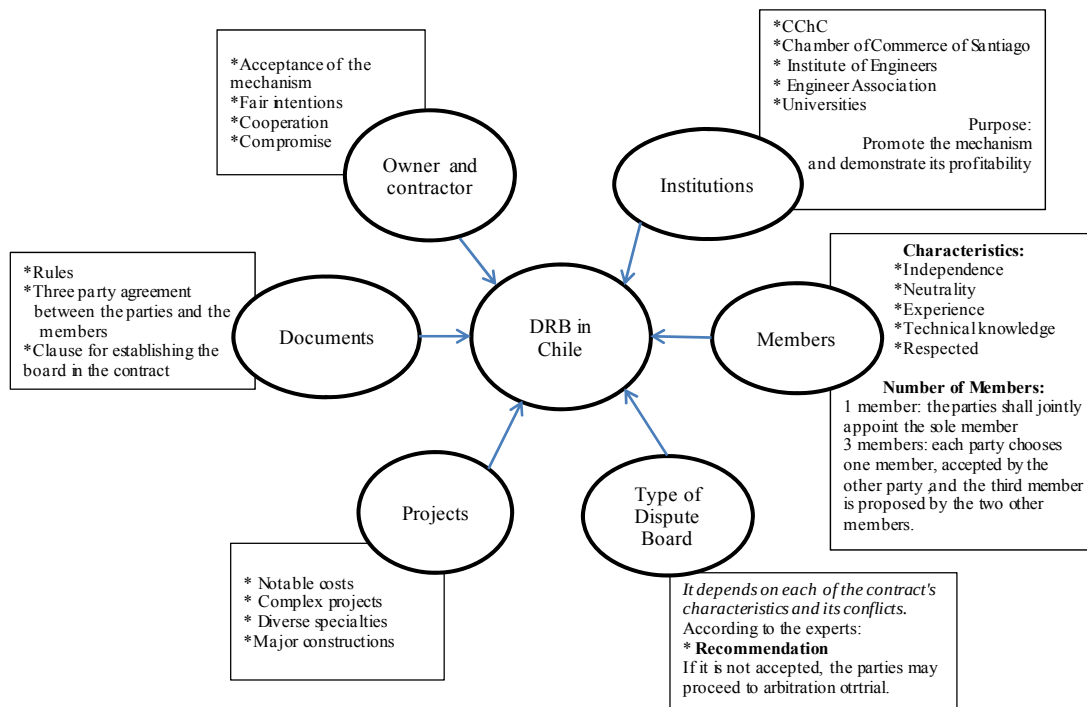


Figure 3: Factors to consider for implementing the DRB

5.4 An approach for DRB application

Based on the information from interviews and considering that all the above factors are met, an approach for the implementation process of the DRB that will achieve this mechanism to be accepted and used in Chile is proposed. The application of the DRB could reduce the impact of contractual disputes.

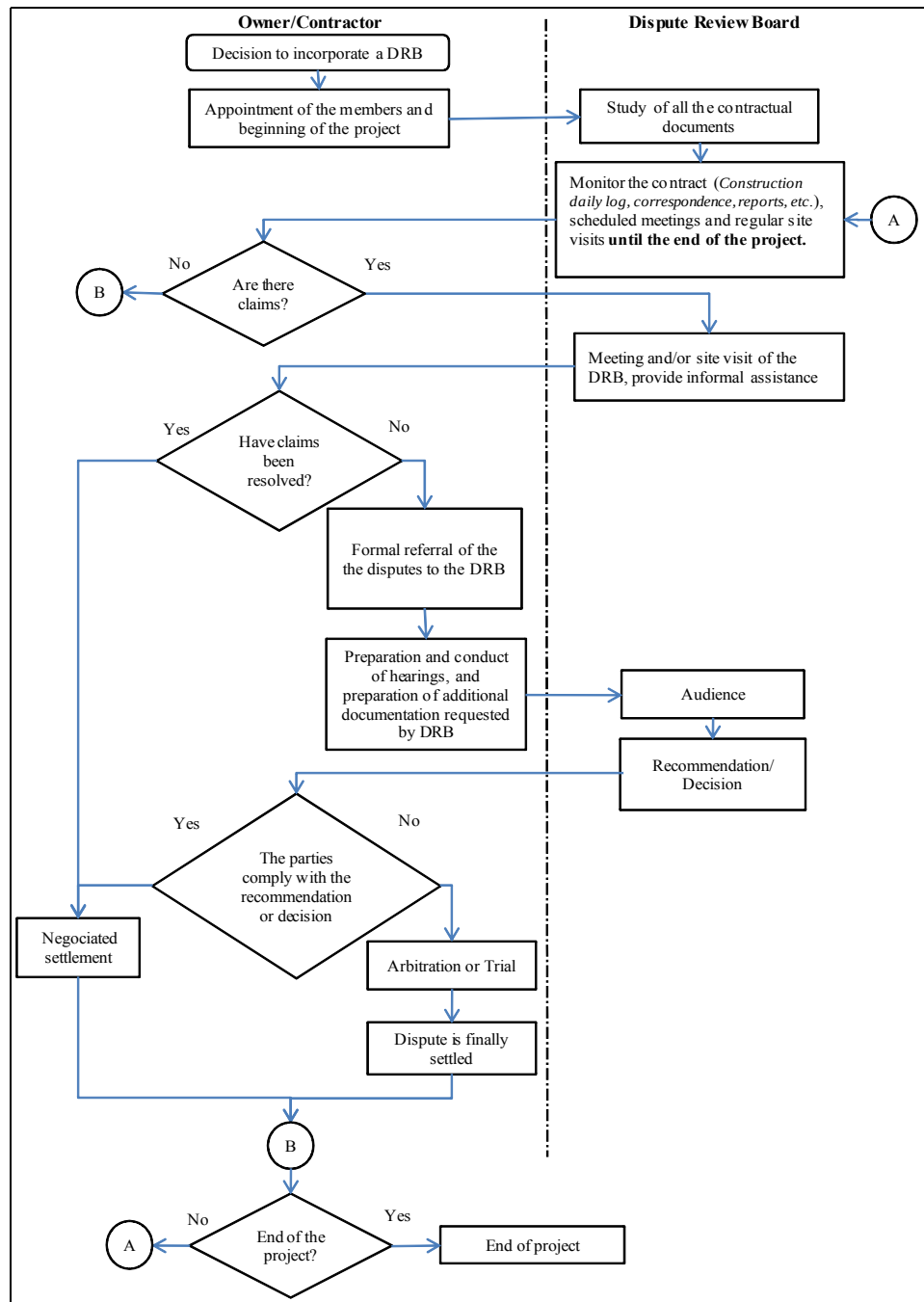


Figure 4 A proposal for implementing the DRB

6. Conclusions

The Dispute Resolution Board mechanism seeks early resolution of differences that occur during the execution of a contract, preventing that by their accumulation, they become complex and costly conflicts that could jeopardize the completion of the work. Currently, traditional mechanisms used in Chile fail to meet with the conditions mentioned above. This research found that there is a very high

potential for this mechanism to reduce disputes and that this pre-arbitration mechanism should be applied in construction contracts.

It has been seen that the DRB can reduce costs, improve compliance of the contract time and improve the contractual relationship, helping the parties to address their differences through good communication and neutralizing the conflicts and solving them by agreement between the parties.

The DRB brings lower costs for the projects due to the potential conflicts that can be avoided during the course of the contract. When compared to the potential costs of arbitration, the system would offer a good alternative, cheaper and with many benefits. The larger the project, it is easier to justify the expenditure on the DRB. For smaller projects, a one-member board can be used at a lower cost.

The main barriers for not implementing the DRB mechanism are the lack of understanding of the mechanism, the lack of experience in their use and the resistance to change. Despite these barriers, there is international evidence that the mechanism works and so the Chilean construction industry need to create awareness of how what is the DRB mechanism and how underutilized is it in Chile and other Latin American countries. The best tool to create a cultural change is the information: to educate lawyers, owners and contractors, and to spread the operation characteristics and benefits of the DRB. The resistance will be reduced as benefits and profitability of the DRB are demonstrated.

For now, experts recommend using the mechanism in projects with high costs, complex projects with multiple disciplines and specialties and large buildings in which conflicts occur frequently to a greater or lesser extent. It is also recommended that the determinations issued by members of the board be not binding on the parties in this initial period of implementation of the mechanism.

The proposed framework will help owners and contractors to comply with the conditions set by the interviewed experts in the field. This will help spread the DRB and the incorporation by contractors

and clients as an effective mechanism for dispute resolution.

References

- Arkin, H. (1993). Pre-Arbitration Dispute Resolution: What is it, where is it, and why?. *International Business Lawyer* , 373-377.
- Arrau, C. (2007). Actualización Bases de Licitación: Diagnóstico de la situación, trabajo realizado y recomendaciones, Cámara Chilena de la Construcción, Santiago, Chile
- Buchanan, M. y Hassal, C. (2009). Dispute Review/ Adjudication Boards. Recuperado el 8 de Octubre de 2009, de www.bjac.org.cn/images/2009_0504/04.ppt
- Bunni, N. (2005). Dispute Boards & Dispute Resolution, International Construction Contracts and the Resolution of Disputes Conference Co-hosted by ICC and FIDIC, 17/18 October 2005
- CAM – Cámara de Arbitraje y Mediación (2009) . Recuperado el 20 de Agosto de 2009, de <http://www.camsantiago.cl/estadisticas.htm>
- Campero, M. & Alarcón, LF (2003). *Administración de Proyectos Civiles* (2ª.ed.). Santiago: Ed. Universidad Católica de Chile
- CDT – Corporación de Desarrollo Tecnológico (2006). *Diagnóstico de la Relación Mandante Contratista, Efectos de la Situación Actual* , Santiago, Chile
- DRBF *Practices and Procedures* (2007). Seattle, EE.UU.: Dispute Resolution Board Foundation
- Figuerola, J. (2009). El arbitraje en los contratos de concesiones de obras públicas en Chile. Incorporación de los Dispute Boards o Paneles Técnicos o de Expertos. *Revista Brasileira de Arbitragem*, No. 20, recuperado de [http://74.125.93.132/search?q=cache:DMV3BwxhSzAJ:www.camsantiago.com/articulos_online/El_Arbitraje_en_Concesiones_\(DRB\).doc+%22El+arbitraje+en+los+contratos+de+concesiones+de+obras+p%C3%ABlicas+en+Chile%22&cd=1&hl=es&ct=clnk&gl=cl](http://74.125.93.132/search?q=cache:DMV3BwxhSzAJ:www.camsantiago.com/articulos_online/El_Arbitraje_en_Concesiones_(DRB).doc+%22El+arbitraje+en+los+contratos+de+concesiones+de+obras+p%C3%ABlicas+en+Chile%22&cd=1&hl=es&ct=clnk&gl=cl)
- FBD – Freshfields Bruckhaus Deringer (2006). *Briefing Dispute Review Boards* . Recuperado el 11 de Agosto de 2009, de http://www1.fidic.org/resources/contracts/drbs_freshfields06.pdf
- Gebken, JM (2006). “Quantification of Transactional Dispute Resolution Costs For The US Construction Industry”, Tesis de Doctorado en Filosofía The University of Texas at Austin
- Harmon, K. (2003a). Resolution Of Construction Disputes: A Review of Current Methodologies. *Leadership and Management in Engineering*. 187-201
- Harmon, K. (2003b). Effectiveness of Dispute Review Boards. *Journal of Construction Engineering and Management* , 129 (6), 674–679
- Harmon, K. (2009). Case Study as to the Effectiveness of Dispute Review Boards on the Central Artery/Tunnel Project, *Journal of Legal Affairs and Dispute Resolution in Engineering and*

Construction , 1 (1), 18-31.

Helmlinger, K y Cruz, J. (2006). *Evolución de la Resolución Alternativa de Controversias Civiles y Comerciales en Chile* , Centro de Arbitraje y Mediación de la Cámara de Comercio de Santiago, recuperado el 27 de Abril de 2009, de http://www.camsantiago.com/articulos_online/61_1_115-126.pdf

Hernandez, R., Fernandez, C. & Baptista, P. (2006). *Metodología de la Investigación* (4a.ed.) Mexico: McGraw-Hill Interamericana

Jones, D. (2006). Construction Project Dispute Resolution: Options for Effective Dispute Avoidance and Management, *Journal of Professional Issues in Engineering Education and Practice* , 132, (3), 225-235.

Paredes, G. y Grey J. (sf). Mecanismos Alternativos de Resolución de Disputas en Construcción, recuperado el 7 de Septiembre de 2009, de <http://www.navarrolegal.com/publicaciones/Mecanismos%20Alternativos%20de%20Resolucion%20de%20Disputas%20en%20Construccion.pdf>

Serpell, A. (1999). A structured methodology for Improving the Owner-Contractor Relationship in Construction Project. En SO Ogunlana & FN Spon (Ed.), *Profitable Partnering in Construction Procurement* pp. 169-178.

Spurin, C. (2003). “Dispute Resolution Boards” Under the Auspices of the DRBF, recuperado el 19 de Julio de 2009, de <http://www.nadr.co.uk/articles/published/DisputeReviewBoards/DisputeReviewBoards2003.pdf>

Thompson, R. (1998). “Efforts To Manage Disputes in the Construction Industry: A Comparison of the New Engineering Contract and the Dispute Review Board”, Tesis de Magíster en Ciencias de la Ingeniería Civil, Virginia Polytechnic Institute and State University.

Yin, R. (2003). *Case Study Research: Design and Methods*. (3a.ed.) California: Sage Publications, Inc.

An investigation of construction lawyer attitudes to the use of mediation in Scotland

Andrew Agapiou¹ and Bryan Clark²

¹Department of Architecture, Strathclyde University, Glasgow, G4 0NG

²Department of Law, Strathclyde University, Glasgow, G4 0LT

Email: andrew.agapiou@strath.ac.uk; bryan.clark@strath.ac.uk

Abstract:

An examination of the literature specific to Alternative Dispute Resolution in the construction industry in Scotland offers a snapshot of a discipline whose research base is in its infancy (Agapiou, 2010). It is widely documented that legal advisers generally perform a gate-keeping role, advising clients on the most appropriate form of dispute resolution for particular cases. Is it reasonable to believe that the attitudes of construction lawyers in Scotland creates a real limit on what could be implemented by a government that seeks to promote modern methods of dispute resolution as part of its civil justice reform agenda? Drawn from a questionnaire survey, this paper seeks to add to the dispute-resolution literature by identifying the attitudes of construction lawyers on the use and effectiveness of mediation to resolve construction disputes in Scotland. Despite the small sample used in this study, there is little evidence within the study that the inherent conservatism of lawyers in Scotland in their approach to the conduct of disputes brought to them by clients has militated against the use of mediation: this is contrary to anecdotal evidence in the construction industry. Neither does it seem, that the lack of knowledge of mediation, far less experience of its operation, along with fear of the unknown as opposed to an adversarial process which for all its imperfections lawyers understand unequivocally has prevented the use of Mediation to resolve construction disputes. Interestingly, there would also appear to be some evidence of a modest bottom up growth of construction mediation according to the research findings.

Keywords:

Attitudes, Dispute Resolution, Lawyers, Mediation, Scotland

Introduction

The construction industry is highly litigious and ongoing disputes can be costly not merely in a financial sense but also in terms of the breakdown of otherwise profitable relationships often engendered by conflict (Oladapo & Onabanjo, 2009).

While arbitration and adjudication are commonly deployed in the Scottish construction sphere as dispute resolution tools, the adversarial nature of such processes may hold deleterious consequences for parties in terms of financial costs, delays, risks and ensuing loss of business. Proponents argue that mediation is a cheaper, quicker, and altogether more harmonious method of dispute resolution than traditional, adversarial methods. Despite evidence of modest growth in the use of mediation within Scottish civil and commercial disputes generally (Clark and Dawson, 2007; Clark, 2009), and evidence of the growing use of construction mediation in other jurisdictions (such as England and Wales) (Brooker, 1998, 2009; Gould et al, 2009), little evidence can be gleaned from the literature regarding construction mediation in Scotland. According to Sarat and Felstiner (1989, p 1664), "lawyers' assimilation, acceptance, rejection, integration, or other response to alternatives to established norms of litigation practice is critical to both the practical consequences and the impact of civil justice reform and innovation." Thus, given the 'gatekeeper' role that lawyers play they may be crucial in helping expedite the development of any innovations in the field. Despite the fact that much has been written about lawyers' role in, and experiences of mediation in commercial disputes in Scotland, and construction matters in other jurisdictions such as England and Wales (Gould et al, 1998; Brooker, 2009) and the USA (Stipanowich, 1994), little is known about construction lawyers' interaction with the process in Scotland. More generally, Scottish civil justice is in a state of upheaval. A fundamental review of the civil court system has recently been undertaken under the auspices of Lord Gill which may lead to radical reforms of the incumbent system¹. Concurrently, a new Arbitration Act has recently been passed by the Scottish Parliament, the intention of which is to create a modern, efficient framework for arbitration and thus position Scotland as an attractive centre for international dispute resolution. Against this backdrop, this paper focuses on the utility of mediation as a process of dispute resolution within the sphere of Scottish construction disputes, as well as the role of lawyers within the process.

The aim of the research was to explore the utility of mediation in the construction industry in Scotland. The objective was to elicit views, practices and experience of mediation techniques rather than an in-depth account of a limited number of randomly chosen case studies. A questionnaire approach was used for this initial stage of the enquiry. The structure of the questionnaire was based on that developed by Clark to assess the attitudes of Scottish Lawyers to Commercial Mediation (Clark & Dawson, 2007) and adapted to the construction context.

Primary data collection and analysis

Having defined the framework for the opinion survey, the next step was to develop the necessary data collection tools in accordance with the research objectives. The questionnaire was distributed to 165 legal professionals randomly selected from the membership lists of professional associations for solicitors, advocates and mediators based and operating in Scotland.

[1] The Report of the Scottish Civil Courts Review was launched by the Lord Justice Clerk, the Rt Hon Lord Gill, on Wednesday, 30 September 2009. The Report, which is in two volumes, is now available on-line @ <http://www.scotcourts.gov.uk/civilcourtsreview/>

Questionnaire Development

Since the questionnaire was to be self-administered, there was a need for it to be self-explanatory. In order to achieve this, a covering letter and an introductory page describing the aims and objectives of the research was attached to the questionnaire. The questionnaire was divided into three sections as described below;

Section One: the use of Mediation in Construction Disputes

A number of variables from the survey were selected from the questionnaires as the basis for assessing the use of mediation in construction disputes including the background and experience of the respondents in the legal profession, their training in mediation and organisational policies and practices towards mediation.

Section Two: Experience of the use of Mediation

In terms of experiences of mediation, respondents will asked to rate their responses to a number of questions ranging from client representation, levels of satisfaction with different elements of the process, factors leading to a failed mediation process, the decisions to recommend mediation to a client and reasons to refuse mediation proposals from the opposing party in a case.

Section Three: Attitudes to Mediation

The respondents were then asked to rate their perceptions of dispute resolution, alternative dispute resolution and mediation and indicate views on key policy issues relative to mediation's development. The purpose of these questions was to ascertain an understanding of the barriers to the utility of mediation in construction disputes.

Data collection process

The questionnaire survey employed sixty-nine (69) items to collect data on lawyers' experience and attitudes to mediation. A brief description of the items used for the purpose of data collection is presented below.

Eleven (11) items on a 5-point Likert scale with anchors ranging from always relevant to never relevant were used to collect data on factors relevant to a decision to recommend mediation. Seven (7) items on a 5-point Likert scale with anchors ranging from always relevant to never relevant were used to collect data on factors relevant to a decision to decline a proposal of mediation from an opposing party. For those with experience of mediation, four (4) items on a 5-point Likert scale with anchors ranging from always satisfied to never satisfied were used to collect data on satisfaction with different elements of mediation. Seven (7) items on a 5-point Likert scale were used to collect with anchors ranging from always to never on factors contributing to a failed mediation. Nine (9) multiple responses items were used to collect data on client representation in mediation by type of case and whether cases were settled, partially settled or not settled. In addition, Eleven (11) multiple response items were used to collect data on professional designation and training in mediatory techniques. Twenty (20) items on a 5 –point Likert scale with anchors ranging from strongly agree to strongly disagree were also used to collect data on attitudes to mediation. Data on organisational policy and practice were also collected.

Questionnaire Data Analysis

The analysis of quantitative data generated from the survey was based on the work of Coakes (2005), Piaw (2006) and Zulhabri et al (2005). This involved descriptive statistics due to the exploratory nature of the study and the non-parametric data generated from the survey. A number of statistical tests were also be applied for the purpose of analysis, including the utilization of Bi-variate analysis using Chi-Square (CS) Tests for independence, correlation analysis (measure of association) using Cramer's V (CV) and Contingency coefficient (CC) for nominal data and spearman's rank (SR) order for ordinal data. Initially, measurements of the coefficient of reliability, Cronbach's alpha, was undertaken to establish the validity and reliability of the survey instrument. This allowed items to be regrouped or excluded from the instrument where the value of alpha was less than 0.65, in line with statistical conventions. The Quantitative data were processed and analysed using SPSS 9.0 software package. The questionnaire was distributed to a total of 165 legal professionals randomly selected solicitors, advocates and mediators based and operating in Scotland. Fifty-questionnaires were completed and returned. This represented a response rate of 30.3%. According to Ellhag and Boussabaine (1999) and Idrus & Newman (2002) such a return rate is sufficient for the purpose of analysis for studies within the construction context. However, this remains a small sample and so the results should be viewed with caution. Below the numerical results for each survey question are given and the proportion of each result as a percentage of all responses to each question is provided in brackets. The survey results were presented in either diagrammatic or tabular form where appropriate.

Professional designation

Of the total, 82 % of the total 50 respondents described themselves as Solicitors, 6% as Commercial Attorneys, and the remainder as Advocates, Solicitor/Advocates and QCs. Figure 1 illustrates the breakdown of respondents by professional designation.

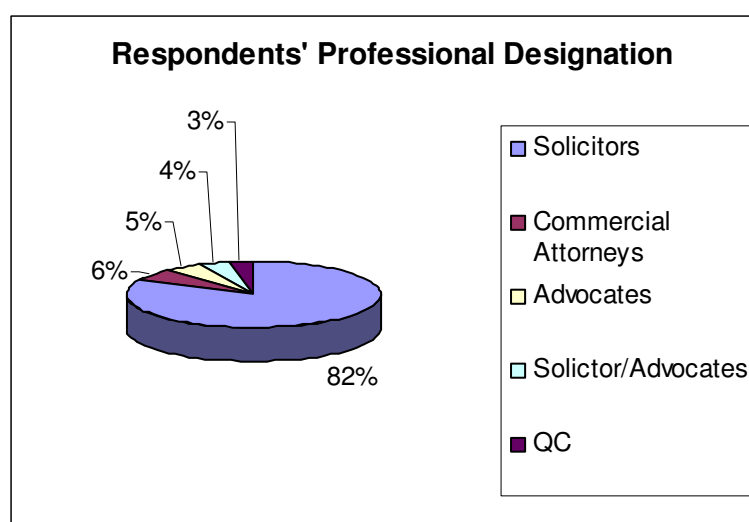


Figure 1: Breakdown of respondents' professional designation

Knowledge of Mediation

In terms of knowledge of mediation, when respondents were asked whether they would be able to explain the process of mediation to a client, an overwhelmingly majority, 90%, were able to provide an explanation of the process with confidence.

Mediation training

Of the total number of respondents' around 82% had receive some form of training in mediatory techniques with 26% of those having attended external courses on mediation and 14% attended in-house training sessions. It seems that only about one-fifth, 18%, were trained as accredited mediators and the same proportion of respondents, 20%, having some exposure to mediation within their degree or diploma courses. Perhaps this is a reflection of the relatively novelty of mediation as a dispute resolution mechanism within Scotland per se and the lack of taught provision within the respective Law Schools more specifically? Nonetheless, the latter figure represents a rise from a 2005 study of Scottish commercial lawyers, in which only 4% of respondents indicated that they had any exposure to mediation in Law School (Clark and Dawson, 2007) Figure 2 illustrates breakdown of respondents' training in mediation.

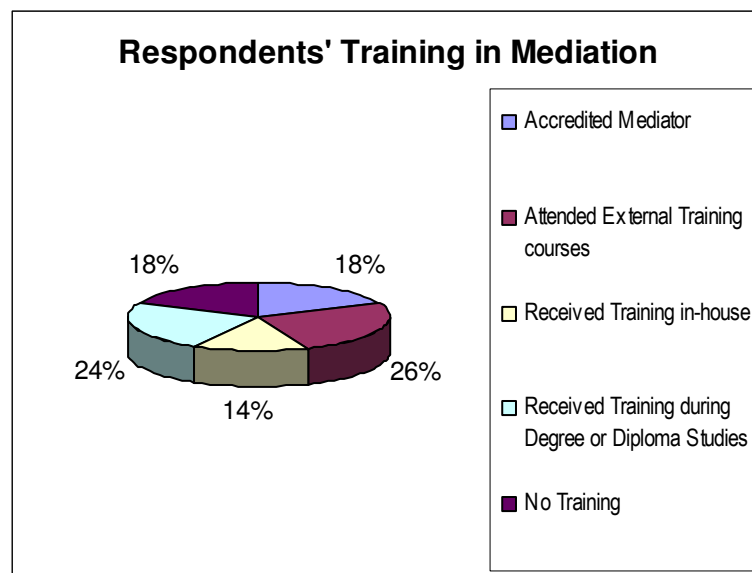


Figure 2: Breakdown of respondents' training in mediation

Years in legal practice

In terms of numbers of years of experience within the legal profession, 46% out of the total respondents were first admitted to practice law between 1991 and 2000, 32% between 2001 and 2009, and 20% between 1981 and 1990. Thus, it seems that those with some level of seniority within the legal profession, at least 9 years in the case of this study, are largely involved with decision making in terms of the settlement of disputes. This is consistent with more than 58% of the total respondents who had acted as a party representative in mediation on a least one occasion, as illustrated Table 1.

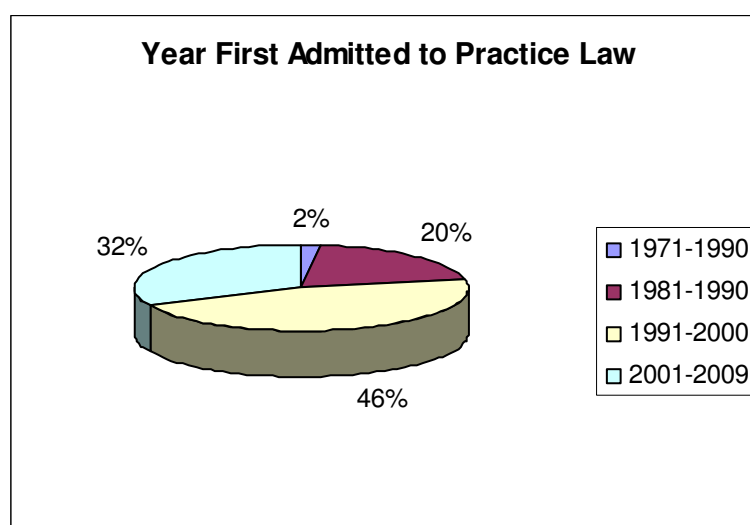


Figure 3: Respondents' years of experience in the legal profession

From a statistical perspective, there does not seem to be a close association however between the number of year in legal practice and respondents' representation in mediation as borne out by analysis in terms of the quantity of mediation cases (CS sig. values $p = 0.00 > \alpha = 0.05$; SR $r < 0.25$).

Perhaps, this is a reflection of the inherent conservatism or the entrenched position of the legal establishment towards the utility of mediation? Equally it may be redolent of the fact that mediation is still a relatively new form of dispute resolution in Scotland.

Organisational policy and practice

In terms of organisational policy and practice in respect of the use of mediations it seems that around two-thirds of firms/organisations would consider them (66%), whereas 18% of respondents' firms have no policy or practice in respect to mediation. This finding chimes with recent work suggesting that mediation is being increasingly seen as a legitimate tool for the resolution of disputes in Scottish legal circles (Clark, 2009)

Experience of the use of Mediation

In terms of experiences of mediation, respondents were asked to rate their responses to a number of questions ranging from client representation, levels of satisfaction with different elements of the process, factors leading to a failed mediation process, the decisions to recommend mediation to a client and reasons to refuse mediation proposals from the opposing party in a case.

2.1.1 Client representation in mediation

In terms of those respondents who had working experience of mediation (58% of the total number of respondents), 97 % had represented a party in 3 or more cases, 44% in 5 or more cases and 21 % in 10 or more mediation cases.

In terms of type of case in which respondents were involved it seems that those relating to payment (24 % of the total number of cases) and damage (22% of total cases) were most common with those related to professional negligence (20% of total cases), change of scope (13% of total cases), delay (10% of total cases) less so. The percentage of disputes relating to payment seems much lower than might perhaps than expected, given the frequency within which these issues are often encountered within the construction context. This may demonstrate the impact of statutory adjudication in reducing the number of payment disputes that are mediated, as for example suggested in a study of construction mediation in England (Gould et al, 2009) . Equally it may be that payment disputes are more amenable to early resolution between lawyers. Of the most common issues payment and damages the total number of cases wholly settled was 35 and 36 respectively. The number of payment-related cases that were partially settled or not settled at all was 16, whereas the number of damages-related cases partially or not settled was 10. In respect of the other types of case, it seems that for professional negligence –related issues, respondents’ indicated that the majority (86 %) of the total numbers of cases were settled. The equivalent figure for change of scope and delay was 60 and 63 percent respectively.

It seems that damage-related cases are by far the highest percentage of cases settled by mediation according to respondents, 90%. The type of case recording the lowest number of settlement as indicated by respondents was for delay-related issues (57%). Seemingly, there is a tendency for the most common types of case encountered, i.e. those related damage, payment and professional negligence to be wholly settled through mediation. This might be because these issues are much less contentious than for example defect-related matters which are inherently more complex and rely more heavily on expert evidence (Gould et al, 2009). Table 1 illustrates the breakdown of the total number of times respondents’ have represented a party in mediation by type of case and the number of those cases that wholly settled and partially settled or not settled.

Table 1: Breakdown of the total number of times respondents’ have represented a party in mediation by type of case and the either number wholly, partially or not settled.

Types of case	Total number of cases	Number of cases that settled	Number of cases not settled	Number of cases partially settled
Professional negligence	37	32	4	1
Change to scope of work (extra work)	28	17	8	3
Payment Issues	51	35	9	7
Delay	19	12	4	3
Damages	40	36	2	2
Others	3	0	3	0
Total	178	132	30	16

Satisfaction with different elements of the mediation process

Overall, it would seem that a large proportion of the respondents who had experience of mediation were satisfied with the mediation process. Around 33.3% were always satisfied and 43.3% often satisfied with the speed of the process. Around 27% and 43% of respondents were either always satisfied or often satisfied with the cost of mediation respectively. A closer examination of the results of the survey also indicates that the respondents were more satisfied than not with the mediator and the outcome of the mediation. In particular, almost 50% of the respondents were often satisfied with the mediator and 63.3% often satisfied with the outcome of mediation. Table 2 provides a breakdown of frequency and percentage response for respondents' satisfaction with different elements of the mediation process.

Table 2: Breakdown of frequency and percentage response for respondents' satisfaction with different elements of the mediation process

Element of the mediation process	Always satisfied	Often satisfied	Sometimes satisfied	Rarely satisfied	Never satisfied	Total (%)
The speed of mediation	10 33.3%	13 43.3%	7 23.3%	0	0	30 100%
The cost of mediation	8 26.6%	13 43.3%	6 20%	3 10%	0	30 100%
The mediator	5 16.6%	14 46.6%	10 33.3%	1 3.3%	0	30 100%
Outcome of mediation	4 13.3%	19 63.3%	7 23.3%	0	0	30 100%

Further statistical analysis of the data would seem to indicate a relationship and positive correlation between the satisfaction of respondents and the number of times respondents have represented a party in a mediation (CS sig value $p = 0.00 < \alpha = 0.05$; CV $r > 0.76$ and CC $r > 0.70$). Thus, it would seem that the level of the satisfaction of respondents increases as their level of use of mediation increases, as would be expected. Equally it is likely that those who have positive experiences with mediation may typically reuse the process.

Factors leading to a failed mediation

In terms of the factors leading to a failed mediation, it seems that there is little agreement from the respondents with experience any failed mediation, 15 in total, as to the contributory factors leading to failure. For instance, despite the fact that client factors were cited more heavily than other factors, only 33.3 % of respondents always attributed the failure of mediation to parties' entrenched or polarised positions and 26.6% to bad feeling between disputing parties. Almost half the respondents, 46.6% and around one-third often attributed failure to parties' positions, bad feeling and unrealistic expectations of the parties to the failure of the mediation process. Almost one-half of the respondents, 42% in total, rarely and never attributed failure of the mediation to the lack of skills of the mediator.

Interestingly, though around one-third, 28% in total did always or often attribute the failure of the process to the skill of the mediator. There would seem to be a polarisation of views in relation to this particular question, particular in regard to the levels of approval for the mediator as a critical success factor in mediation, illustrated in Table 2. It would be interesting to see if there is some correlation between this view and that of industry's clients in respect to the role of the mediator, particularly given that there is available evidence from England that industry participant perceive the involvement of a third party as detrimental to the whole process of dispute resolution in the construction industry (Brooker, 2009)

Table 3 provides a breakdown of frequency and percentage response of the contributory factors leading to a failed mediation.

Table 3: Breakdown of frequency and percentage response of the contributory factors leading to a failed mediation

Contributory factors leading to a failed mediation	Always	Often	Sometimes	Rarely	Never	Total (%)
Unrealistic expectations of parties	1 7%	4 29%	8 57%	0 0%	1 7%	14 100%
Parties entrenched/polarised in their position	5 33.3%	7 46.6%	3 20%	0 0%	0 0%	15 100%
Bad feeling between the parties	4 26.6%	5 33.3%	6 40%	0 0%	0 0%	15 100%
Lack of skills of the mediator	1 7%	3 21%	4 29%	3 21%	3 21%	14 100%
Mediation was only being used tactically	0 0%	1 8%	5 38%	3 23%	4 31%	13 100%
Conflict of evidence	0 0%	1 8%	2 15%	5 38%	5 38%	13 100%

The decision to recommend Mediation to a Client

In terms of the decision to recommend the mediation process to a client, it seems from the analysis of the respondent data that there are several critical and determining factors. It can be seen from Table 4 that these include reduction in client legal costs, achieving a speedier settlement and the possibility of achieving a creative settlement to a dispute. Indeed, respondents were more likely to have always considered reduction in legal costs, a speedier settlement and a creative one to the decision to recommend mediation to a client as relevant than any other factors.

These findings are consistent with those shown in Table 2. However, other factors associated with the decision to recommend mediation such as gaining information on the other side's case, narrowing the issues in a dispute and weakness's in a client case seem much less relevant and critical in respect of a decision to recommend mediation.

Nevertheless, the respondents considered that weakness in a client case, narrowing issues as well as enabling continuation were sometimes relevant to a decision. Interestingly, the low size of financial sum was sometimes considered relevant to the decision to recommend mediation to a client. The fact that speed and cost were considered relevant and critical factors in the decision to recommend mediation may suggest a close association or correlation between the decision to recommend and a representation of a client. The analysis does indeed indicate that there is a statistical association and relationship between the decision to recommend mediation to a client and the representation of a client in a mediation (CS sig values $p = 0.00 < \alpha = 0.05$; CV $r > 0.70$; CC $r > 0.70$). This analysis would seem to suggest that those more likely to recommend mediation to a client are more likely to represent a client in mediation. Thus, experience and a track record in mediation would seem to be critical factors for legal counsel in respect to their recommendation to clients. Table 4 illustrates the frequency and percentage response of relevant factors to the decision to recommend mediation to a client.

Table 4: Breakdown of the frequency and percentage response of relevant factors to the decision to recommend mediation to a client

Factor relevant to the decision to recommend mediation to a client	Always relevant	Often relevant	Sometimes relevant	Rarely relevant	Never relevant	Total (%)
A reduction of legal costs for your client	20 57%	11 31%	4 11%	0 0%	0 0%	35 100%
Low size of the financial sum in dispute	7 20%	3 9%	16 46%	6 17%	3 9%	35 100%
Achieving a speedier settlement	17 49%	16 46%	1 3%	1 3%	0 0%	35 100%
The possibility of reaching a creative settlement	8 23%	14 40%	13 37%	0 0%	0 0%	35 100%
The possibility of assessing the risk of continuing the dispute	7 20%	11 31%	11 31%	5 14%	1 3%	35 100%
A weakness in a client's case	3 3%	4 11%	18 51%	8 23%	2 6%	35 100%
Narrowing the issues in dispute during mediation	5 14%	8 23%	15 43%	4 11%	3 3%	35 100%
Enabling continuation of a business relationship	6 17%	8 23%	15 43%	4 11%	2 6%	35 100%
Gaining information on the other side's case	4 11%	4 11%	7 20%	16 46%	4 11%	35 100%
The privacy of mediation	7 20%	9 26%	9 26%	7 20%	3 3%	35 100%

The decision to refuse a proposal for mediation from the opposing party in the dispute

In respect to the decision to refuse a proposal for mediation, around one-third, 31%, of the respondents always considered the clients' wishes not to use mediation as relevant. Interestingly, it could be that at least some clients are averse to the whole process of mediation involving a third party, preferring rather to settle disputes on an amicable basis through negotiation or failing that seeking a ruling from an authoritative third party. It could also be that clients may have a negative perception of the process or harbor misconceptions of the process. Further research probing the views of the client base may be required to learn more about the perception of disputants regarding mediation. Any client reluctance to use mediation begs the question as to who controls the decision to mediate or otherwise – lawyer or client? While the issue of the power relationship between lawyers and clients in the Scottish construction industry will be explored further by the authors in follow up interviews, for a general discussion of this issue, see Clark (2009). About one-fifth of the survey respondents considered the belief that the opposing party would not take part in good faith as always relevant. There also seems to be a general belief among the sample that negotiation was capable of settling a case, but only on some occasions was this considered relevant. By far the largest proportion of respondents, 50%, considered belief in the strength of a client's case as a relevant factor in the decision to refuse a proposal, but only sometimes. The belief that recovery of documents was essential before reaching settlement was considered much less relevant by the respondents, with 38% considering it a relevant factor only sometimes and 38% considering it rarely relevant.

Table 5: Breakdown of frequency and percentage response of factors determining the decision to refuse a proposal for mediation from the opposing party in the dispute

Factors determining the decision to refuse a proposal for mediation from the opposing party in the dispute	Always relevant	Often relevant	Sometimes relevant	Rarely relevant	Never relevant	Total (%)
Client did not want to use mediation	8 31%	7 27%	9 35%	2 8%	0 0%	26 100%
Belief in the strength of the client's case	1 4%	5 19%	13 50%	5 19%	2 8%	26 100%
Belief that the opposing party would not take part in good faith	5 19%	6 23%	9 35%	4 15%	2 8%	26 100%
Case type not appropriate for mediation	3 12%	8 31%	9 35%	4 15%	2 8%	26 100%
Belief that negotiation was capable of settling the case	2 8%	6 23%	12 46%	5 19%	1 4%	26 100%
Belief that recovery of documents was essential before reaching settlement	1 4%	1 4%	10 39%	10 39%	4 15%	26 100%

Attitudes to Mediation

The purpose of this section was to ascertain respondents' attitudes to various key policy issues relative to construction mediation's development. A summary of initial findings in this respect are set out below. The analysis was undertaken using ratings based on 5-point Likert scale with anchors ranging from strongly agree to strongly disagree in which respondents were asked to rate their response to a number of statements in respect to their attitudes to dispute resolution more generally and to mediation specifically. Table 6 illustrates the frequency and percentage response of respondents' attitudes to dispute resolution and mediation. In terms of different types of dispute resolution mechanisms available, 23 % of the respondents somewhat disagreed that litigation was generally well adapted to the needs and practices of the construction community, whereas around one-third, 32%, somewhat agreed. Perhaps, this reflects in part clients' perceptions of the lengthy time and high costs associated with litigation and the public nature of the process as compared to mediation.

There was no indication that the use of mediation in the construction context was a sign of weakness. A large proportion of the respondents somewhat disagreed, 40% or strongly disagreed, 52%, with suggesting mediation to an opponent was a sign of weakness (52%). A high proportion of the respondents considered Arbitration unsuited to the needs and practices within Construction. A striking figure, some 80% of the respondents somewhat disagreed and strongly disagreed that Arbitration was generally well adapted to the needs and practices of the construction community. Whether the recent reforms to the process heralded by the new Arbitration Act 2010 will alleviate such concerns remains unclear (see Dundas, 2010). By contrast, 84% strongly or somewhat agreed that Adjudication was generally well adapted to the construction context. At face value then it could be suspected that the popularity of this default mechanism in many construction contracts may render mediation unnecessary in the construction sphere. Nevertheless, only one-third of respondents, 34%, considered that the default to Adjudication in many construction disputes would render mediation obsolete. This finding may suggest that in some circumstances respondents would recommend mediation to clients as an alternative to more embedded methods of dispute resolution such as adjudication. The authors will endeavour to ascertain more about the relationship between adjudication and mediation in follow-up interviews. In terms of measures designed to aid the further entrenchment of mediation as a dispute resolution process in construction matters, a large proportion of the respondents, 62%, strongly and somewhat agreed that construction contracts should have a mediation clause, such as those provided within JCT Standard Forms of Building Contract as well as ICE contract versions². A small majority of the respondents, 54 %, strongly or somewhat agreed that judges should refer more cases to mediation. Nevertheless, opinion was split on whether making mediation a mandatory first step would be a positive development. Whereas, 44% strongly and somewhat agreed with the statement, 46% somewhat and strongly disagreed.

[2] Examples of how mediation is addressed in Construction Contracts include JCT Standard Form of Building Contract (2005) Section 9.1 and ICE Conditions of Contract – Measurement Value 7th Edition (1999) Clause 66A(2)(a)

The issue of court referral to mediation is of long standing controversy. Unlike other jurisdictions, e.g. England and Wales and many parts of the USA, where court initiation and even compulsion (in the USA) is commonplace, Scotland has had only very limited experience with court referral of the process. While many would argue that in keeping with the original ethos of mediation, participation in the process should remain purely voluntary, others have pointed to the shot in the arm that court initiation can provide mediation.

Clearly the current research indicates a significant opinion within Scottish legal circles desirous of increased court promotion. For a summary of the debates around court promotion of mediation see Clark (2008). Interestingly, opinion was split equally on Scots lawyers' awareness of mediation. Equal numbers, 44 %, of respondents strongly or somewhat agreed with the statement and also strongly or somewhat disagreed. Although the vast majority of respondents felt able to explain mediation to a client, it may be that a higher proportion of those that did not respond were less familiar with mediation, thus reflecting the current finding. In order to alleviate such reported ignorance, interestingly, only a small majority, 52%, strongly and somewhat agreed that mediation training should be compulsory. In fact only 18% strongly agreed with the proposition. In terms of the how mediation fits with cultural norms and the usual *modus operandi* of lawyers, it seems that an overwhelming majority disagreed that widespread use of mediation would be detrimental to the legal profession in Scotland. Indeed, when respondents were asked to whether a lawyer's standing amongst colleagues would suffer as a consequence of their involvement with mediation, a large majority, 74%, disagreed that there would be a negative perception. It could be that the perception of the legal fraternity as inherently conservative towards non-court sanctioned dispute resolution is unfounded, or that perhaps this was the most socially-acceptable response? Other opinions expressed by respondents seem to suggest that former rather than the latter. Indeed, a large majority, 88%, strongly or somewhat agreed that mediation would provide lawyers with an opportunity to offer further services to their clients. Similarly, many of the respondents also considered mediation as an opportunity to enhance their fee earning potential, with 78% strongly disagreeing that the growth of mediation would be detrimental to future earnings. While such findings may lend weight to the argument that lawyers may seek to 'milk' mediation there is nothing inherently wrong with professional seeking to access new markets, and there currently exists scant evidence of lawyer highjacking of construction mediation in Scotland.

Table 6: Breakdown of frequency and percentage response of respondents' attitudes to mediation

Attitudes	Strongly agree	Somewhat agree	Somewhat disagree	Strongly disagree	Don't know	Total (%)
If a lawyer participated more often in mediation his/her standing amongst colleagues would suffer	0 0%	1 2%	11 22%	37 74%	1 2%	50 100%
Mediation is detrimental to the development of law	3 6%	19 38%	13 26%	12 24%	3 6%	50 100%
Mediation is inappropriate where there is a power imbalance between the parties	1 2%	8 16%	27 54%	13 26%	1 2%	50 100%
Judges should refer more cases to mediation	2 4%	26 52%	5 10%	12 24%	5 10%	50 100%
Making mediation a mandatory first step would be a positive development	2 4%	20 40%	7 14%	16 32%	5 10%	50 100%
Legal practitioners make the best mediators	1 2%	16 32%	21 42%	1 2%	11 22%	50 100%
Litigation is generally well adapted to the needs and practices of the construction community	2 4%	16 32%	23 46%	9 18%	0 0%	50 100%
Arbitration is generally well adapted to the needs and practices of the construction community	1 2%	8 16%	23 46%	17 34%	1 2%	50 100%
Adjudication is generally well adapted to the needs and practices of the construction community	12 24%	30 60%	7 14%	1 2%	0 0%	50 100%
Default to adjudication in many construction disputes renders mediation obsolete	0 0%	17 34%	17 34%	14 28%	2 4%	50 100%
Mediation suffers from a lack of coercive power	1 2%	12 24%	20 40%	14 28%	3 6%	50 100%
Mediation is an opportunity for lawyers to offer further services to their clients	11 22%	33 66%	2 4%	2 4%	2 4%	50 100%
Lawyers will lose money if mediation grows	0 0%	3 6%	23 46%	16 32%	8 16%	50 100%
Suggesting mediation to an opponent is a sign of weakness	0 0%	1 2%	20 40%	26 52%	3 6%	50 100%
Construction contracts should contain a mediation clause	6 12%	25 50%	9 18%	5 10%	5 10%	50 100%
A barrier to mediation's development in Scotland is its negative perception among clients	2 4%	17 34%	15 30%	4 8%	12 24%	50 100%
A barrier to mediation's development in Scotland is its negative perception among lawyers	1 2%	12 24%	26 52%	5 10%	6 12%	50 100%
Mediation training should be compulsory for lawyers	9 18%	17 34%	14 28%	5 10%	5 10%	50 100%
There is a lack of awareness regarding mediation amongst the legal fraternity in Scotland	1 2%	21 42%	17 34%	5 10%	6 12%	50 100%
Mediation is of more utility in low value disputes	1 2%	4 8%	20 40%	19 38%	6 12%	50 100%

Summary and Conclusions

Construction disputes by their very nature are often complex, sometimes multi-party disputes, many of which are not suited to either adjudication or traditional forms of dispute resolution (these being slow and expensive). UK surveys of construction lawyers have so far confirmed mediation a suitable forum for such disputes, the opinion being it can be effective in all types of construction disputes irrespective of the relationships involved. Also, with the advantages of mediation in construction dispute resolution having been long recognised in other jurisdictions, it is difficult to imagine similar recognition not developing within Scotland. Our study suggests that this process is at least beginning and while the total recorded cases remain low³, increasing numbers of construction lawyers hold many positive perceptions regarding the benefits of using mediation, seeing it as a well established part of their business. It should be noted that this initial study serves as a starting point for our future research. The initial study is cross-sectional in nature in that it represents a *snapshot* in time. It is recognized that the introduction of mediatory techniques into construction disputes will have a cumulative effect on the Scottish legal fraternity over time. Cross-sectional studies are often unable to yield information about the direction of causal relationships between variables that are interrelated in a complex way. Neither do cross-sectional studies permit researchers to assess the effectiveness of intervention strategies. Therefore, the need for follow-up studies, which identify the effect of policy and organizational change over time, is recognized. Although, it is also recognised that on-going changes to the dispute resolution landscape will add more uncertainties into the context, the findings of the initial study will act as a springboard from which a more extensive follow-up study will be undertaken. In-depth interviews exploring selected relationships in depth and detail will be a key component of the follow-up study. While it is also recognized that the interview sample will be much smaller and the findings less generalisable than those from the questionnaire survey, these qualitative data will supplement the quantitative data to offer insightful explanations of the use of mediation in construction disputes.

References

- Agapiou, A (2010) Scottish Construction Lawyers and Mediation, Robert Gordon University, Unpublished LLM Dissertation.
- Boussabaine A. H., Elhag, (1999) A Neural Network approach for forecasting Cost-Flow for water industry projects, *Journal of Engineering, Construction and Architectural Management*, Vol. 6 No. 3, pp.213-224.
- Brooker, P and Lavers, A (1998) Perceptions of alternative dispute resolution as constraints upon its use in the UK Construction industry, *Construction Management and Economics* 15, 519-529.
- Brooker, P (2005) Construction lawyers' experience with mediation post-CPR 21 *Construction law Journal* 19-46

[3] Clearly there may be double counting in the figures, in that different respondents may be referring to the same case. Equally our methodology in no way has allowed us to track all construction mediations in Scotland that have taken place.

- Brooker, P (2009) Criteria for the appropriate use of mediation in construction disputes: Judicial statements in the English Technology and Construction Court, *International Journal of Law in the Built Environment*, Volume 1 (1), p82-97
- Clark, B and Dawson, C (2007) Scottish Commercial Lawyers and ADR: Awareness, Attitudes and Experience. *Civil Justice Quarterly*, April
- Clark, B (2009) Mediation and Scots Lawyers: Past, Present and Future, May, *Edinburgh Law Review*
- Clark, B (2008) Institutionalising Mediation in Scotland, 3, *Juridical Review* 193
- Coakes, S.J. (2005) *SPSS: Analysis Without Anguish: Version 12.0 for Windows*, Wiley
- Dundas, H (2010) The Arbitration (Scotland) Act 2010: Converting Vision into Reality, *Arbitration*, 76, p 2-15
- Gould, N. and Ors (2009) *The Use of Mediation in Construction Disputes: Summary Report of the Final Results*, King's College London and HMCS, 7 May 2009
- Gould, N, Capper, P, Dixon, G and Cohen, M (1999) *Dispute resolution in the construction industry*, 1st Edition, Thomas Telford, ISBN 0727728369.
- Idrus, A.B. and Newman, J.B. (2002) 'Construction related factors influencing choice of concrete floor systems', *Construction Management and Economics*, 20(1), pp 13-19
- Oladapo, A & Onabanjo, B (2009) A study of the causes and resolution of disputes in the Nigerian Construction Industry, RICS COBRA Research Conference, University of Cape Town, 10-11th September, pp 7-22.
- Piaw, C. H (2006) *Basic Research Statistics* (2006), McGraw-Hill Education
- Sarat, Austin D. and William L.F. Felstiner. (1988) Law and Social Relations: Vocabularies of Motive in Lawyer/Client Interaction. *Law & Society Review*. 22:4:737-770.
- Stipanowich (1994) What's Hot and What's Not. DART Conference Proceedings, Lexington, Kentucky, US. October 16-19.
- Zulhabri, J, Jaalunlaili Abdullah, and Rosli Mohamad Zin (2005) Findings of Alternative Dispute Resolution Application and obstacles towards active development of ADR in the Malaysian Construction industry, 3rd Conference of Law and Technology.

Mediators: Towards a code of good conduct for construction mediators

Penny Brooker

Director of Law Research Centre
School of Law Social Science and Communication,
University of Wolverhampton, Wolverhampton, WV1 1SB ,
United Kingdom

Email: P.J.Brooker@wlv.ac.uk.

Abstract:

Purpose: This paper considers the desirability and practicality of a standard code of conduct for construction mediators in England and Wales, which would accommodate both facilitative and evaluative mediator orientation.

Content: A critical analysis is made of construction mediator codes of conduct in England and Wales and compares these to codes from other common law countries, the EU Mediators' Code of Conduct and the International Mediation Institute. The uptake of construction mediation grew significantly in the English and Welsh jurisdiction with the introduction of court rules, which permit the award of costs against a successful party if they unreasonably refuse to mediate (Brooker and Laver, 2005a, b). Construction mediation is now dominated by a small group of experienced mediators from a legal background (Gould et al, 2009; 2010) and this predominance of lawyers is reported to be connected to a rise in evaluative mediator techniques (Brooker, 2007). In the USA, the growth of evaluative mediators led to a critical debate on mediator orientation and the call for standards of good practice, which reflect these differences (See for example; Riskin, 1996). The status of mediation is enhanced as the courts increase opportunities for mediating which has led to the development of a mediator profession concerned with standards of 'acceptable' or 'appropriate practice' (See for example in the US, Menkel-Meadows, 2009; In England and Wales Boon et al, 2007). In the English jurisdiction there is no 'overarching' set of standards in place but as mediation growth continues with 'court approval' (Clark and Mays, 1996; Roberts, 2002; Boon et Al, 2007; Dolder, 2004), the schism between mediator approaches has become both more apparent and critical to the ongoing debate for a nationally recognised standard of mediator practice.

Conclusion: This paper considers whether the 'Codes of Conduct' currently in use are sufficiently well defined to ensure participants are informed about mediator style (evaluative or facilitative) thus enabling self-determination and concludes that a lack of clear guidelines may obscure informed consent.

Key Words: Mediation; Mediator Codes of Conduct; standards; training; self-determination; informed consent.

1. Introduction

1.1 Mediation Expansion

Many countries have introduced mediation to provide an alternative to the traditional or 'formal' dispute resolution procedures such as litigation and arbitration, which often involve complex procedural and evidential rules and a reliance on costly legal practices (See for example: Woolf, 1995; Alexander, 2006). The Civil Procedure Rules (CPR) in England and Wales now support mediation because ostensibly it offers an inexpensive, speedier and less confrontational procedure for settlement whilst the parties maintain control over their dispute and their settlement outcome (Genn, 2006; Roberts, 2000). Reforms to civil litigation in many countries are reported to have led to a substantial growth in construction mediation and groups of construction mediators are reported to be developing practice and standards of provision (Brooker and Wilkinson, 2010; Gould, 2010).

This paper considers the expansion of construction mediation in England and Wales and the specific advent of two 'styles' of mediator practice (For an explanation see, Riskin, 1996): facilitative and evaluative, which are now utilised in the industry (Brooker, 2007; Gould et al, 2009; 2010). In facilitative practice, mediators use their skills of listening, questioning and reframing to help the parties 'communicate' and examine their case in order to reach their own settlement outcomes (See National Mediation Network Website, 2010; Riskin, 1996; Menkel Meadows, 1993; Folberg and Taylor, 1984 at 7; Roberts, 2002). An evaluative mediator offers 'information and opinions on the substance of the dispute' (See Guthrie, 2001: 146; Riskin, 1996). 'Interventions' might include 'imparting a view over the merits of the case or of particular issues'; (CEDR Website, 2010) assessing the strength and weaknesses of presented cases; making predictions about a court outcome or providing settlement proposals with the party's consent (See Riskin, 1996).

In some construction literature in the UK the term 'evaluative mediation' has been used to describe a process or model where the mediator gives a recommendation at the end of the procedure if the parties fail to reach a settlement (See Brooker, 2007: 221; Gaitskell 2005a, b; Kendall, 2000), which is also sometimes called 'conciliation'. In Australia, which has implemented a voluntary accreditation system for mediators, processes where a neutral 'gives advice' are referred to as 'conciliation' or 'advisory mediation' or 'evaluative mediation' (Alexander, 2008a: 119). In the context of this paper a distinction is drawn between 'evaluative mediation' as a model and 'evaluative mediators' who use interventions or approaches (Alexander, 2008a; Riskin, 1996) which are characterised as evaluative (See discussion in Brooker, 2007).

The significance of the distinction in these two approaches is that when mediators use evaluative interventions it said to 'undermine' the parties' self-determination on which mediation is predicated (Alfini, 2008; Kovach and Love, 1996, 1998; Stulberg, 1997; Alfini, 2008; Welsh, 2001; Hadeen, 2005; McCorkle, 2005). Welch (2001: 4) categorizes the attributes of self-determination to be when the parties: 'actively participate' in the process; 'control the decision making'; 'create settlement outcomes' and finally 'decide whether to settle or not'. Alfini (2008: 830) believes that self-determination is what differentiates mediation from 'adjudicatory' dispute resolution procedures such as litigation and arbitration, and now construction

adjudication, where the parties do not usually establish the process, rules and outcome. Mediator skills, he believes, should be employed to support the party's self-determination and their 'actions' should not be 'inconsistent with this core value' (Alfini, 2008: 831).

In the USA empirical evidence suggests lawyers become more involved in mediation as it becomes connected to the courts which is linked to a rise in evaluative mediators (See for example; McAdoo and Hinshaw, 2002; Wissler, 2004; Thompson, 2004; Hensler, 2001, 2002, 2003; Welsh, 2001a, b; Menkel-Meadows, 1997; Nolan-Haley, 1998, 2002; Brad Reich, 2002; Stipanowich, 1998; Roberts, 2002; Coben and Thompson, 2006; Press, 2003). Some theorists argue that if evaluative mediators are 'here to stay' (Baruch Bush, 2002: 124) then training, standards of practice and Codes of Conduct should 'reflect' these different approaches (See for example, Riskin, 1996). This would raise awareness about 'mediator style' and consequently inform the parties' choice about which mediator to appoint (See Riskin, 1996: 48-49; Barush Bush, 2002; Barush Bush & Pope, 2002). Parties who are 'educated' (Coolie, 2000 at 77-78) about the process and a mediator's approach exercise self-determination with informed consent (See for example; Riskin, 1996; Cooley and Love, 2008; Nolan-Haley, 1999).

1.2 Legal Involvement in Mediation in England and Wales

The predominance of the legal professions in mediation developments in England and Wales was observed as early as the 1990s (Roberts, 1992, 1993) and this position does not appear to be radically changing (Roberts, 2009). Sixty per cent of mediators responding to CEDR's Audit (2010) (one of the first organisations to provide training in the UK) come from a legal background. These findings are replicated in TCC mediation survey where the majority of mediations were undertaken by lawyers (41 per cent by solicitors; 34 per cent by barristers and 7 per cent by judges) and only 16 per cent were reported to be undertaken by construction professionals (Gould et al, 2010: 10-11; See also Bucklow, 2007). Both the CEDR Audit (2010) and the TCC survey suggests that as parties become more familiar with mediation they begin to shun appointing bodies by making direct appointments either with mediators they have used before or from recommendations (Gould et al, 2010). In the TCC survey only 20 per cent of recorded mediations went through ADR providers and the findings suggest that a relatively small core of legal experts are operating in this forum (Gould et al, 2010: 61). Brooker (2007: 233) suggests that the influence of the legal professions is likely to be associated with a rise in evaluative mediators and an increase in adversarial behaviour in construction mediation and to avoid 'future censure' lawyers should 'safeguard self-determination by raising their clients' awareness of mediator style' which would support a '*bone fide*' choice based on informed consent.

1.3 Professionalisation of mediation

As mediation has been given a prominent role by the courts this raises its status and those who practise as mediators are seeking to professionalise their work (Boon et al, 2007; Gould et al 2009; 2010; Menkel Meadows, 2009; Brooker and Wilkinson, 2010; Grossman, 2003; Miles, 2004; Brady, 2009; Nolan-Haley, 2002). Boon et al (2007: 49) recognise the 'hallmarks' of a mediator profession in England and Wales, '*where the proliferation of competing groups is common. The field is rapidly*

acquiring social importance: there is a body of practitioners with expertise that is distinctive; and many are subject to training and ethical codes’.

Professional status is also associated with groups limiting others from practising as they develop systems of ‘restrictive accreditation’ which are aligned to codes of ‘best practice and ethics’ with the aim of excluding others from their domain (See for example, Menkel Meadows, 2009: 195; Clark, 2009).

1.3 Aim of paper

The purpose of this paper is to first, establish which ‘Codes of Conduct’ are observed by construction mediators in England and Wales with the aim of assessing whether they support party self-determination and are sufficiently well defined and prominent to ensure that users are informed about the mediation process and mediator’s style. Second, to compare the Codes with those from other Common Law countries; The European Union (EU) Mediator Code of Conduct (EU Code) and the International Mediation Institute’s Code (IMI Code) in order to determine the desirability and practicality of a common Code of Conduct in the English jurisdiction, which would accommodate both facilitative and evaluative mediators’ orientations (Riskin, 1996).

2. Background

2.1 Civil Procedure Rules & the TCC

The uptake of construction mediation grew in the UK with the introduction of the Civil Procedure Rules (CPR) and Construction and Engineering pre-action protocol (CEPP) which encourage litigants to contemplate and use mediation in ‘appropriate cases’ (See for example; Brooker, 2009, 2010a, b; Brunsdon-Tully, 2009; Shipman, 2006; Sorabji, 2008). Substantial interest in mediation was not realized until the court applied a costs penalty under CPR against a successful party who had unreasonably refused to mediate in *Dunnett*¹ (See for example; Brooker and Lavers, 2005a, b; Gould et Al, 2009; 2010). Although Genn (2007: 8) argues, that the ‘cautious’ approach by the Court of Appeal in *Halsey*² may have negated expansion by providing guidelines for ‘unreasonable refusals’ rather than continuing the ‘relentless judicial pressure’ of earlier cases. However, a recent survey of mediation in the TCC suggests that more mediation is taking place than had been appreciated (Gould et Al, 2007, 2009 and 2010).

The TCC Court Guide (2005 revised 2007) states that ‘Legal representatives in all TCC cases should ensure that their clients are fully aware of the benefits of ADR and that the use of ADR has been carefully considered prior to the first CMM (Court Management Meeting)’ (7.1.3). The court does not recommend ‘specific’ ADR procedures, or particular organisations or individual mediators, although the Court Guide does indicate where information is available such as ‘TeCSA (Technology and Construction Solicitors Association); Technology and Construction Bar (TECBAR); the Civil Mediation Council and from some TCC court centres outside London’ (7.3.4) (See for example, Brooker, 2009).

¹ *Dunnett v Railtrack PLC* (2002) C.P. Rep. 35

² *Halsey v. Milton Keynes General NHS* [2004] EWCA Civ 576

Mediation therefore remains undefined by the court rules leaving the choice of mediator approach to the parties' discretion. It is likely that parties who reach an impasse when negotiating a dispute or have not used any form of ADR or mediation previously may rely on their legal advisors, who may be 'repeat players' (Gould et Al, 2010: 68). Therefore, at this stage the selection of mediator may be influenced by the lawyers' preference or experience (Gould et Al, 2010; Brooker, 2007; 2009; 2010a, b).

2.2 Regulation and accreditation of mediation

Alexander (2008b: 2) considers the regulation debate to be 'one of the most controversial topics' in the Professionalisation of mediation, primarily because of the 'diversity – consistency dilemma' where the divergent interests of the mediator to maintain 'flexibility' and 'innovative practice' competes with the need to protect consumers through assuring quality. Alexander (2008b: 2) suggests that delineating a 'consistent definition' for mediation is central to the debate because this may exclude some mediation practices (See for example, Mills, 2005) and further, specific to the question of accountability is whether mediator providers should inform users of the 'character of the mediation process' and 'qualifications and skills of their mediators'.

In Australia the steer for regulation led in 2008 to a voluntary scheme where mediators attain accredited status by registering with a Recognised Mediator Accreditation Body whose training and standards comply with the national system (See for example, Alexander, 2008b: 8). A major area of contention before implementation was arriving at a definition of mediation (See Sourdin, 2007). The final version of the 'Mediator Guidelines' (2008) defines mediation as 'primarily facilitative' (3) and the mediator's role one of 'support(ing) the participants to make decisions about further actions and outcomes' (Clause 2.3). Part 1.2(e) specifically states that mediators should engage in mediation 'in a way that addresses the participant's mutual needs with respect to their individual interests based on the principle of self-determination'. Mediators who provide an 'advisory component' or 'expert information' are required to show a 'higher level' of training and must 'inform' the parties about this approach and 'obtain clear consent' (section 2.4) (See Alexander, 2008a: 119). Mediators who use a 'blended approach' must provide detailed evidence of their qualifications to undertake an evaluative role which includes 'registration' or 'membership' in 'the professional area in which advice is to be given' (S5(4)(a)).

2.2.1 Regulation Debate

Brady (2009: 392) notes the 'plethora' of organisations in the UK offering training and panels of mediators many of which, she states, have developed their own Codes of Practice. Commentators observe that the State does not regulate mediation in England and Wales, nor does any one professional body have 'overall control' over accreditation and there is no single 'code of conduct', which is observed by mediators (See Boon et Al, 2007; Dolder, 2004; Brady, 2009). Arguments against regulation are frequently based on 'market demand' which, it is suggested, controls development and the use of mediation (Boon et Al, 2007: 48-49). For example, in construction, Gould et al (2010: 66) report that evidence from the TCC mediation survey does not indicate poor mediator standards necessitating state regulation or accreditation from

specific training organisations. Users selected mediators through ‘well informed choice’ and there was a ‘pattern’ of repeat appointments indicating that mediators do not ‘abuse their role’ (Gould et Al, 2010: 66).

Arguments formulated against regulation include (See Boon et Al, 2007):

- ‘inhibiting’ the market
- ‘stifling development’
- ‘commitment to one type of practice
- ‘increasing the costs of qualification’
- ‘excluding talent’ and;
- ‘prioritising settlement as the main outcome of mediation’

Alexander (2008b: 2) however, observes that allowing ‘market demand’ and mediators a free reign on developing practice results in an ‘abundance’ of different approaches which consumers may not always be ‘sufficiently informed’ about. Many argue that mediators should be ‘accountable’ when users are ‘vulnerable’ or there is an ‘unequal power difference’ (See for example; Alexander 2008b; Boon et al 2007) and Clark and Mays (1996: Conclusion) suggest that consumers may need protection as mediation becomes ‘enmeshed within the legal framework’ either as a requirement for litigation or refusing legal aid.

Arguments for regulating mediation practice are concerned with (See Boon et Al, 2007):

- ‘Protecting consumers’
- ‘Excluding incompetent practice’
- ‘Clarifying personnel, procedure and costs’
- ‘Standardising practice’
- Helping parties ‘to seek compensation from the courts’

Construction mediation may involve ‘sophisticated’ participants where there is no clear ‘power in-balance’ thus not reaching the litmus test of ‘vulnerability’ for protection (Gould et Al, 2010: 66). Nevertheless, such research as exists in the construction field supports the contention that mediator practice is not uniform (Boon et Al, 2007; Brooker, 2007; Gould et Al, 2009; 2010; Bucklow, 2007) either in England and Wales or elsewhere (Stipanowich, 1996; Brooker and Wilkinson, 2010). When parties lack information about construction mediation then Bowie (2004: 33) suggests they can either feel ‘threatened’ when the mediator take an evaluative stance or ‘disappointed’ if the mediator remains facilitative which can be avoided if ‘*all parties enter into the process with the same set of expectations and understandings*’.

Lawyers are likely to be involved in construction mediation at a number of stages: advising on its use, representing their clients during the process and finally as the chosen mediator. Brooker (2007: 233) suggests that if mediation is to achieve its ‘maximum potential’ in England and Wales choice of mediator style must be addressed by basing ‘selection on information formed on the basis of transparent standards of practice and training’.

2.3 Accreditation in England and Wales

In 2003, the CMC set up a voluntary accreditation scheme for the registration of organisations offering mediator training which must reach a minimum standard covering such issues as: ‘ethics; mediation theory; negotiation and role play’ (See for example: Brady, 2009; Brooker, 2007; Brooker, 2010b: 162-163). Specific to this enquiry is the requirement for accredited organisations to observe a Code of Conduct, which either ‘embraces or adopts’ the EU Mediator Code: 7(b). In 2009, the CMC attempted to introduce a new ‘self-certification scheme’ but elements of the proposal met with disapproval (See, Brooke, 2009 a, b) and the current scheme remains in place until further consultation (Brooke, 2010).

2.4 European initiatives

There are a number of initiatives in Europe which are relevant to or influencing mediation practice in England and Wales, which will be briefly reviewed (For detail see Allen, 2008; Brady, 2009; Brooker, 2010a: 162-163).

2.4.1 EU Mediator Code and self-determination

The EU introduced a Mediator Code in 2004 as a voluntary set of ‘broadly drawn principles’ which mediators and organisations can ‘commit’ to (Brady, 2009). In the preamble, the code recognises that mediation providers ‘may wish to develop more detailed codes which they have adapted to their specific context or the types of mediation services that they offer’. The text of the EU Code does not explicate self-determination or ‘informed consent’ but the mediator is required to ensure ‘that the parties...understand the characteristics of the mediation process and the role of the mediator and the parties in it’ (3.1) and further to ensure ‘all parties have adequate opportunities to be involved in the process’ (3.2). Brooker (2007: 225) notes that the EU Code is ‘silent’ on ‘mediator style’ but mediators are required to ‘conduct the mediation in an appropriate manner taking into account the circumstances of the case, including possible imbalances of power and any wishes the parties may express’ (3.1). The EU Code supports ‘market choice’ by stating that parties are ‘free to agree with the mediator, by reference to a set of rules or otherwise, on the manner in which the mediation is to be conducted’ (3.1) (See, Brady, 2009; Gould et Al, 2010: 11-13, 2009: 33).

2.4.2 EU Mediation Directive

An EU directive came into force in 2008 which requires each member state to put in place a ‘predictable legal framework’ (para 7) for mediation in cross border disputes (See Allen, 2008; Brady, 2009; Brooker, 2010a). Each member state should ‘encourage’ the ‘development and ‘adherence to voluntary codes of practice’; provide ‘effective quality control mechanisms’; and foster ‘initial and further training of mediators’ (article 4). Like the EU Code, the directive does not delineate the approach a mediator should use (Brooker, 2010a, b) nor does it highlight self-determination or informed consent but section 13 supports these principles by stating that ‘the parties are themselves in charge of the process and may organise it as they wish and terminate at any time’. Thus the directive supports a ‘market based approach’ by indicating that the ‘mechanisms should aim at preserving the flexibility of the mediation process and the autonomy of the parties’ (At 17) (See Brady, 2009; Gould et al, 2010).

2.4.3 IMI International Standards of Practice

The IMI has taken the initiative to introduce an international ‘mediator competency standard’ (Meyer and Leathes, 2008) and its mission is “To develop mediation into a free-standing profession over all fields worldwide based on high competency and ethical standards and to ensure transparency via Profiles including mandatory information and a summary of feedback by prior users’ (IMI, Website, 2010). The IMI Code of Conduct, which is available on line, requires mediators ‘to explain the process to the parties and their advisors, and be satisfied that they consent to the process used’ (at 3.2.1). There is a Professional Conduct Assessment process, which parties can activate if they believe there has been ‘a lack of compliance with the code’.

Accredited mediators add their profiles to a searchable database, which must include information about, their experience; style; code of conduct; complaints process; professional indemnity and ‘feedback digest’ (IMI Website, 2010).

2.5 Regulation versus ‘free market’ debate

Regulation and accreditation developments are in a state of fluctuation and the CMC proposal to ‘coordinate’ more formalised accreditation is in abeyance until further consultation has been undertaken (Brooke, 2010). There are leading voices resisting this move, for example, South (2009) reports that Dr Karl Mackie (Chief Executive of CEDR) is leading the ‘counter challenge’ by supporting a ‘free market’ for the ‘evolution of ideas and variety of training approaches’ (See also, Gould et Al, 2009, 2010; Boon et Al, 2007).

Just over half (52.8 per cent) of mediators responding to CEDR’s Mediator Audit (2010: 10) indicate approval for a ‘single standard of basic professional training’ and 54.9 per cent believe there should be ‘single body’ responsible for ‘setting and monitoring professional standards but there is a preference for an Association of Mediators over a regulatory body (CEDR Audit, 2010: 11). These figures represent little change from the attitudes reported in 2007 but a substantial change from 2005, when 61.7 per cent were in favour of a ‘single standard’ of training and over 70 per cent supported a ‘single’ body regulating activity (CEDR Audit 2010: 10-11. The report notes that previous audits had not provided ‘don’t know’ or ‘don’t care’ categories). The Audit reports that the CMC is now the frontrunner, both for training and for setting standards, but there is little enthusiasm for bodies outside England and Wales taking the lead (CEDR Audit, 2010).

Whether mediators operating in England and Wales will support uniformity or closer scrutiny of their practice is yet to be determined, although South (2009) maintains that there will be a substantial ‘lobby’ against ‘over restrictive’ regulation in preference for ‘self regulation’ with ‘rigorous Codes of Good Practice or Codes of Conduct’. This research is therefore a timely review of the Codes of Conduct observed by construction mediators in order to assess the level of conformity and the commitment to party self-determination and informed consent.

3. Methodology

The objective of the study is to determine the codes of conduct that construction mediators observe by investigating the information provided to consumers on the websites of mediation providers (See McCorkle, 2007³). Evidence from the CEDR Audit (2010: 4) indicates that many mediators are on multiple panels with respondents recording affiliation with over 50 organisations and half of CEDR mediators with ‘at least’ one other provider. As noted earlier the CMC has a voluntary accreditation system for organisations that train mediators and publishes a list of members online. In order to extend the scope of mediation provision, three lists were constructed.

1. First, organisations accredited by the CMC (50 members) and National Mediation Helpline (NMH) (45 members) were collated and cross-referenced. Some organisations on the finalised list offer only work place, family, children, community and, in one case, Sports mediation, which were removed leaving 41 providers involved with commercial and/or construction mediation (See Table 1).
2. The second list is of professional bodies and association, which have members working in construction, building and engineering but which are not on the CMC or the NMH registers. This includes the legal professional bodies (the Bar Council and Law Society); the Royal Institute of British Architects (RIBA); the Institute of Chartered Engineers (ICE) and two further legal associations for lawyers working in the TCC: Tecbar (The Technology and Construction Bar) and TeCSA (Technology and Construction Solicitors Association). All these organisations have accredited panels or provide lists of construction mediators on their websites (See Table 2).
3. The final list encompasses organisations with a long history of mediation provision, which are not accredited members of the CMC but which are documented in the mediation literature (See for example, Mistelis, 2006; the TCC Mediation Study (Gould et al, 2010) and CEDR’s Mediator Audit (2010) (See Table 3).

An inspection was made of the websites of each organisation between the 7th to the 25th June 2010 to review the information on their Codes of Conduct, which were available to the public. The aim of project was to examine the codes that construction mediators observe so additionally, where available, an investigation of the providers’ mediator panels or lists were reviewed in order to establish the numbers working in the construction field. There is a debate in the literature on whether mediators should have ‘subject-matter expertise’ (Riskin, 1996: 46) or whether ‘process expertise’ alone is more important for mediation success (See for example, Imperati, 1997: 702; Barush Bush, 1996). Talk Mediation’s website (2010) (See Table 1) states that their ‘specialist’ mediators ‘primarily use facilitative mediation’ which negates the necessity for ‘subject matter’ expertise because they deal with ‘real interests’ and which ‘are not limited to party rights’. This distinction has relevance to parties in construction disputes because a preference for evaluative mediators may lead to an inclination to choose ‘subject experts’ who come from either the legal professions or other construction related fields (See Brooker, 2007 at 232-233; Fullerton, 2005; Bickerman, 2005; Hart, 1992; Cheung et al, 2002).

³ McCorkle reviewed ‘impartiality’ and ‘neutrality’ in Codes of Conduct in USA courts.

A review of the organisations' websites from Table 1 indicate that they offer commercial mediation and many include construction as an example of one the areas they work in. Ten (10) organisations from Table 1 did not identify their mediators online and some did not have searchable profiles requiring users to make contact through email or phone enquires. In contrast, others publish which of their mediators specialise in construction or have construction experience. For example, UK Mediation advertises 'Re Build Construction', which has two specialist construction mediators. After searching the websites, it was not always possible to establish whether some providers have construction mediators or their level of experience as 'lead' mediators in this field. Further, some mediators may only indicate commercial mediation to distinguish this from other contexts such as family or community. Where organisations provide information on 'construction mediators' this is recorded, however, the numbers are only indicative of the extent of construction mediation provision.

As noted earlier, research indicates that the legal professions (Brooker, 2007, 2009, 2010a, b; Gould et Al, 2009, 2010; CEDR Audit, 2010) dominate construction mediation in the England and Wales. One objective of this study was to review the level of provision by lawyers. Therefore, each table also indicates when a mediation provider is primarily set up for lawyer-mediators and whether they are barristers, solicitors or a mixed composition.

4. Findings

4.1 CMC accredited providers' Code of Conducts (Table 1)

Fourteen (35 per cent) mediation organisations accredited by the CMC (accessible at the time of the review) explicitly refer to the EU Mediator Code on their websites in the information provided for potential users. For example, Clerksroom provide a downloadable brochure for interested users with a section on quality, which identifies that their mediators follow the EU Code. In addition to a statement about the Code, which the organisation follows, some also provide the text or a link to the EU Code. An investigation of other areas of the providers' websites for available documentation also found references to codes of conduct. For example, the Royal Institute of Chartered Surveyors' (RICS) information pack for accredited mediators; MTVM's application form for mediator accreditation; East Midland Dispute Resolution Centre's Policy Document for Mediators; and ADR Group, who have an ADR Net register of mediators, all state that their mediators either follow or must follow the EU Code. A further six providers made a general reference about following a code of practice. In total, the study found that sixty per cent of the organisations identify a code of conduct.

Table 1: CMC and NMN (National Mediation Helpline) accredited mediation organisations (Lists available online on 17th June 2010).

	Organisation	Lawyer ϕ (B) (S)	C M C	N M H	EU Code nf $\phi\phi$	Own Code	No. mediators ng $\phi\phi\phi$	Const. Mediators $\phi\phi\phi\phi$
1	The Academy of Experts		*	*	Nf		71	25 ⁴
2	ADR Group		*	*	5	*	ng	34

⁴ 25 profiles on the Academy of Expert's panel use the term 'construction'.

3	Association of Cambridge Mediators		*	*	Nf		20	8 ⁶
4	Association of Midlands Mediators		*	*	Nf		58 ⁷	9
5	Association of Northern Mediators		*	*	*		127 ⁸	35
6	Association South West Mediators		*	*	*		ng	
7	BL Resolve	L(S)	*	*	*		12	2 ⁹
8	CEDR		*	*		*	150+ 29 ¹⁰	15
9	Civitas Law, Civil & Public Law Barristers	L(B)	*	*	Nf		10	2
10	Clerksroom Mediation	L(M)	*		*		495	24
11	Commercial and Medical Disputes Resolutions (CMDRS)		*	*	Nf		6	2
12	Commercial Mediation West Wales ¹¹		*	*			6	
13	Devon & Somerset Law Society	L(S)	*	*	Nf		13	3
14	Disputes Resolved		*	*	Nf		150 ¹²	
15	The Dispute Mediation Consultancy LLP		*	*	*		15	
16	East Midlands Mediation Centre Ltd	L(B)	*		13		7 ¹⁴	
17	Effective Dispute Solutions Ltd (EDSL)		*	*	Nf		19 ¹⁵	5 ¹⁶
18	Eternal Alliances		*	*	*		8	
19	Focus Mediation Ltd ¹⁷		*	*	Nf		10	
20	Garden Court Mediation	L(B)	*	*	Nf		20	
21	Global Mediation Limited		*	*		* ¹⁸	ng	
22	Gray's Mediation Services LLP (RICS)		*	*	*		13	7
23	IDRS Ltd ¹⁹		*	*	Nf		ng	
24	The Institute of Chartered Accountants in England & Wales		*		Nf		20	3
25	InterMediation (part of InterResolve) ²⁰		*	*	*		ng	
26	Keep Talking		*	*		* ²¹	9	5
27	LADR (LambBuilding Alternative Dispute Resolution)	L(B)	*	*	Nf		19	

⁵ ADR Group has a panel of registered mediators called ADR Net and state their mediators must follow the EU Code.

⁶ Two mediators from the Association of Cambridge Mediators indicate wide experience in commercial disputes in 'any' or 'nearly' any area.

⁷ The Association of Midlands Mediators has over 60 mediators but only 58 profiles were on their website. 29 profiles were online at the time of the search was undertaken.

⁸ Association of Northern Mediators did not have all their mediator profiles online at the time of the search.

⁹ One profile on BL records experience of construction disputes but this is not down as specialism.

¹⁰ CEDR state they have over 150 mediators but they also have a 'Direct Referral Panel' of 29.

¹¹ Commercial Mediation West Wales' website was unattainable at the time of the search. A telephone call on 18th June provided information that it is a small group of six mediators some with a construction background. They adopt the EU Code of Practice.

¹² Clerksroom Mediation gave the number of mediators only.

¹³ A reference to the EU Code was in a downloadable policy for KCH mediators.

¹⁴ One profile for East Midlands Mediation Centre indicates the mediator is 'able to take on the role of any type of civil dispute'.

¹⁵ Effective Dispute Solutions was updating its website at the time of the search but has 36 mediators on the panel (21st June 2010). 17 profiles were accessible.

¹⁶ The profiles on the Effective Dispute Solutions' website indicate commercial mediation with some experience in construction or building disputes.

¹⁷ Focus's website states that it provides construction mediation.

¹⁸ Global Mediation refers to a strict code but no further information was obtainable online.

¹⁹ Division of the Institute of Chartered Arbitrators

²⁰ InterMediation mainly works in insurance claims.

²¹ Keep Talking state that they have adopted the CMC Code of Conduct.

28	<i>Landmark Chambers (Property specialists)</i>	<i>L(B)</i>	*	*	<i>Nf</i>		17	
29	<i>Law South Mediation</i>	<i>L(S)</i>	*	*	*		9 ²²	2
30	<i>Littleton Dispute Resolution Services Limited</i>	<i>L(B)</i>	*	*	<i>Nf</i>		23	7
31	<i>Lyons Davidson Dispute Resolution</i>	<i>L(S)</i>	*	*	<i>Nf</i>		9	1
32	<i>Mediation-1st</i>		*	*	*		4	1 ²³
33	<i>Mediation Solve</i>		*	*	*		5	2
34	<i>MTVM (Middlesex & Thames Valley Mediators)</i>		*	*	²⁴		ng	
35	<i>Northern Dispute Resolutions</i>		*	*	<i>Nf</i>		ng	
36	<i>Oxford Mediation</i>		*	*	*		21	13
37	<i>Royal Institution of Chartered Surveyors</i>		*	*	²⁵		ng	
38	<i>Solent & Wessex Civil Mediation</i>		*	*	*		14	3 ²⁶
39	<i>Specialist Mediators LLP</i>		*	*	<i>Nf</i>		22	2
40	<i>Talk Mediation Ltd</i>		*	*	*	* ²⁷	ng	
41	<i>UK Mediation</i>		*	*		* ²⁸	ng	2 ²⁹

- ϕ Composition L = Legal professionals; (S) = solicitors firm; (B) barristers chambers or group.
- $\phi\phi$ nf = not found.
- $\phi\phi\phi$ ng = not given.
- $\phi\phi\phi\phi$ Number of construction/building mediators where information provided on the website.

4.2 Legal and professional bodies and associations' Code of Conducts (Table 2)

Table 2 provides information from professional bodies which have panels of commercial and construction mediators but which are not accredited by CMC or the NMH. (The RICS is a recognised provider with the CMC and is in Table 1.) The Law Society regulates the practice of solicitors in England and Wales. It has a commercial and civil mediation accreditation panel at two levels: General Members (two years duration prior to application as a practitioner) and Practitioner Mediators. Members must comply with the Law Society's Code of Practice for civil and commercial mediators and must have completed an approved training course. Solicitors who do not register for the scheme are not prevented from acting as mediators and can practise 'direct(ly)' or be on other organisations' panels. However 'scheme members' who engage in 'direct' mediation are required to provide each party with 'the codes of practice or conduct' that they 'adhere to' (Law Society Scheme, 2009: 8).

A search of English firms listed in the civil and commercial mediation specialism group found 1223 in this category but only 38 with accredited mediator practitioners on the Law Society panel. This decreases to 28 firms specialising in construction and engineering. A review of this list shows multiple entries for firms with regional offices and the same mediator profiles in each entry. Only two profiles specify construction and engineering as a specialism but solicitors may define their practice more broadly as commercial litigators or commercial work. A review of the Legal 500 Series Mediator List, which ranks law firms and lawyers nationally and

²² Law South Mediators record 6 mediators and 3 associate mediators

²³ Mediation 1st had one profile indicating 'experience' with construction disputes.

²⁴ MTVM's application form states that panel members must adopt the EU Code.

²⁵ Reference to a code was in the RICS' Mediator Information Pack.

²⁶ One profile from Solent & Wessex Mediation records experience with one construction case.

²⁷ Talk Mediation's website indicates that it mediates in a wide number of areas. They have a panel but it is not online and contact is by email or telephone. The website states that their mediators are all 'Obligated to adhere to our very strict Code of conduct, incorporating the EU Code of Conduct'.

²⁸ UK Mediation states that mediators 'all work to our codes of ethic and practice'.

²⁹ UK Mediation advertises 'Re Build Construction' with two specialist construction mediators.

internationally (Legal 500 Website, 2010), found two solicitors named as construction specialists but neither appear on the Law Society's practitioner mediator list.

A second search of the Law Society panel found 30 Welsh firms specialising in civil and construction mediation but only one firm with an accredited mediator practitioner who was not a construction specialist. These finding may indicate that many solicitors do not seek to advertise their work as mediators through the Law Society practitioner scheme but prefer to gain appointments through accreditation with other panels or from personal recommendations (See Gould et Al, 2010). Solicitors registering civil and commercial mediation as a specialism with the Law Society may be offering this to clients in an advisory or representative capacity.

Table 2: Professional organisations' (not on CMC list) Codes of Mediator Conduct

<i>Organisation</i>	<i>Lawyer (B) (S)</i>	<i>EU Code</i>	<i>Own Code</i>	<i>No. Mediators</i>	<i>No. Constr. mediators</i>
<i>Bar Council</i>	<i>L(B)</i>			<i>Total number mediators not available</i>	<i>128</i>
<i>Law Society (England search)</i>	<i>L(S)</i>		<i>*</i>	<i>17 practitioner mediators</i>	<i>2</i>
<i>Law Society (Wales search)</i>	<i>L(S)</i>		<i>*</i>	<i>1 practitioner mediators</i>	
<i>Tecbar</i>	<i>L(B)</i>	<i>nf</i> ^Φ			<i>44</i>
<i>TeCSA</i>	<i>L(S)</i>	<i>nf</i>	<i>*</i>		<i>34</i>
<i>ICE</i>		<i>nf</i>			<i>22</i> ³⁰
<i>RIBA</i>		<i>nf</i>		<i>Not given</i>	

Φ nf = not found

The Bar Council recognises the increase in 'barristers training as mediators' and that their 'skills' make them particularly apt mediators but some of these skills, for example, 'evaluating strengths and weaknesses' and 'advising on likely outcome' (Bar Council Website, 2010) are evaluative in nature. Barristers can register as mediators with the Bar Council if they have completed a training course which is 'at least 25 hours in length and includes teaching in mediation procedure and ethics and must include role playing sessions' (Bar Council Website, 2010). The registration form identifies several training organisation found in Table 1 but some are specific to family mediation or other specialist fields and the School of Psychotherapy and Counselling does provide a mediator panel. The remaining provider on the form is ADR Chambers (UK) Ltd, which is in Table 3. Applicants can also indicate other training organisations, which are not on the form but these must reach the standards set by the Bar Council.

The Bar Council's mediator panel can be searched either by looking for an individual name or by searching under a specialist area and/or the number of mediations undertaken as a 'lead mediator' (See Table 4). 126 barrister mediators are registered as 'lead' mediators specialising in construction (22-23 June 2010), although many list more than one specialism. Nearly 40 per cent have mediated in fewer than five mediations although over a quarter reported experience in over 30 mediations (Table 4). The Bar Council 'registration form' does not specify required information on the code of conduct observed.

Table 4: Barristers registered with the Bar Council specialising in construction mediation

³⁰ The ICE mediation panel only allows searches for mediators and conciliators together.

Number of mediation	Numbers of mediators
30+	33
15-29	18
5-14	31
Under 5	44
Total	126

TeCSA is an association of solicitors who work in the fields of ‘construction, engineering technology and/or IT’ (TeCSA Website, 2010). It’s ‘objective is to encourage higher standards and greater efficiency in the conduct of ADR, arbitration and litigation in the TCC’. TeCSA ‘maintains’ a list of ‘specialist mediators’ and 34 profiles were on their website at the time of the search. The association also provides a ‘nominating service’. The Association’s ‘Mediation Rules’ are published online which state that the mediator ‘will abide by the terms of the Model Procedure and any relevant code of conduct’ (TeSCA, 2010).

Tecbar represents barristers ‘practicing in the TCC’ (Tecbar Website, 2010) and membership rests on reaching a set criteria. One of the stated objectives of the association is ‘to accredit and/or nominate suitably qualified members as arbitrators or adjudicators or mediators’ (f). Application for membership is through a downloadable form and the website provides a list of 44 members who are mediators, however, an inspection of the website did not find a Mediators’ Code of Conduct online.

The legal profession have gone a long way from the early allegations that they lacked knowledge about mediation or that they were ‘hindering’ its progress because they feared losing their dominance in the dispute resolution market (See for example; Genn, 1998; Brooker and Lavers, 2000). Lawyers are now reported to be positioning themselves either as mediators or within the process as advisors representing their clients’ cases (Gould et al, 2009; Brooker, 2007; 2010a, b), which is said to explain the predominance and preference for evaluative mediators in construction (Brooker, 2007). In total, an examination of the websites of 51 organisations indicates that a third of these represent the legal professions either as barristers, solicitors or as a combination of both. It is also evident from this investigation that other providers have a substantial number of legally qualified mediators on their panels.

Members of the construction industry may also access mediation information from other professional bodies for example ICE has a panel of mediators who have undertaken ‘rigorous accreditation’ (ICE Website, 2010) and a search provides profiles for 22 conciliators/mediators although it does not allow for ‘mediator only’ searches. ICE distinguishes the processes of mediation and conciliation in an ICE contract where the conciliator provides a recommendation in the event of a failure to reach a settlement. A search of the website did not provide further information on a Code of Conduct.

RIBA has a panel of mediators and its online Mediation Scheme describes a process where the ‘overall intention is to reach an agreed solution’ or if ‘impossible’ the mediator will attempt to ‘narrow the differences’ and will at the parties’ request make a recommendation (2). RIBA’s website provides online information on mediation guidance and a copy of its ‘Mediation Scheme’ but neither document referred to a Code of Conduct.

4.3 Other highly regarded mediation providers' codes of conduct (Table 3)

Table 3 reports on a number of providers with highly regarded reputations in England and Wales. The TCC mediation survey found that mediators were appointed from CEDR (5 cases), In Place of Strife (3 cases), ADR Group (3 cases), Consensus (1), Independent Mediators Ltd (1), Association of Midlands Mediators (1) and from three unnamed sets of barristers' chambers. Of these, ADR Chambers; In Place of Strife; Independent Mediators and Consensus do not appear in the earlier tables and therefore are in Table 3.

Table 3: Leading Mediator Providers

Organisation	Lawyers (L)	EU Code	Own Code	No. mediators	Construction mediators
ADR Chambers	L	*		99	39
In Place of Strife		*		28 ³¹ (23 May 2010)	18
Independent Mediators Ltd	L	Not found		6	5 ³²
Consensus		*		35 (22)φ	11φφ

φ Consensus identifies a list 22 'principal' mediators who have wide mediation experience.

φφ 20 Consensus mediators indicate experience in the construction field. Consensus also has a specialist construction panel with 11 mediators.

ADR Chambers withdrew from the CMC Board of Members in April 2010 (CMC Website, 2010). It was set up in 1999 to support the civil justice reforms formulated by Lord Woolf and has a 'judicial panel of Law Lords, Lords of Justices, High Court, Circuit and other retired judges' (ADR Chambers Website, 2010). The Bar Council recognises ADR Chambers' mediator training programme for registration. ADR Chambers has an expert panel, a specialist panel, an international panel and mediator panel for practicing barristers and solicitors. A search of the mediator panel returned 99 profiles that narrowed to 39 mediators using the term construction. Two further profiles indicate experience with building (1) and (marine) engineering (1). ADR Chambers states on its website that it 'requires all members to adhere to the EU Code of Conduct' and provides a link to this document.

Five mediators from Independent Mediators are in the top tier (13 names) of the Legal 500 Series 'Mediators List' (Legal 500 Website, 2010). A review of the website did not find reference to a Code of Conduct. Eight mediators from In Place of Strife are in the top five tiers of the Legal 500 Mediators Series. In Place of Strife has a Quality Assurance section on its website and has 'adapted' the EU Code and its text is provided online. Finally, Consensus provides the text of the EU Code and has a panel of experts in engineering, building and construction with 11 mediator profiles on its website.

4.4 Mediation agreements and rules of procedure

The public seeking information on mediation can also sometimes access other sources such as the organisation's Mediation Agreement or Mediation Rules if published online. An inspection of these available documents found that they covered such issues as fees; confidentiality, impartiality, commencement and termination of mediation but some documents provide details on the conduct of the mediation

³¹ In Place of Strife, list 20 mediators plus a further 8 recommended mediators (23rd May 2010).

³² All mediators from Independent Mediators indicate extensive experience in wide areas of commercial disputes.

process and the role of the mediator and the parties. For example, ADR Group's 'Mediation Procedure and Rules' state that the 'mediator will: determine the procedure' (at 3.1), however at 3.2 the mediator will not: analyse legal positions, impose settlement or offer legal advice to the parties'. CEDR's 'Model Procedure for Mediation' (2010) states that the mediator 'will take responsibility for determining the procedure in consultation with the parties' (Section: The Mediation) as does Independent Mediators' 'Mediation Agreement' (at 5.2). Law South Mediation's agreement states that the mediator 'will (b) after consultation with the Parties, determine the procedure to be followed at the mediation' (at 10). The Association of Northern Mediators' 'Supplementary Information Document (no 12 at 4.1)' states that the mediator 'may conduct the mediation in any manner he or she considers appropriate' and 'will take particular account of ... (c) the parties' wishes'. Landmark Chambers' Mediation Agreement states that the mediator will 'chair and determine the procedure for, the Mediation' (3.1).

Other agreements might imply that the mediator has more 'control' over the determination of the process than the parties do and they sometimes show a high degree of legalistic language. For example, the IRDS' Mediation Agreement' provides that before the mediation begins 'the mediator will tell all the parties how he or she will conduct the session and will ask the parties to sign a mediation agreement to confirm that they understand the process and their responsibilities within the mediation' (2 h). The DMC 'Agreement' indicates that the mediator shall determine in consultation with the parties the procedure (at para.1) and at paragraph 3 the 'mediator may in his or her absolute discretion give such directions for the conduct of the Mediation as he or she thinks fit'. Later in the same document the 'Rules of Procedure' are given and under the 'Mediator's Position' it states, 'The ultimate right to determine the procedure at the Hearing rests with the Mediator' (1). Mediation 1st states that the mediator will 'determine the procedure at and for the Mediation to include the submission of any documents to the mediator or between the Parties prior to the Mediation' (3.1) and later in the document that the 'Parties will comply with any directions made by the Mediator under paragraph 3 of this Agreement as to the conduct of the Mediation or as to the submission of any documents to the Mediator or to any other Party prior to the Mediation' (7).

4.5 Discussion

As this review shows not all mediator providers, raise awareness of the EU Mediator Code or put their own codes on their websites. At the time of the web search forty per cent of providers accredited by the CMC (Table 1) did not refer to a Code of Conduct or this was not easily accessible. In some cases, a diligent search in 'Mediation Agreements', 'Rules of Procedures' or in information for mediators who want to gain accreditation with the organisation found references to codes of conduct. Although the EU Code recognises that providers may 'develop' or 'adapt' the Code to reflect the 'context' of provision this study indicates that organisations have generally not exercised this discretion. This does not imply that mediators or providers ignore the Codes that they work under and they may issue other documents to users before they mediate or raise the matter in discussions prior to engagement.

Very few mediators are able to make a full time living mediating in England and Wales (CEDR Audit, 2010) but if they aspire to gain professional status the 'first step' might be 'bringing inspirational, higher purpose language into their codes of

conduct' (Alfini, 2009: 837; See also Oberman, 2008). The EU Code and the Mediation Directive do not place emphasis on the principle of self-determination in contrast to Codes from other jurisdictions. For example; in the preparatory note to the Uniform Mediation Act (2001) in the USA the importance of informed self-determination is highlighted as part of the 'drafting philosophy' (note 6) not only because the principle 'guarantees fairness but also because it contributes to party satisfaction'. The Model Code of Practice for Mediators (2005) formulated by the American Arbitration Association, the American Bar Association and the Association for Conflict Resolution purposely identifies that 'the mediator shall conduct a mediation based on the principle of party self-determination' which not only 'includes the outcome' but also 'free and informed choice as to process' (STANDARD 1: Self-Determination at A).

The CMC produced a new Code of Conduct in 2009, which stated 'the mediator shall conduct the mediation under the principle of party self-determination' but the code went further by adding 'by facilitation' (2). This was objected to in the consultation process on the grounds that it was 'too prescriptive', ignored other models of mediation including evaluative mediation and Early Neutral Evaluation and did not give credence to either the 'flexibility of mediation' or to the parties' informed choice about the role of the mediator (Brooke, 2009b: 52-53). Suggestions made for amendments denote that highlighting self-determination may be acceptable by removing 'by facilitation' or adding 'otherwise agreed'. (Brooke, 2009 a: 53). A clear statement indicating that process selection is through party self-determination would be a simple measure to take without compromising the 'free market' approach favoured by commercial mediators (South, 2009; Gould, 2010; Brady, 2009) which would allow mediators to fashion an appropriate process without compromising on informed consent.

The EU Code at 3.1 requires the mediator to 'conduct the mediation in an appropriate manner, taking into account the circumstances of the case, including possible imbalances of power and *any wishes the parties may express*, (italics for emphasis)'. Further the Code specifies that 'the parties may agree with the mediator on the manner in which the mediation is to be conducted' but the 'mediator must ensure that the parties to the mediation understand the characteristics of the mediation process and the role of the mediator and the parties in it' (at 3.1). Mediation Agreements and Codes of Procedure, which indicate that mediators will 'determine' the procedure in consultation with the parties, tacitly acknowledge party self-determination. However, this is somewhat short of the obligation placed on nationally accredited mediators in Australia who are required to undertake steps to make certain that the parties enter mediation with 'informed consent' about the procedure that they are going to engage in (Alexander, 2008a, b; Sourdin, 2007). Under the Australian National Mediator Standards (2007) 'mediators will ensure' that parties are given 'an explanation of the process and have had an opportunity to reach agreement about the way in which the process is to be conducted' (2).

Coolie (2000: 77-78) believes that mediators have an 'ethical duty' to 'educate the parties on the mediation process and the role and responsibilities of the mediator in that process' otherwise he maintains the parties do not 'meaningfully consent to engage in it':

‘If one party believes that the mediation will be totally facilitative and the other party believes that the mediation will have an evaluative element, then their minds have not met on, nor have they consented to engage in a particular process’.

There are actions that mediators can take to make certain that consumers are cognisant of the process prior to mediating which, if put into practice, have the advantage of endowing ‘clarity’ on informed consent (See Imperati, 1997 at 705). For example, Imperati (1997: 705) suggests that mediators should ‘prepare statements’ on the ‘approach or approaches’ that they adopt with illustrations ‘of what he or she does, or more importantly, does not do’ in mediation and further suggests these could be disseminated through ‘letters’, ‘mediator’s websites’, ‘Mediation Agreements’, or discussion prior to mediating.

An example of ‘informative’ practice is the IMI published Code of Ethical Conduct (IMI Website, 2010) which expects mediators to ‘make reasonable efforts to determine that the process used is explained to the parties and their counsel and that they consent to the process being used and to the mediator selected’ (4.2(a)). Further, mediators in their online profile have to include information not only on the Code of Practice they observe but also on their approach in mediation. IMI’s panel allows users to search for mediators with specialist subject expertise and to select the style they prefer. This includes, ‘any style; evaluative; facilitative or transformative’. A search for UK mediators returned 33 profiles of which 18 have construction experience. Of these, 17 mediators specify that they adopt a facilitative style, 10 an evaluative style and 5 a transformative approach (Date of search 1st July 2010), which indicates that some mediators offer multiple approaches to mediation.

This review reveals variable practices on how providers impart information about mediation and mediator approaches on their websites. Many organisations provide an explanation of the process in a ‘question and answer section’ where they inform potential users that mediation is different to litigation because they are in ‘control of the process’ and ‘the outcome’. Some providers give significant emphasis to quality assurance and highlight that their mediators observe a Code of Conduct and there are examples of organisations making the full text of the EU Code readily available to the public. However, the EU Code itself does not expressly refer to the principle of self-determination, although the language used alludes to a process of consultation and agreement between the parties and mediators prior to engagement. Codes of Conduct from other countries or international organisations are found to place specific requirements on their mediators to ‘educate’ or ‘inform’ the parties about the process and their style before mediating which gives weight to the principle of party self-determination and supports informed consent. Adopting these practices in England and Wales would provide transparency about mediators’ styles, would raise the profile of self-determination as the ‘guiding principle’ (Alfini, 2008: 831) by which mediators practise this newly developing profession and would promote a ‘reference point’ (Mason, 2009: 202) for good practice.

5. Conclusion

In England and Wales, access to court is now constrained by court rules and litigants face costs sanctions for unreasonably rejecting mediation, which has paved the way for mediators to develop careers and move towards the recognition of mediation as a

profession, but it also raises the bar for mediators to assure the quality and standards of their practice. This study shows a lack of conformity on how mediation providers publish their Codes of Conduct and self-determination is not given prominence either in the EU Mediator Code or other codes available online. The majority of providers recognise the EU Code, which is required for accreditation with the CMC but few have gone beyond a perfunctory statement to this effect. The websites rarely acknowledged self-determination as the primary principle governing the practice of mediation although it is noted that their 'Mediation Agreements' usually advise that the mediator shall determine the procedure in consultation with the parties. If party self-determination sets mediation apart from other dispute resolution mechanisms, then the Codes of Conduct observed by mediators should give prominence to this principle in order that 'informed consent' forms the basis of engagement.

Codes of Conduct from other jurisdictions provide exemplars of good practice which could assist with illuminating the importance of self-determination by presenting clear guidelines of the steps mediators should taken to ensure that parties enter mediation with informed consent. A further measure would be to assign mediators an 'ethical duty' (Coolie 2000) to educate consumers about their approach through marketing information (Imperati, 1997), which would aid transparency, ensure parties enter mediation with informed consent and would go someway to 'enshrining' party self-determination as the cornerstone of construction mediation without stultifying mediators' style or innovation.

Case List

Dunnett v Railtrack PLC (2002) C.P. Rep. 35.

Halsey v Milton Keynes General NHS Trust [2004] EWCA Civ 576

Statutory References

Civil Procedure Rules 1996 (UK)

Uniform Mediation Act 2001 (US)

References

American Arbitration Association, the American Bar Association and the Association for Conflict Resolution (2005) *The Model Code of Practice for Mediators* <http://www.adr.org/si.asp?id=3827> (Accessed on 9th July, 2010)

ADR Chambers' Website, (2010) <http://www.adrchambers.co.uk/> (Accessed on 23 June, 2010)

ADR Group Website, 2010 <http://www.adrgroup.co.uk/adrnet-howtojoin.asp> (Accessed on 9th July, 2010)

Alfini J., (2008) Ethics in the Expanding World of ADR: Considerations, Conundrums, and Conflict *S. Tex. L. Rev.*, 49, 829-839

Alexander N., (2006) *Global Trends in Mediation* (2nd Edition) Kluwer International, BV

- Alexander N., (2008a) Mediation Metamodels: Understanding Practice *Conflict Resolution Quarterly*, 26, 1, 197-118
- Alexander N., (2008b) Mediation and the Art of Regulation, 8 *Queensland U. Tech. L. & Just. J.* 1, 1-23
- Allen, T., (2008) Implementing the EU Directive on mediation
http://www.cedr.com/index.php?location=/library/articles/20081006_244.htm
 (Accessed on 13th February, 2009)
- Association of Northern Mediators' Supplementary Information (2010)
[http://www.northernmediators.co.uk/Downloads/Supplemental%20conditions%20to%20agreement%20to%20mediate%20\(11\).pdf](http://www.northernmediators.co.uk/Downloads/Supplemental%20conditions%20to%20agreement%20to%20mediate%20(11).pdf) (Assessed on the 25 June, 2010)
- Australian National Mediator Standards (2007)
<http://www.leadr.com.au/documents/Practice%20standards.pdf> (Accessed on the 11th May, 2010)
- Bar Council (2010)
<http://www.barcouncil.org.uk/about/find-a-barrister/BarristerMediators/>.
<http://www.barstandardsboard.rroom.net/standardsandguidance/codeofconduct/>
 (Accessed on the 25 June, 2010 and 1st July, 2010)
- Baruch Bush R., and Pope S., (2002) Changing the Quality of Conflict Interaction: The Principles and Practice of Transformative Mediation *Pepperdine Dispute Resolution Journal*, 3, 67-96
- Baruch Bush R., (2002) Substituting Mediation for Arbitration: The Growing Market for Evaluative Mediation and What it Means for the ADR Field *Pepperdine Dispute Resolution Law Journal* 3, 111-131
- Boon A., et Al (2007) Regulating mediators? *Legal Ethics*, 10 (1), 26-50
- Bowie J., (2004) Ethical Issues in Construction Mediation: Are there any Rules? *Construction Law* 24, Spring, 33-39
- Brad Reich J., (2002) Attorney v. Client: Creating a Mechanism to Address Competing Process, *Southern Illinois University Law Journal*, 26, 183-222
- Brady A., (2003) *Alternative Dispute Resolution (Mediation) Development for Non-Family Civil Disputes in England and Wales*, paper presented at the World Jurist Association 21st Biennial Conference on Law of the World, Sydney/Adelaide: WJA Publication, August, 17-23
- Brady A., (2009) Mediation Developments in civil and commercial matters within the European Union, *Arbitration*, 73(3), 390-399
- Brooke H., (Sir) (2010) *Mediation in the UK today: An authoritative review of the UK mediation scene today from the CMC's perspective* CMC Academic Seminar, 20th January 2010, London

<http://www.civilmediation.org/files/pdf/Mediation%20in%20the%20UK%20today%20by%20Sir%20Henry%20Brooke%20Jan%202010.pdf> (Accessed on the 6 July 2010)

Brooke H., (2009a) *Chairman's Report on Consultation CMC*
<http://www.civilmediation.org/files/pdf/Responses%20to%20the%20Consultation%20Process%20final.pdf> (Accessed on the 6th July, 2010)

Brooke H., (2009b) *Responses to the Consultation Report*
<http://www.civilmediation.org/files/pdf/Responses%20to%20the%20Consultation%20Process%20final.pdf> (Accessed on the 6th July, 2010)

Brooker P., (2010a) Judging Unreasonable Behaviour at the Interface of Mediation in the English Jurisdiction *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, Vol. 2, No 3, 148 (August publication)

Brooker P., (2010b) *Construction Mediation in England and Wales* in Brooker P and Wilkinson S., (2010) *Construction Mediation: an international review*, Spon Press an imprint of Taylor and Francis, London and New York

Brooker P., and Wilkinson S., (2010) *Construction Mediation: an international review*, Spon Press an imprint of Taylor and Francis, London and New York

Brooker P., (2009) Criteria for the Appropriate Use of Mediation in Construction Disputes: Judicial Statements in the English Technology and Construction Court *International Journal of Law in the Built Environment*, Vol. 1, Issue 1, 82-97

Brooker P., (2007) An Investigation of Evaluative and Facilitative Approaches to Construction Mediation, *Structural Survey*, Vol. 25, Number 3/4, 220-238

Brooker P., and Lavers A., (2000) Issues in the Development of ADR for Commercial and Construction Disputes *Civil Justice Quarterly*, 19 (Oct), 353-370

Brooker P., and Lavers A., (2005a) Construction Lawyers' Experience with Mediation Post-CPR, *Construction Law Review*, No 1, 19-46

Brooker P., and Lavers A., (2005b) Mediation Outcomes: Lawyers' Experience with Mediation, *Pepperdine Dispute Resolution Journal*, Vol. 2, 161-213

Brunsdon-Tully M., (2009) There is an ADR but does anyone know what it means any more? *Civil Justice Quarterly*, 28(2), 218-237

Bucklow A., (2007) The 'Everywhen' Mediator: The Virtues of Inconsistency and Paradox: The Strength, Skills, Attributes and Behaviours of Excellent and Effective Mediators, *International Journal of Arbitration, Mediation and Dispute Management* 73, 1, 40-48

CEDR (2010) *Model Procedure for Mediation* 11th Edition
http://www.cedr.com/library/documents/MMP_11thEd.doc (Accessed on 5th July, 2010)

CEDR *Code of Conduct for third party neutrals* (Revised 2008)

http://www.cedr.com/library/documents/CodeOfConduct_revised_2008.pdf
(Accessed on the 28 May, 2010)

CEDR's *Third Mediator Audit* (2007)

<http://www.cedr.com/gfx/TheMediationAudit2007.pdf> (Accessed 3rd March, 2009)

CEDR's *Fourth Mediator Audit* (2010)

<http://www.cedr.com/docslib/CEDRMediatorAudit2010.pdf>
(Accessed on 1st June, 2010)

Cheung S., Suen H., and Lam T., (2002) Fundamentals of Alternative Dispute Resolution processes in Construction", *Journal of Construction Engineering and Management*, September/October, 409-417

Clark B., (2009) Mediation and Scottish Lawyers: past, present and future, *Edin. L.R.* 2009, 13(2), 252-277

Clark B., and Mays R., (1996) Regulating ADR – The Scottish Experience, 5 *Web JCLI* <http://webjcli.ncl.ac.uk/1996/issue5/clark5.html> (Accessed on 26 July, 2010)

CMC 2009 *CMC Interim Report on Consultation Process*

<http://www.civilmediation.org/files/pdf/INTERIM%20REPORT%20ON%20THE%20CMC%20QUESTIONNAIRE.pdf> (Accessed on the 13 June, 2010.)

CMC *Code of Conduct for Mediators* (2009)

<http://www.civilmediation.org/files/pdf/The%20CMC%20Code%20of%20Good%20Practice%20for%20Mediators%202009.pdf> (Accessed on 10 May, 2010)

Coben J., & Thompson N., (2006) Disputing Irony: A Systematic Look at Litigation about Mediation *Harvard Negotiation Law Review*, 11, 43-145

Consensus Website, 2010 <http://www.consensusmediation.co.uk/mediation.html>
(Accessed on 8th July, 2010)

Cooley J., (2000) Mediators and Advocates Ethics *Dis. Resol. J.*, 55, 73-79

Cooley J., and Love L., (2008) Midstream Mediator Evaluations and Informed Consent *Dispute Resolution Mag.*, 14, 11-14

DMC Website *Mediation Agreement Processes* (2010) Accessed 16 June, 2010

http://www.dispute-mediation.co.uk/the_process.aspx (Accessed on 16 June, 2010)

Dolder C., (2004) The Contribution of Mediation to Workplace Justice *Industrial Law Journal*, 33, No 4, 420-442

East Midlands Mediation Centre Website (2010) *Mediation Policy*

<http://www.eastmidlandsmediationcentre.co.uk/EMMC-POLICY.pdf> (Accessed on 16th June, 2010)

Folberg J., and Taylor A., (1984) *Mediation* Jossey (Bass: San Francisco)

Gaitskell R., (2005a) Current Trends in Dispute Resolution. Society of Construction Arbitrators' Annual Conference, May 2005,
<http://constructionarbitrators.org/pdfs/DrGaitskellstalk1.pdf> (Accessed 19th February 2009)

Gaitskell R., (2005b) Adjudication: Its Effect on other Forms of Dispute Resolution (the UK experience)
http://www.keatingchambers.co.uk/resources/publications/2005/rg_kl_adjudication_its_effect.aspx (Accessed on the 13th February, 2009)=

Genn H., (1998) *Central London County Court Pilot Mediation Scheme: Evaluation Report*, Lord Chancellor's Department, Research Series No 5/98. July

Genn H., (2007) "Twisting Arms: Court Referred and Court Linked Mediation under Judicial Pressure" <http://www.justice.gov.uk/publications/research210507.htm>
 (Accessed on 25th February, 2008)

Gould et Al., (2010) *Mediating Construction Disputes: An Evaluation of Existing Practice*. Kings College London, Centre of Construction Law and Dispute Resolution (London)

Gould N., et Al (2009) The Use of Mediation in Construction Disputes: Summary Report of the Final Results May 2009 (Kings College London) 2009 at
<http://www.mondaq.com/article.asp?articleid=79894> (Accessed on 20th May, 2010)

Gould N., (2007) *Mediation in Construction Disputes: An Interim Report* at
http://www.fenwick-elliott.co.uk/articles/ADR/mediation_interim_report.htm
 (Accessed on 2nd November, 2007)

Grossman A., (2003) Professionalising Commercial Mediation: Discarding the Baggage of Idealised Professions *Mediate.com*
<http://www.mediate.com/articles/grossmanA.cfm#bio> (Accessed on the 6th May, 2010)

Hadeen T., (2005) Coercion and Self-Determination in Court-Connected Mediation: All Mediations are Voluntary, but some are more voluntary than others *Justice System Journal Special Issue: Making Dispute Resolution Work*, 26, 273-286

In Place of Strife Website (2010)
http://www.mediate.co.uk/quality_assurance/code.html (Accessed on the 10 May, 2010)

IMI Code of Ethics <http://www.imimediation.org/imi-code-of-professional-conduct>
 (Accessed on the 9th July, 2010)

Hensler N., (2001) "Making Deals in Court Connected Mediation: What's Justice Got to do with it?" *Washington University Law Quarterly*, 79, 787-861

Hensler N., (2002) Suppose It's Not True: Challenging Mediation Ideology, *Journal of Dispute Resolution*, (Volume 2002), 81-99

Hensler N., (2003) Our Courts, Ourselves: How the Alternative Dispute Resolution Movement Is Re-shaping Our Legal System, *Penn State Law Review*, Summer, 108, 165-197

<http://www.civilmediation.org/files/pdf/INTERIM%20REPORT%20ON%20THE%20CMC%20QUESTIONNAIRE.pdf> (Accessed on the 13 June, 2010.)

<http://www.civilmediation.org/files/pdf/Responses%20to%20the%20Consultation%20Process%20final.pdf> (Accessed on the 9th July, 2010)

<http://www.civilmediation.org/files/pdf/Responses%20to%20the%20Consultation%20Process%20final.pdf> (Accessed on the 9th July, 2010)

ICE Website, 2010 <http://disputeresolution.ice.org.uk/Search.aspx>. (Accessed on 25th June, 2010)

Imperati S., (1997) Mediator Practice Models: The Intersection of Ethics and Stylistic Practices in Mediation *Willamette L. Rev*, 33, 703- 744

Kendall J., (2000) Choosing a System for Resolving Commercial Disputes *International Company and Commercial Law Review*, 11, 82-86

Kovach K., and Love P., (1996) Evaluative Mediation in an Oxymoron, *Alternatives to High Cost Litigation*, 14, 31-32

Law Society (2009) *Civil and Commercial Mediation Accreditation Scheme: Criteria and guidance notes*,

http://www.lawsociety.org.uk/new/documents/accreditation/ccmas_guidance.pdf (Accessed on the 23rd June, 2010)

Landmark Website (2010)

http://www.landmarkchambers.co.uk/__data/assets/pdf_file/0004/82876/Landmark_Mediation_Agreement.pdf (Accessed 24 June, 2010)

Law Society, 2010 website

<http://www.lawsociety.org.uk/productsandservices/accreditation/accreditationcivilcommercialmediation.page> (Accessed on 23 June 2010)

Legal 500 Website (2010) <http://www.legal500.com/> &

<http://www.legal500.com/c/london/dispute-resolution/mediators> (Accessed on the 23 June, 2010)

Mason J., (2009) Ethics in the construction industry: the prospects for a single professional code, *International Journal of Law in the Built Environment*, 1, 3 194-204

McAdoo B., and Hinshaw A., (2002) "The Challenge of Institutionalising ADR: Attorneys' Perspectives on the Effect of Rule 17 on Civil Litigation in Missouri", 67 *Missouri Law Review*, 67, 473-572

McCorkle S., (2005) The Murky World of Mediator Ethics: Neutrality, Impartiality and Conflicts of Interest in State Codes of Conduct *Conflict Resolution Quarterly*, 23, No 2, 165-183

Menkel Meadow C., (2009) Are there systematic ethics issues in dispute system design? And what we should [not] do about it: Lessons from International and Domestic Fronts (2009) *Harv. Negot. L. Rev.*, 14, 195-231

Menkel Meadows C., (1997) Ethics in ADR: New Issues, No Answers from the Adversary Conception of Lawyers; Responsibilities *S. Tex. L. Rev.* 407 *South Texas Law Review*, 38, 407-453

Menkle-Meadow C., (1993) "Lawyer Negotiation: Theories and Realities – What we learn from Mediation" *Modern Law Review*, 56, 362-379

Meyer J., and Leathes M., (2008) How mediators can obtain professional certification and thereby elevate their profession: A look at IMI's voluntary credentialing programme *Disp. Resol. J.*, 63, October, 22-27

Miles R., (2004) Mediation: A Fledgling Profession or a Pot Puri of Good Intentions? *ADR News: The NADR Quarterly Newsletter*, Vol. 4 Issue No 3 October

Mills K., (2005) Can a single ethical code respond to all models of mediation? *Bond Dispute Resolution News*, Vol. 21, 5-41

Mistelis L., (2006) *ADR in England and Wales: A Successful Case of Public Private Partnership* in Alexander N., (Ed) (2006) *Global Trends in Mediation* (2nd Edition), Kluwer International, BV

National Mediators Helpline, (2010)
<http://www.nationalmediationhelpline.com/what-is-mediation.php> (Accessed on May 7, 2010)

Nolan-Haley J., (2002) Lawyers, Non-Lawyers and Mediation: Rethinking the Professional Monopoly from a Problem-Solving Perspective, 7 *Harvard Negotiation Law Review*, 7, 235-299 (Spring)

Nolan-Haley J., (1999) Informed Consent in Mediation: A Guiding Principle for Truly Educated Decision Making *Notre Dame L. Rev.*, 74, 775-839

Nolan-Haley J., (1998) Lawyers, Clients and Mediation, *Notre Dame Law Review*, 73, 1369-1390

Obermann S., (2008) Style Vs Model: Why Quibble? *Pepp. Disp. Resol.*, 9, 1-62

Press S., (2003) Institutionalization of Mediation in Florida: At the Crossroads, *Penn St LL Rev*, 108, 43-66

RIBA's Website *Code of Conduct* (2010)

<http://www.architecture.com/TheRIBA/Organisation/Constitution/CodeOfConduct.aspx> (Accessed on the 25 June 2010)

Riskin L., (1996) Understanding Mediators' Orientations, Strategies and Techniques: A Grid for the Perplexed *Harvard Negotiation Law Review*, 1, 7-51

Roberts S., (1992) "Mediation in the Lawyers' Embrace" *Modern Law Review*, 55(2), 258-264

Roberts S., (1993) "ADR and Civil Justice: An Unresolved Relationship" *Modern Law Review*, 56(3), 452-470

Roberts S., (2002) Institutionalised Settlement in England: A contemporary Panorama *Willamette Journal of International Law and Dispute Resolution*, 10, 17-35

Roberts S., (2009) "Listing concentrates the mind": The English civil court as an arena for structured negotiation *Oxford Journal of Legal Studies*, 29(3), 457-479

Roberts S., (2000) "Settlement as Civil Justice" *Modern Law Review*, 63(5), 739-747

TCC Court Guide 2nd Edition of the TCC Court Guide (2005), (revised with effect from 1st October 2007) at http://www.hmcourts-service.gov.uk/docs/tcc_guide.htm (Accessed on 23 May 2009)

Shipman S., (2006) Court Approaches to ADR in the Civil Justice System, *Civil Justice Quarterly*, 25(April), 181-218

Sorabji J., (2008) Costs – further developments from Halsey: Nigel Witham Ltd v. Smith & Isaacs and S v. Chapman, *Civil Justice Quarterly*, 27(4), 427-432

Sourdin T., (2007) Australian National Mediator Accreditation Scheme Report on the Project

<http://www.nswbar.asn.au/docs/professional/adr/documents/AccreditationReportSept07.pdf> (Accessed on the 3rd March 2010)

South J., (2009) Development of Commercial Mediator Training in England and Wales, <http://www.mediate.com/articles/southJ1.cfm> (Accessed on the 30th May 2010)

Stipanowich, T., (1996) Beyond Arbitration: Innovation and Evolution in the United States Construction Industry, *Wake Forest Law Review*, 31, 1, 65-182

Stipanowich T., (1998) "The Multi-Door Contract and Other Possibilities", *Ohio State Journal on Dispute Resolution*, 13, 303-404

Stulberg J., (1997) "Facilitative Versus Evaluative Mediator Orientations: Piercing the Grid", *Florida State University Law Review*, 24, 985-1005

Tecbar Website (2010) http://www.tecbar.org/adr_index.asp (Accessed on 28th June, 2010)

TeCSA Website (2010) http://www.tecsa.org.uk/mediation_adr-1328.htm (Accessed on 28 June, 2010)

TeCSA Mediation Rules (Website http://www.tecsa.org.uk/mediation_rules-1336.htm (Accessed on the 28 June, 2010)

Welsh N., (2001a), "Making Deals in Court Connected Mediation: What's Justice got to do with it?" *Washington University Law Quarterly*, 9, 787-861

Welsh N., (2001b) "The Thinning Vision of Self-Determination in Court Connected Mediation: The Inevitable Price of Institutionalisation?" *Harvard Negotiation Law Review*, 6, 1-96

Wissler R., (2002) "Court-connected Mediation in General Civil Cases: What we Know From Empirical Research", *Ohio State Journal on Dispute Resolution*, 17(3), 641-703

Woolf (Lord) (1995) (The Right Honourable The Lord Wolf, Master of the Rolls) *Access to Justice: The Interim Report* London: HMSO.

Opinions of Legal Professionals Regarding the Selection of Appropriate Resolution Method in Construction Disputes

Dr. Deniz ILTER¹

¹Istanbul Technical University,
Turkey

Email: artande@itu.edu.tr

Abstract:

The construction industry has managed to develop and adopt many unique methods to resolve disputes that occur in projects. However, the justification for implementing these methods has been primarily upon contractual requirements, government regulations, court order or basic reactionary instinct but not on measured utility. Despite being an industry keenly focused on quantitative results, many researchers in the field argue that parties involved in a dispute fail to analyse the actual utility associated with each dispute resolution method in a particular situation. Indeed, due to the fragmented and complex nature of construction projects, there is no one best way of dealing with disputes, as often they are different in scale, complexity and nature. In deciding which dispute resolution method to apply, there may be a need to take into consideration various external factors, such as technical, political, financial, social, economical and legal. In this context, this paper aims to examine the current practice in dispute resolution method selection in the Turkish construction industry through interviews with legal professionals. The results reveal remarkable findings regarding the opinions of the legal interviewees on the importance of method selection, parties' behaviour patterns in the selection process and how legal professionals advise their clients in determining the appropriate method.

Keywords:

Dispute resolution, legal professionals, method selection, Turkey.

1 Introduction

For decades, the construction industry has been mired in adversarial relationships between the parties. Each party's priorities are unsurprisingly at conflict with each other's, establishing a repetitive cycle of hostilities. The owner usually wishes to obtain maximum quality, functionality and capacity at minimum cost. The contractor, while hoping to develop a satisfied client, must in the long run achieve financial goals that are advanced by expending the minimum resources required to meet a minimum scope of work (Gibson and Gebken, 2006). When the profit position of a party in a construction project is threatened, dispute will emerge (Hibberd and Newman, 1997).

The true cost of any dispute includes any time spent by staff and consultants in preparation and attendance at any resolution attempt, in addition to the direct cost to the

project and indirect costs to the organisations involved. A major factor in the total end cost of any project will be how effectively disputes can be avoided, managed and resolved (Carmichael, 2002). As the cost of any project is one of the most important issues to any owner, it is paramount that all parties prescribe to meaningful and structured methods to deal with disputes.

Construction industry has managed to develop and adopt many unique ways to resolve disputes. However, the justification for implementing these methods has been primarily upon contractual requirements, government regulations, court order, limited previous experience or basic reactionary instinct but not on measured utility. Despite being an industry keenly focused on quantitative results, parties involved in a dispute fail to analyse the actual utility associated with each dispute resolution method in a particular situation (Harmon, 2003).

Selecting a dispute resolution process is the first step to resolve a dispute and this is an important decision because of the resource implications. Due to the fragmented and complex nature of construction projects, there is no one best way of dealing with disputes, as often they are different in scale, complexity and nature. In deciding which dispute resolution strategy to apply, there is a need to take into consideration various external factors, such as technical, political, financial, social, economical and legal. In current practice, decisions concerning dispute resolution normally are based on qualitative assessment and refer back to previous experience. To a large extent, qualitative assessment is subjective in nature, and may be biased towards the personal preference of the assessor (Cheung and Suen, 2002).

In this context, this paper aims to examine the current practice in dispute resolution method selection in the Turkish construction industry through interviews with legal professionals in the quest for describing the factors affecting the selection, parties' behaviour patterns in the selection process and how legal professionals advise their clients in determining the appropriate method.

2 Methodology

Expert sampling was used in the exploration of the opinions of legal professionals regarding the current practice in dispute resolution method selection in the Turkish construction industry. This method was chosen because the number of legal professionals (lawyers, claim consultants and legal scholars) who have actual experience of implementing various methods (especially Alternative Dispute Resolution – ADR) is still limited in Turkey. With the aim of reaching the experts, a letter was sent to request an interview to the members of Alternative Dispute Resolution Centre of the Ankara Bar, members of the Arbitration - Mediation and Expert Appraisal Centre of the Istanbul Chamber of Commerce and legal scholars who have publications about ADR. There were 12 interviews in total, which comprised 4 claim consultants (working for prominent contractor firms in Turkey), 5 construction lawyers and 3 legal scholars specialized on ADR. In depth interviewing methodology was used in collecting data from the legal professionals, which enabled detailed discussion of the issues tackled in the research.

3 Importance of Selecting the Appropriate Method

The first issue discussed was the importance of selecting the appropriate method in the resolution process. This was an open ended question and legal interviewees were asked to explain their opinions and experience related to method selection.

All interviewees stated that selection of the method is of critical importance in the resolution process. Below are some quotations from the interviewees:

“Employing an appropriate resolution process is critical to resolve construction disputes. This is because having an appropriate resolution process should pave the path to success. In other words, one of the key decisions in dispute resolution is choosing an appropriate resolution process.”

“Selection of the appropriate method is half the way to a successful resolution since every single dispute is different in nature and requires a resolution method suited to this nature.”

One of the interviewees pointed out that selection of different methods may bring on different results for the dispute. *“While selection of a particular method may bring an advantage in the short run, another could bring advantage in the long run, thus the importance of selecting the appropriate method is unquestionable.”* It was also stated that selection of the appropriate resolution method not only affects the outcome of the dispute but also the future of the relationship of the parties, which is frequently even more important.

Another interviewee stated that the method selected by a party in the resolution process is usually seen as an indicator of their intentions regarding the resolution by the adverse party. *“If a party offers mediation, he/she can be said to be more willing to resolve the dispute promptly than a party who offers arbitration.”*

It was also stated that if method selection is performed by a professional with appropriate knowledge and experience, this could contribute significantly to the success of the resolution process.

4 Current Practice Regarding Method Selection in the Turkish Construction Industry

All of the interviewees stated that rarely a resolution method is selected with an analytic approach that includes assessment of the utilities associated with each method. Ten out of twelve interviewees used the word “unconsciously” to characterize the current practice regarding method selection in the Turkish construction industry. It was stated that, in most disputes, parties try negotiation and if this is not successful, dispute is taken to arbitration or litigation right away. Legal interviewees stated that this is due to lack of knowledge, especially of various ADR methods. One of the interviewees stated that *“this is unfortunate given the fact that mediation and some other ADR methods are more suitable for construction and family disputes than any other types of disputes. If mediation is publicised and practiced in a right way in Turkey, it will be successful ...”*

As a part of this question, the interviewees were asked to rate the “frequency of the use of an analytical approach in dispute resolution method selection” in “preparation of the contract” and “execution of the contract” phases on a scale of 1 (very rare) to 5 (very often). Figure 1 shows the average of the responses given by the interviewees. The interpretations of the interviewees are that analytical approach is “rarely” used in dispute resolution method selection in both preparation of the contract (1.86) and execution of the contract (2.14) phases.

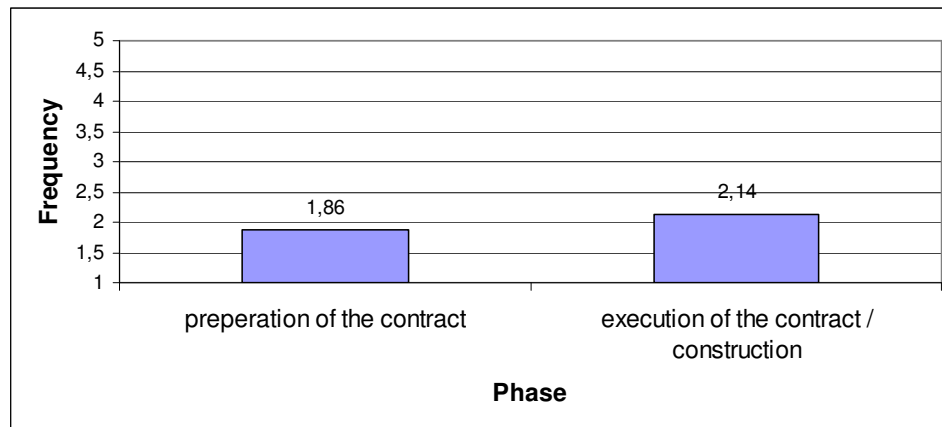


Figure 1. Frequency of the use of an analytical approach in dispute resolution method selection.

The interviewees listed the common reasons behind poor assessment in the selection of a dispute resolution method as follows:

1. Parties and/or their advisers are not aware of various alternative dispute resolution methods available.
2. Parties tend to punish the other party instead of analysing the utility associated with different methods.
3. Some legal advisors / lawyers tend to recommend methods that are more profitable for themselves but not necessarily suitable for the dispute.
4. Parties and/or their advisers do not know how to assess the utility associated with each method.
5. Parties and/or their advisers do not have the necessary skills, such as communication techniques, observation, conflict management etc to implement alternative dispute resolution methods.
6. Prejudice against alternative dispute resolution methods among some legal advisors / lawyers. One of the interviewees stated that “*Architects and engineers working in the construction industry are more interested in and willing to use ADR methods whereas some lawyers have a prejudice against these methods.*”
7. ADR methods are perceived as intermediate techniques and are not considered in the event of a dispute.

5 Parties' Behaviour in the Selection of Dispute Resolution Method

Another issue discussed was parties' behaviour in the selection of dispute resolution method. Interviewees were asked to rate the "frequency of behaviour patterns adopted by parties in the selection process" on a scale of 1 (very rare) to 5 (very often). Figure 2 shows the average of the responses given by the interviewees. Inertia (unwillingness to change thought patterns that have been used in the past in the face of new circumstances) was rated as most frequent. Negligence (negligence of important decision criteria and decision alternatives mostly due to lack of knowledge), selective perception – prejudice (actively screening-out information that is not thought to be salient), focusing on short term results (focusing on short term achievements with a high motivation for "winning") and emotional selection (selection of a method with a high motivation for punishing the other party) follow.

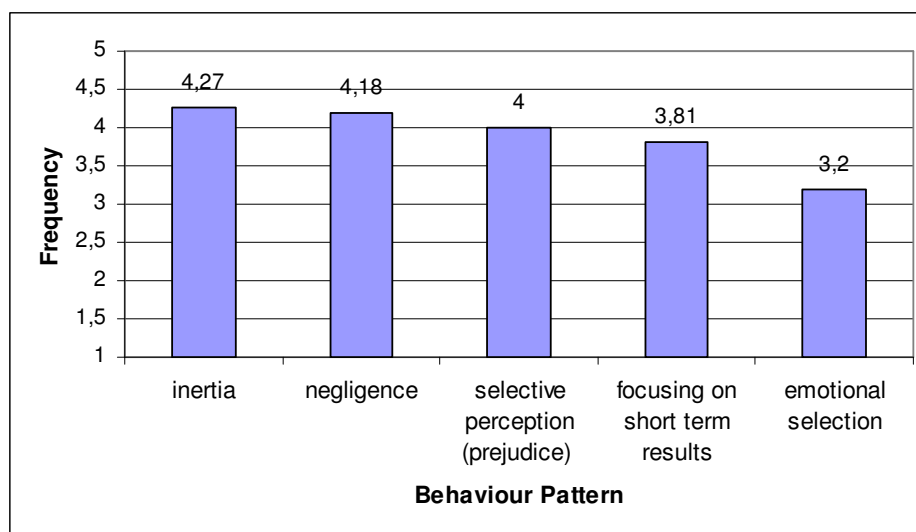


Figure 2. Frequency of behaviour patterns adopted by parties in the selection process.

One of the interviewees stated that *"Provoked disputants do not objectively consider their cases. In those circumstances, it is difficult to make reasoned decisions, hence the parties should decide on a method of dispute resolution depending on several criteria."*

Another stated that *"Professional advisers and/or analytical decision making tools are needed for logical decision making, which is an important part of all science-based professions, where specialists apply their knowledge in a given area to making informed decisions."*

6 Factors Affecting the Recommendation of Legal Professionals in the Selection

The last issue discussed was factors affecting the recommendation of legal professionals in the selection. This was an open ended question and legal interviewees were asked to list the factors they consider as important in selecting a resolution method.

The interviews revealed that, in current practice, the most important factor in determining which resolution method will be used in the event of a dispute is the “related contract clause”. This is mainly because parties who seek legal assistance try negotiation and if this is not successful, they consider adversarial methods to resolve their dispute and refer to the contract, which would in most cases include litigation or arbitration. However legal interviewees stated that if a party agreed to consider other dispute resolution methods (other than negotiation and litigation / arbitration), there would be a need to assess many factors. Following factors have been specified by the interviewees:

1. Will of the party they are representing (if the party they are representing has already selected a method or methods).
2. Current relationship between the parties.
3. Their will in continuing the relationship with the other party.
4. Nature of the dispute; whether a court order is seen as a necessity in resolving the dispute.
5. Financial position of the party, the amount they are willing to pay for the resolution of the dispute.
6. Time granted by the party for the resolution of the dispute.
7. Risk of delay associated with the dispute.
8. Value / amount of the dispute.
9. Value / amount of the project.
10. Legal aspects of the dispute (Whether an enforceable court order or a precedent is needed to resolve the dispute)
11. Technical expertise required to resolve the dispute.
12. Privacy needed / trade secrets etc.
13. Parties’ degree of knowledge of and trust in various methods.
14. Degree of control the party wants to have over the resolution process / flexibility needed.
15. Whether a binding or non-binding decision is preferred.
16. Whether the party wants to be able to choose the third party to resolve the dispute.

Most of the experts also added that they prefer using ADR methods to resolve construction disputes because of the win-win solutions these can offer¹. ADR methods are also seen as “especially suitable” for construction disputes due to the fact that disputes hinder the progress of a construction project causing a considerable cost for all the parties and require prompt resolution. Interviewees also stated that they do not prefer litigation due to the inadequacy of the court system (“*in terms of both quantity and quality*”) and lack of a specialized construction court in Turkey.

7 Conclusions

Selecting a dispute resolution process is the first step to resolve a dispute and this is an important decision because of the resource implications. Due to the fragmented and complex nature of construction projects, there is no one best way of dealing with disputes, as often they are different in scale, complexity and nature. In deciding which dispute resolution strategy to apply, there is a need to take into consideration various external factors, such as technical, political, financial, social, economical and legal. In current practice, decisions concerning dispute resolution normally are based on qualitative assessment and refer back to previous experience. To a large extent, qualitative assessment is subjective in nature, and may be biased towards the personal preference of the assessor.

In this context, the current practice in dispute resolution method selection in the Turkish construction industry is examined in this paper through interviews with legal professionals in the quest for describing the factors affecting the selection, parties’ behaviour patterns in the selection process and how legal professionals advise their clients in determining the appropriate method.

The results of the interviews revealed that;

- Legal professionals think selection of the method is of critical importance in the resolution process.
- Selection of the appropriate resolution method not only affects the outcome of the dispute but also the future of the relationship of the parties, which is frequently even more important.
- Selection of different methods may bring on different results for the dispute.
- The method selected by a party in the resolution process is usually seen as an indicator of their intentions regarding the resolution by the adverse party.
- Rarely a resolution method is selected with an analytic approach that includes assessment of the utilities associated with each method in the Turkish construction industry.

¹ It should be noted that all of the interviewees are ADR experts and their views regarding ADR do not represent the view of legal professional in general in Turkey.

- In most disputes, parties try negotiation and if this is not successful, then the dispute is taken to arbitration or litigation right away in the Turkish construction industry.
- Inertia, negligence, selective perception – prejudice, focusing on short term results and emotional selection are often seen in parties' behaviour in the selection of a dispute resolution method.
- Legal professionals consider diverse factors while recommending a resolution method such as, will of the party they are representing, current relationship between the parties, parties' will in continuing the relationship with each other, nature of the dispute, financial position of the party, time granted by the party for the resolution of the dispute, risk of delay associated with the dispute, Value / amount of the dispute and the project, legal aspects of the dispute, technical expertise required to resolve the dispute, privacy needed, parties' degree of knowledge of and trust in various methods, flexibility needed in the process, whether a binding or non-binding decision is preferred and whether the party wants to be able to choose the third party to resolve the dispute.

Given the industry's reliance on the subjective approach to dispute resolution selection, there is a need for a more systematic approach. Selecting a dispute resolution process is the first step to resolve a dispute and this is an important decision because of the resource implications. The industry's approach in the selection of dispute resolution method has been heavily criticized, where too much reliance is placed on intuitive judgments rather than on rational approach. Considering the large damages that result from construction conflicts due to delays, cost overruns, and decreasing productivity, the construction industry needs to employ new decision support technologies for evaluating and comparing alternative decisions.

8 References

- Carmichael D (2002) Disputes and International Projects, The Netherlands, Swets & Zeitlinger.
- Cheung SO and Suen HCH (2002) A Multi Attribute Utility Model for Dispute Resolution Strategy Selection Construction Management and Economics 2002 (20).
- Gibson GE and Gebken RJ (2006) Quantification of costs for dispute resolution procedures in the construction industry Journal of Professional Issues in Engineering Education and Practice 2006 (07)
- Harmon KMJ (2003) Resolution of Construction Disputes: A Review of Current Methodologies, Leadership and Management in Engineering 3(4) 187-201.
- Hibberd P and Newman P (1999) ADR and Adjudication in Construction Disputes, Oxford, Blackwell.

The development of Statutory Adjudication in the UK and its relationship with construction workload

Peter Kennedy¹, Janey Milligan², Lisa Cattanach² and Edward McCluskey¹

¹School of the Built and Natural Environment,
Glasgow Caledonian University, Glasgow G4 0BA,
United Kingdom

² Construction Dispute Resolution Ltd.,
Glasgow G69 6GA
United Kingdom

Email: P.Kennedy@gcu.ac.uk; jlm@cdr.uk.com; lhc@cdr.uk.com;
Edward.McCluskey@gcu.ac.uk

Abstract:

This paper tracks the development of Statutory Adjudication in the UK since its inception in 1998. It presents data on numbers of referrals to adjudication in the UK, particularly those which were routed through Adjudicator Nominating Bodies (ANBs). The paper is based on data from the Adjudication Reporting Centre at Glasgow Caledonian University which draws its information from questionnaires received from ANBs and from samples of practicing adjudicators. Trends in the causes of disputes referred, parties in dispute, procedures adopted by adjudicators, etc are also included in this review. The paper considers how the nature of the disputes has changed over the course of 12 years as parties have grown familiar with the process; how it has been utilised during a period of recession in construction and the re-emergence of the criticality of cash flow as firms attempt to cope with increased competition and reduced margins. The study found that there appeared to be a relationship between the UK workload and the number of adjudication referrals. In the immediate aftermath of a downturn in workload adjudications rose, but one year later the ongoing reductions in workload were followed by a reduction in referrals. The views of a number of experts were sought and they confirmed that adjudication was still highly regarded and that they had not experienced any diminution in the use of adjudication.

Keywords:

Construction Disputes, Dispute Resolution, Statutory Adjudication.

1 Introduction

The data used in compiling this paper comes from the Adjudication Reporting Centre (ARC) which collects data from Adjudication Nominating Bodies (ANBs) in the UK and also from adjudicators. This study refers to data from ANBs, published government statistics and commentaries from leading authorities. Statutory Adjudication was introduced to the UK through the Housing Grants, Construction and Regeneration Act

1996¹ (HGCR Act) and enabled through The Scheme(s) for Construction Contracts² and their respective Exclusion Orders³. At the time of its introduction there were concerns about the willingness of weaker contracting parties (mostly subcontractors) to initiate adjudication proceedings against more powerful parties (Kennedy, Morrison and Milne 1997). The major concern expressed by subcontractors, albeit their experience was almost entirely limited to adjudications in respect of set-off claims in domestic subcontracts, was that if they referred a dispute to adjudication they might be denied future opportunities to tender for work. It was believed by many that weaker parties (usually further down the contracting chain) would be deterred by the threat of commercial power from utilising this new and powerful form of resolving disputes (in the interim) and allowing vital cash to flow through the subcontracting chain. This view is supported by Gaitskell (2007) who reported that during the pre-statutory adjudication era, resistance by main contractors and employers meant that contractual provisions which aided cash flow were not widely used. Gaitskell suggested that the inclusion of such provisions would depend on the 'negotiating strength' of the parties. If the objecting party was strong enough, the contractual provisions were removed. Mason (2010), reflecting on a report by Huxtable (1983) records the atmosphere between main contractors and subcontractors which prevailed at that time. The report by Huxtable was entitled 'The corruption of the commercial process' and it detailed a number of evolving practices, several involving bespoke contracts, which many would consider unfair.

Despite these concerns the statutory adjudication process was adopted actively. Until relatively recently (around 2007) the UK construction industry experienced growth in its workload. This coincided with the settling-in period of adjudication which was subject to various legal challenges and practices which have led to the need for review and proposed changes to improve the effectiveness of the process. Many of these challenges were the result of more powerful parties seeking to increase the barriers to the use of adjudication and to dissuade others from adjudicating⁴. Over this period, the adjudication process operated in a time of relatively stable workload in the construction industry. It is only within the last three years or so that it may be possible to see if and how the adjudication process may be impacted upon by the reduced financial circumstances in which the UK construction industry has found itself. This paper attempts to address this impact.

2 Data collection

The data for this paper was collected by the ARC which collects data through questionnaires from ANBs throughout the UK. ARC also collects data from participating adjudicators which provides insights into the operation of the adjudication process. The work of ARC started in 1998 when the 1996 HGCR Act came into force,

¹ Housing Grants, Construction and Regeneration Act 1996

² The Scheme for Construction Contracts (Scotland) Regulations 1998 and The Scheme for Construction Contracts (England & Wales) Regulations 1998

³ The Construction Contracts (Scotland) Exclusion Order 1998 and The Construction Contracts (England and Wales) Exclusion Order 1998

⁴ *Bridgeway Construction Ltd v Tolent Construction Ltd* (2000) CILL1662 also see *Yuanda (UK) Co Ltd v WW Gear Construction Ltd* [2010] EWHC 720 (TCC)

and it has collected data continuously from then to the present date. ANBs are asked to provide information on a wide array of areas of interest, only a few of which are referred to in this paper.

3 Development of statutory adjudication

3.1 Role of the Adjudicating Nominating Bodies (ANBs)

Adjudication referrals in the UK are mostly handled by ANBs. Disputes are usually referred to ANBs which then appoint the adjudicators. ANBs also administer the training and qualifications of adjudicators who are registered with them. There are a number of ANBs (as shown below) but there has been some consolidation in the provision.

Table 1 below indicates the responses of a number of ANBs over a period of years with regard the number of registered adjudicators. This indicates that some ANBs are no longer operating and in many cases the number of adjudicators registered with individual ANBs has reduced. This may be because they moved their registration to another body or have reduced the number of ANBs with which they register – some adjudicators are listed with several (typically 2 or 3) to maximise their opportunities for appointments. Each ANB has its own criteria that has to be met for an adjudicator to be listed.

Table 1. Numbers of adjudicators registered with ANBs

ADJUDICATOR NOMINATING BODY	May 2002	October 2002	April 2003	Feb 2004	October 2004	October 2005	April 2006	October 2006	October 2007	April 2008	April 2010
Association of Independent Construction Adjudicators	194	194	176	150	155	116	50	50	49	48	48
Chartered Institute of Arbitrators	147	n	n	164	158	169	180	172	171	164	164*
Confederation of Construction Specialist	25	24	14	21	43	43	43	43	43	43	15
Construction Industry Council	170	144	149	154	152	132	136	136	102	102	94
Institution of Chemical Engineers	NR	6	13	13	13	14	18	15	15	15	15
Institution of Civil Engineers	80	82	87	84	84	82	84	81	81	81	72
Royal Institute of British Architects	70	69	69	68	63	68	68	68	67	67	50
Royal Institution of Chartered Surveyors	107	116	117	122	122	122	116	116	104	115	120
3A's Polycon AIMS Ltd	33	NR	54	NR	NR	NR	NR	NR	NR	NR	NLO
Institution of Mechanical Engineers	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Chartered Institute of Building	53	52	50	53	58	56	56	54	46	43	43
Construction Confederation	43	43	43	43	43	43	28	28	28	27	NLO
Scottish Building	12	12	12	12	12	9	11	11	9	9	9
Royal Incorporation of Architects in Scotland	14	14	14	14	14	14	10	10	10	10	13
Royal Institution of Chartered Surveyors in Scotland	47	45	35	38	38	38	25	23	23	23	20
Centre for Effective Dispute Resolution	48	46	NR	40	NR	41	41	42	42	42	32
Institution of Electrical Engineers	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Technology and Construction Solicitors Association	128	128	128	149	142	142	142	142	133	133	128
Chartered Institute of Arbitrators (Scottish Branch)	22	17	17	17	20	21	21	17	16	16	17
The Law Society of Scotland	10	11	13	16	18	18	18	28	28	28	33
Technology and Construction Bar Association	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR	NR
Adjudication.co.uk	NR	7	7	NR	NR	NR	NR	NR	NR	NR	NR
TOTALS	1203	1010	998	1158	1135	1128	1047	1036	967	966	921

NR - not reporting

NLO - no longer operating

* provisional

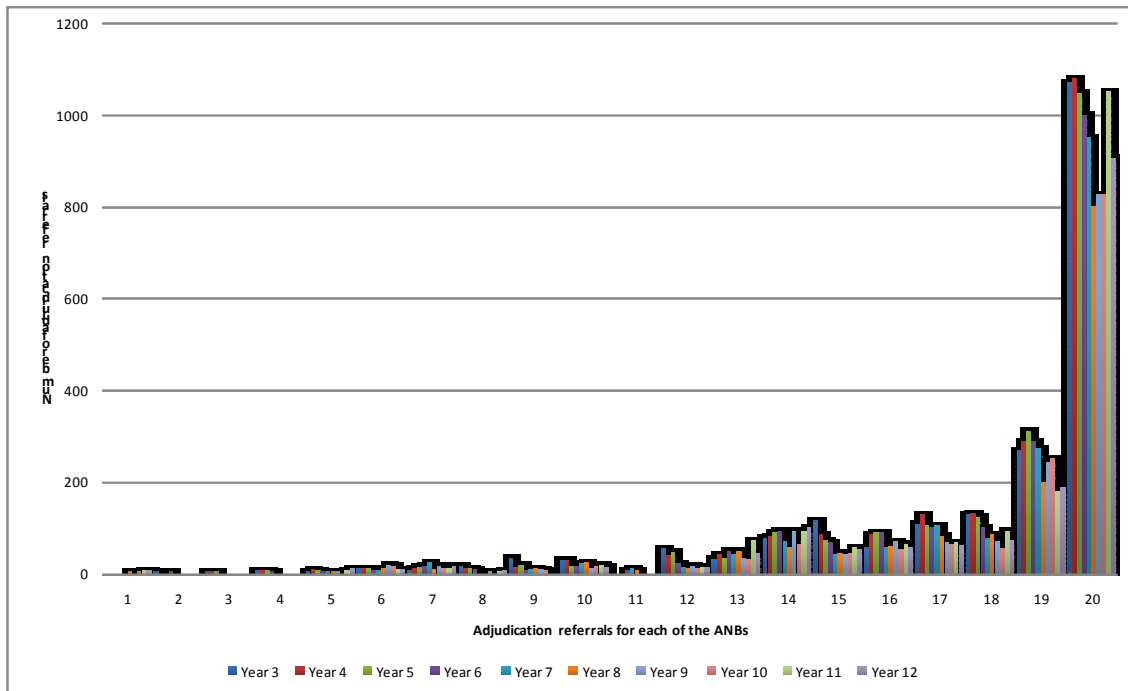


Figure 1. Numbers of adjudication referrals for each ANB

Some ANBs have more referrals coming to them than others as Figure 1 illustrates. In this histogram the identity of ANBs has been removed but it can be seen that there is one dominant body which accounts for approximately 60% of all adjudications in the UK.

Whilst the ANBs are involved in the appointment of the majority of adjudicators, they do not have a monopoly in the appointing process. The 1996 Act in the UK does not restrict appointments through an ANB. The sources of appointment of adjudicators are shown in Table 1 below.

Table 1. Sources of appointment of adjudicators

Source of appointment	April 2002	July 2004	Oct 2005	Oct 2007	April 2008
Through an ANB	89.6%	83.1%	83.2%	85.0%	87.3%
By agreement between the parties	9.9%	16.6%	16.8%	12.0%	12.0%
Named in the contract	0.6%	0.3%	0.0%	3.0%	0.7%

It will be observed here that, while being named in the contract is unusual, it is a fairly common practice for parties to agree on the adjudicator when the dispute has arisen.

3.2 The adjudication process

The professional skill base of adjudicators has fluctuated over time but the two principle groups are quantity surveyors and lawyers. The increasing proportion of lawyers reflects perhaps the matters referred in adjudications as they have evolved over time. These professionals are followed by a group comprising civil engineers, architects and professional builders.

Table 2. The primary professional disciplines of the adjudicators

DISCIPLINE	May 2002	Oct 2002	April 2003	Feb 2004	Oct 2004	Oct 2005	April 2006	Oct 2006	Oct 2007	April 2008	April 2010
Quantity Surveyors	28.9%	39.1%	43.8%	41.6%	38.0%	38.8%	35.7%	35.1%	34.5%	31.4%	33.5%
Lawyers	22.1%	21.9%	22.1%	21.6%	26.1%	26.3%	25.6%	26.6%	26.6%	28.4%	15.6%
Civil Engineers	14.6%	17.3%	13.2%	11.1%	11.6%	11.0%	15.8%	15.1%	15.0%	14.5%	14.1%
Architects	7.8%	8.9%	10.2%	9.3%	9.6%	9.6%	9.0%	8.8%	8.7%	9.6%	8.1%
CIOB/Builders	3.4%	3.4%	2.6%	5.2%	5.0%	4.9%	4.7%	5.3%	4.9%	7.5%	4.7%
Building Surveyors	2.1%	1.7%	1.4%	1.0%	1.3%	1.4%	1.2%	1.2%	1.3%	2.5%	0.2%
Construction Consultants	2.5%	0.3%	0.9%	4.1%	5.3%	4.7%	4.6%	4.8%	5.6%	1.9%	0.5%
Structural Engineers	2.1%	3.4%	0.8%	2.2%	1.3%	1.3%	1.2%	0.9%	1.1%	1.3%	1.6%

Table 3. Procedure adopted by adjudicators

Procedure Adopted	to October 2001	to October 2002	to July 2004	to Oct 2005	to October 2007	to April 2008
Employ a documents only procedure	56.0%	52.0%	56.9%	46.8%	57.6%	77.9%
Employ an interview procedure with one party present	3.0%	0.3%	0.8%			
Employ an interview procedure with both parties present	35.0%	21.0%	24.6%	30.3%	15.2%	10.6%
Carry out a full hearing procedure	6.0%	6.0%	8.1%	15.6%	14.1%	7.7%
Carry out a conference call			5.8%	6.4%	10.9%	3.9%
Site Visit		11.0%	1.9%	0.9%	1.1%	
Legal debates			1.5%		1.1%	
Interview with contract administrator present			0.4%			
Other		1.0%				

The procedures adopted by the adjudicators have also evolved as can be seen from Table 3 above. The principal procedure adopted is that of a ‘documents only’ which gives the adjudicators greater protection from natural justice or procedural challenges. Statistically, site visits and legal debates are not favoured by adjudicators.

3.3 The Disputes

Table 4. Disputing parties

Contracting Parties	October 2005	October 2007	April 2008
Client and consultant	3%	5%	4%
Client and main contractor	35%	27%	37%
Client and nominated subcontractor	1%	0%	0%
Main contractor and domestic subcontractor	51%	52%	47%
Main contractor and nominated subcontractor	2%	0%	0%
Sub-contractor and sub sub-contractor	3%	8%	4%
Consultant and contractor	1%	2%	1%
Trade contractor and employer	2%	2%	3%
Management contractor and package contractor	2%	3%	4%
Consultant and sub-consultant	0%	0%	1%

The most common pair of parties in dispute remains main contractor and subcontractor but the client and main contractor account for a significant number. Others which have changed over time include nominated subcontractors who have disappeared from the scene and those involving package or trade contractors has increased, most likely due to the changing patterns of procurement choices made by clients.

Table 5. Subjects of the disputes

Subject	July 2004	October 2005	October 2007	April 2008
Valuation of Final Account	12%	14%	22%	22%
Failure to comply with payment provisions	19%	14%	8%	19%
Valuation of interim payments	15%	13%	15%	16%
Withholding monies	10%	11%	10%	10%
Extension of time	8%	8%	8%	9%
Loss and expense	9%	10%	2%	7%
Valuation of variations	15%	17%	11%	5%
Defective work	4%	5%	7%	4%
Determination	2%	3%	4%	4%
Non-payment of fees	2%	1%	7%	2%

Table 5 above includes adjudicators' reported views on the main causes of the disputes. This is clearly a matter of judgement as the issues to be decided may contain many elements. 'Final account' features prominently and has grown over the reporting periods to be the most common but that might simply be because the vast majority of adjudication referrals are initiated after Practical Completion and at that stage of projects any dispute is likely to impact in some way on the Final Account.

3.4 Growth trends of adjudications in UK

The rate of adoption of statutory adjudication since its introduction in 1998 is shown in Table 6 and Figure 2 below. This illustrates a rapid expansion over the first two years before it then levelled out for two years and then declined by around 25% to reach a new level.

Table 6 Adjudication referrals per year from 1998 to present

TIME PERIODS	ALL ANBs ADJUDICATIONS
YEAR 1 - May 1998 - April 1999	187
YEAR 2 - May 1999 - April 2000	1309
YEAR 3 - May 2000 - April 2001	1999
YEAR 4 - May 2001 - April 2002	2027
YEAR 5 - May 2002 - April 2003	2008
YEAR 6 - May 2003 - April 2004	1861
YEAR 7 - May 2004 - April 2005	1685
YEAR 8 - May 2005 - April 2006	1439
YEAR 9 - May 2006 - April 2007	1506
YEAR 10 - May 2007 - April 2008	1432
YEAR 11 - May 2008 - April 2009	1737
YEAR 12 - May 2009 - April 2010	1528

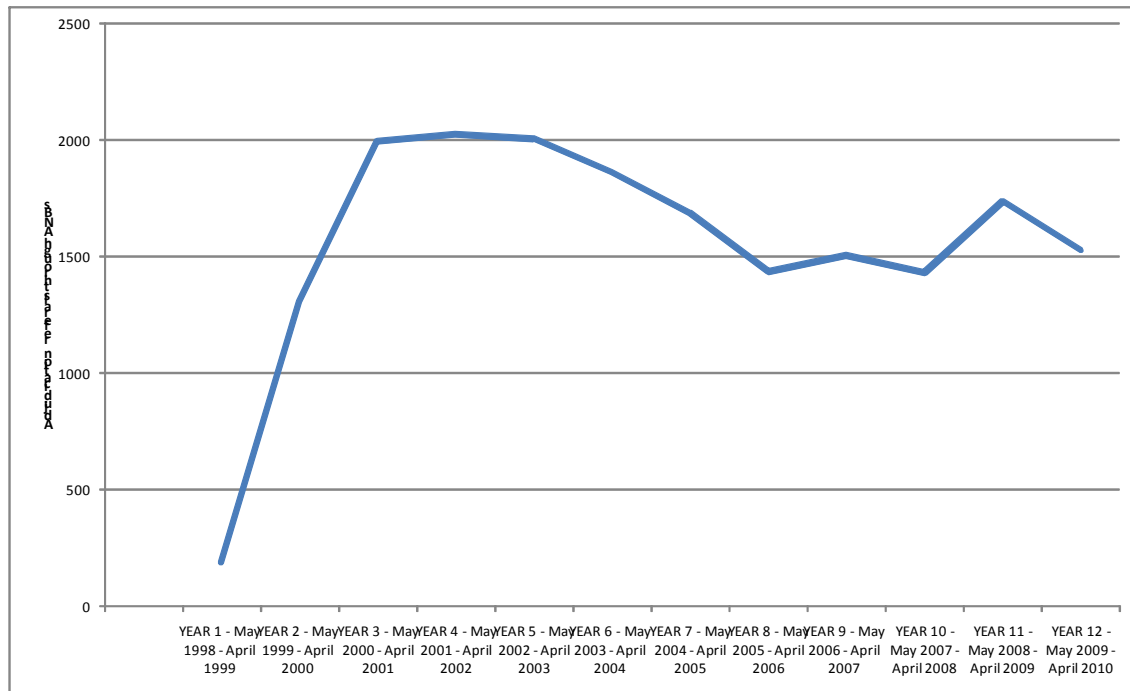


Figure 2. Graph of the number of adjudication referrals each year from 1998

This graph shows minor fluctuation between 2005 to 2007. However, there was a clear spike in demand during 2008/09 with this settling back down the following year. There has always been some fluctuation over the course of the years but since 1998 there has been some consistency in which months see peaks and others troughs. Figure 3 below illustrates this pattern.

The purpose of looking at the monthly variation in adjudication referrals was to see if there was any evidence of 'ambush' by parties. There are certainly peaks in November and March with troughs in September and December but the reasons may have more to do with financial year ends than 'ambush' which one might expect near holiday periods.

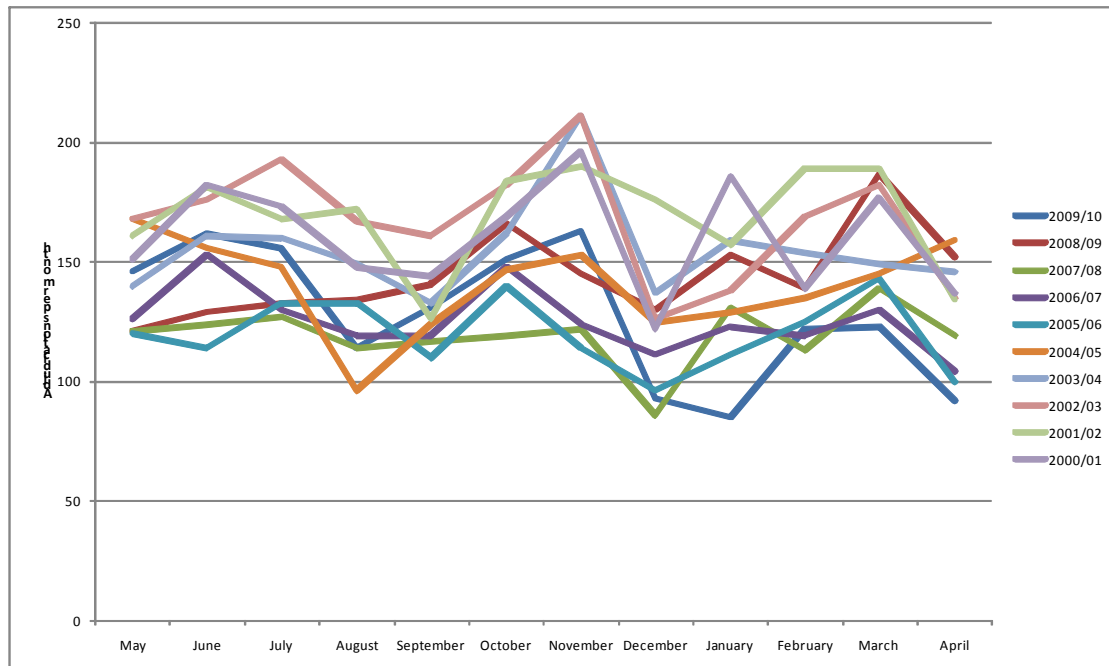


Figure 3. Monthly variations in adjudication referrals

4 Relationship with construction workload

From Figure 4 below, it can be seen that from year 1 to 3 there was a dramatic increase in the use of adjudication as might be expected from a novel form of dispute resolution following its introduction. From years 3 to 5, the number of referrals remained steady, which mirrored the UK construction workload over this period. However, it is interesting to note that - as the UK construction workload increased from years 5 – 7 the number of adjudication referrals started to decrease. From years 8 – 9, the UK workload decreased slightly and then started to increase whilst the number of referrals did the opposite on that occasion. From years 11- 12 the UK construction workload started to decline sharply as did the number of adjudication referrals. It appears from the graph - that when there was an increase in the UK construction workload, followed by a slight decrease - the number of adjudication referrals continues to increase. However, when the UK construction workload dropped more dramatically in 2010, the number of adjudication referrals also dropped noticeably.

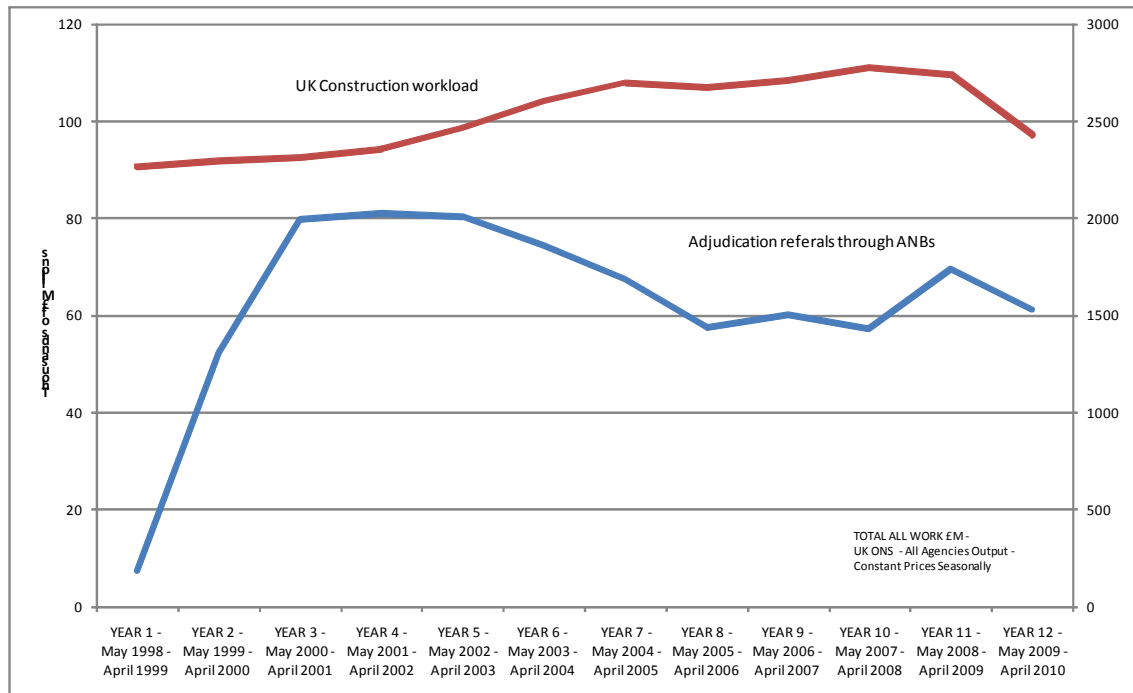


Figure 4. Adjudication referrals and UK construction workload

The last three years in this series are when the downturn in the UK construction workload begins to take effect. This coincided with a downturn in the economy as a whole and access to funding became more problematic. Figure 5 below shows the situation quarter by quarter. It has to be borne in mind that there may well be a time lag effect in that disputes may manifest themselves a year or more after the point at which

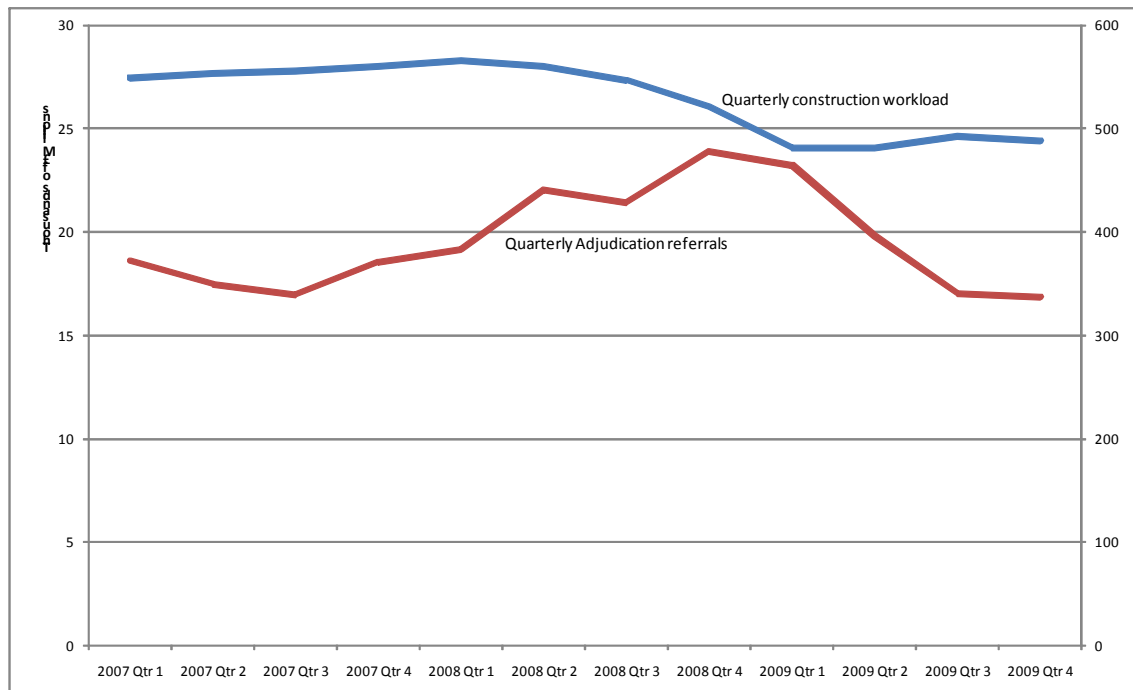


Figure 5. Quarterly profile of adjudication referrals and UK construction workload

the contract commenced when it would have been included as part of the workload statistics. It may also be the case that the motivation to pursue an adjudication may be influenced to a greater or lesser degree by the immediate requirements of cash flow and continuity of work.

The above graph shows a more detailed account of the UK construction workload and the number of adjudication referrals. The referrals have the normal peaks and troughs that occur throughout the year (see Figure 3); however, the trend follows that stated in the aforementioned paragraphs. It can be seen that Q1, 2008, shows a slight decrease in the UK construction workload with a correspondingly sharp increase in the number of adjudication referrals. This is interesting as it suggests that, at this point and even up to Q4 of 2008 with the workload declining and fewer opportunities to tender for subcontract work, subcontractors were increasingly pursuing adjudications: perhaps driven by the need for cash flow and having more time on their hands. However in Q4 of 2008, the decrease in workload becomes sharper and the corresponding number of referrals starts to decrease. This is likely to coincide with staff being paid off and money being tighter. Even when from Q1 to Q3, 2009, the UK construction workload increases slightly, there is a steep decline in the number of referrals. It is interesting to contrast the response here with that in 2008 when, despite a drop in workload, the referrals increased. In this case the referrals dropped when the same circumstances arose again. Perhaps this is down to the old adage of 'not biting the hand that feeds'. There would appear to have been a change in response pattern at the end of 2008. In Q4, 2009, the UK construction workload levels out and the number of referrals does likewise.

5. Expert commentary

A number of people with considerable experience in adjudication and who have extensive knowledge of the landscape of construction disputes were shown the initial findings of the research and asked to comment in an attempt to understand why these patterns were developing.

Geoff Brewer of FII Brewer Consultancy who provides Adjudication services, stated:-

'I'm not surprised by the research - I understand that the RICS has seen a 20% drop in adjudication nominations over this past year alone. There is little doubt that this is a consequence of the market, particularly the downturn in the private development sector. This is of course set to continue with further severe cuts to the public purse. Another factor however is the spiralling (and unrecoverable) cost of adjudication which has caused parties and their advisors to consider arbitration and/or litigation, both of which have responded well to the 'competition' provided by adjudication. There is also a steadily growing belief in mediation.

Finally I would add that many law firms now have a list of experienced and well regarded adjudicators from which they will try to appoint by agreement. This trend will continue to reduce the appointments recorded by the ANBs.'

On a positive note, Mark Entwistle Chair of the Association of Independent Construction Adjudicators, one of the ANBs stated:-

‘The use of adjudication is self-evidently linked to overall construction output, but the popularity of the process can now be said to be secure as the preferred formal dispute resolution mechanism for the construction industry. The relatively few adjudication cases that get referred to the courts also bears witness to its success. Though arbitration cases seem to be on the increase, especially for major disputes, there is little sign that the construction industry has any growing disaffection for adjudication and the Government’s plans to down-size the court service will likely serve to increase its use.’

John L Richies, Arbitrator & Adjudicator, Henry Cooper Consultants Ltd comments on trends and noted the increase of referrals in 2008 but questioned the quality:-

‘The trend has been that there have been more direct appointments. I think that is because particularly lawyers have become more sophisticated and have learnt to deal with the animal they get. I think that is the reason for direct appointments. Generally I think Adjudication is slower, that might sound like an odd thing, but it is because people are much more considered and we are getting more sensible submissions. There was a spell in 2008 where everybody jumped in and panicked because the bottom dropped out the housing market and submissions were very poor.

We moved away from any old thing and I do not think, based on my own experience, that there has been an actual real drop off for about 12 months. I think what we are dealing with at the moment is our workloads, live workloads, and until the government cuts bite we won’t know how small this market place can get. It is certainly my view at the moment that obviously if they reduce the amount of work there must be fewer disputes per se because there is less work. I think we are now moving into the area of silly pricing again and that will see a high level of disputes in the future. It is the same way that points make prizes “silly prices make disputes” inevitably. If you screw somebody down at the start, I do not care if it is the main contractor or the subcontractor, they will spend the whole job trying to get that money back.’

In summary, the trend is that in 2008 the increase in referrals was because the housing market dropped and everyone looked for their money and there were poor quality referrals at the peak. Now although the construction workload has gone down and referrals have gone down in line with that they have possibly gone further down because more people are agreeing adjudicators.’

Tony Bingham, Barrister stated:-

‘Attempting to pin down facts, draw conclusions, from workload data/dispute referrals requires caution at the best of times. In a recession (or whatever we call these times) requires an even broader brush. What I can say for sure is that I am very busy as Adjudicator (and arbitrator). Some folk are rummaging for the last penny and will fight for it. Some are fighting for work then in 2 years time will fight about it. The instant

referee, the adjudicator, is the most successful commercial answer and has been welcomed by industry...lawyers too."

With particular regard to the numbers of referrals the following points were made:-

Christopher Dancaster, Adjudicator and Arbitrator stated:-

'It has been suggested that the current downturn in activity in the construction industry might result in a major upsurge in adjudications. There would appear to be no immediate evidence of this. On a personal note the number of appointments as Adjudicator does not appear to be dropping dramatically but their complexity seems to be on the increase and there is a noticeable increase in the number of disputes referred that do not go through to a Decision.'

Ian Strathdee, Adjudicator and Arbitrator expressed a similar view:-

'The pattern and trend comparing the financial downturn of construction workload with the total ANBs adjudication referrals as indicated does not represent my experience from late 2008 through to 2010. I received far more adjudication referrals in that period than any other earlier period'.

Len Bunton of Bunton Consulting takes the consideration a step further to beyond the Adjudication being commenced. He stated

'There is no doubt that the desire to proceed to Adjudication is being influenced by the current downturn in the construction industry. Parties are keen to compromise to assist cash flow, and are reluctant to incur advisers and potentially adjudicators' fees. Two other issues are of concern. The immediate prospect of a jurisdictional challenge, and there is no doubt respondents are doing all they can to set the scene to contest enforcement, following the adjudication. Even if a successful Decision is received, there is no guarantee of payment, and there is also the strong prospect that the loser won't have the funds to pay, and the winner receives nothing, and has a costly bill from as I say the advisers and the adjudicator.

My message is tread carefully before you go down this route.

Concentration in the UK industry has to be on effective procurement, more care being given to contract terms, and efficiency on site to deliver a quality product.'

6. Discussion

These experts clearly believe that the adjudication process is still very well regarded by the construction industry, although it is suggested that more parties may be accessing adjudicators directly by agreement between the parties rather than through ANBs. This would partially explain the drop in ANB referrals. There is also the suggestion as discussed by Gould et al (2010) that other forms of dispute resolution may be being used more than previously and to some extent substituting for adjudication. This would

also result in some reduction in adjudications. Clearly however, these experts think that there is no noticeable reduction in the number of disputes in the industry and in the enduring popularity of Adjudication.

Gaitskell (2007) made the point that after the introduction of Adjudication in 1998 and the Civil Procedure Rules came into force on April 26 1999, there was an impact upon other forms of dispute resolution. Trials in the Technology and Construction Courts went down as did arbitrations. At the same time mediations and adjudications increased. There appears to be some dynamic relationship between the forms of dispute resolution in construction as well as some fluctuation in the overall number of disputes being submitted for resolution.

Salmond (2010) has shown the wide range of challenges being used by parties to resist the enforcement of adjudicators' decisions. These (almost routine it would appear) challenges to the adjudicator's decision raise the threshold of costs to the parties which may consequently render some smaller claims uneconomic to pursue. They also extend the period of time during which the referring party has to wait for payment, mitigating against speedy resolution advocated by Latham (1994).

One expert pointed to the actions taken by parties as a compromise to release cash. He also points to the difficulty of enforcement and to the possibility that the paying party may not have the funds. These latter concerns may help to explain some of the downturn in adjudications.

7. Conclusions

Over the past 12 years adjudication as a dispute resolution process has matured. It started slowly then accelerated to reach a peak after two years, fell away by about a quarter and then levelled out. The kind of disputes which are dealt with by adjudication has changed from being simple payment problems where the payment regime laid down in the Act was not being followed, to the present day when many disputes are concerned with large sums of money and complex legal questions. One possible explanation is that certain sectors of the disputes spectrum (ie smaller disputes) may be reducing disproportionately. Further research is required to determine if this is the case.

The adjudication process has been widely considered to have been successful but it is only in the past two to three years, when the industry has experienced a significant downturn in its workload and access to working capital has been restricted by banks, that it has been possible to witness this novel dispute resolution system working under stress. Whilst the adjudication process may reach a conclusion with the decision, that is not the end of the dispute as far as the parties are concerned. There is the issue of enforcement and if that becomes prohibitively expensive or protracted this, in itself, could impact on the attractiveness of adjudication as a means of resolving the dispute.

It is probably too early to discern a pattern of behaviours in response to the financial situation and restricted workload but it would be a retrograde step if the industry was to slip back into old (pre 1998) solutions to deal with its endemic problem of cash flow.

5 References

- Gaitskell, R., (2007), 'International statutory adjudication: its development and impact', *Construction Management and Economics*, **Volume** 25, Number 6, pp. 777-784.
- Gould, N., King, C., Britton, P., (2010), 'Mediating Construction Disputes: *An Evaluation of Existing Practice*' Kings College London, Centre of Construction Law & Dispute Resolution.
- Huxtable, J., (1983), 'The Corruption of the Commercial Process', Camberley; Confederation of Construction Specialists.
- Kennedy, P., Morrison, A. and Milne, D.O. (1997), 'Conflict between Main Contractors and Subcontractors: The resolution of disputes arising from the application of Set-Off', *Construction Management and Economics*, **Volume** 15, Number 6, pp. 527-537.
- Latham, M., (1994), 'Constructing the Team', HMSO, London.
- Mason, J., (2010), 'The corruption of the commercial process revisited', *Construction Law Journal*, Volume 26, Number 4, pp. 235-248.
- Salmond, K.T., (2010), 'The enforcement of adjudicator's awards under the Housing Grants Construction and Regeneration Act 1996: Part 36', *Arbitration*, Volume 76, Number 3, pp 538-551.

6 Acknowledgements

The research team wish to thank the Adjudicating Nominating Bodies of the UK and reporting adjudicators for their cooperation in submitting data to the Adjudication Reporting Centre at Glasgow Caledonian University. Thanks must also go to the group of experts who gave of their time to consider and comment upon this research.

A proposal for a ‘Dual Scheme’ of statutory adjudication for the building and construction industry in Australia

Michael C Brand¹ and Philip Davenport²

¹Faculty of the Built Environment,
The University of New South Wales, Sydney, NSW, 2052,
Australia

²Solicitor of the Supreme Court of New South Wales,
Australia

Email: michaelb@fbe.unsw.edu.au; p.davenport@bigpond.com

Abstract:

In Australasia there are two different models for adjudication in the building and construction industry. The first is the ‘Defined Scheme’. It is the scheme adopted under the ‘Australian Model’ of adjudication as applied in New South Wales, Victoria, Queensland, South Australia, Tasmania and the Australian Capital Territory. The Defined Scheme is also applied in Singapore. The second is the ‘Non-specific Scheme’. It is the scheme adopted under ‘UK Model’ of adjudication as applied in Western Australia, Northern Territory and New Zealand. This paper describes how a ‘Dual System’ of adjudication could work and how it would resolve many of the existing problems with the ‘Defined Scheme’ of adjudication in the Australian building and construction industry. The proposed Dual Scheme is a combination of the Defined Scheme and the Non-specific Scheme of adjudication. It allows for adjudication of Progress Claims (just as they are adjudicated now under the Defined Scheme), but also provides for separately adjudicated Money Claims in a similar way to that done under the Non-specific Scheme. The main advantage of the proposed Dual Scheme is that both parties to a construction contract can take advantage of adjudication like in ‘Non-specific Scheme’ of adjudication whilst simultaneously maintaining the integrity, and the relative time and cost benefits, of the Defined Scheme in dealing with progress payment claims. Examples of how the Dual Scheme works, as proposed here, are provided in the Appendix.

Keywords:

Australia, Dual Scheme, security of payment, statutory adjudication.

1 Introduction

1.1 Security of payment problem in the Australian construction industry

In August 2001, the Honourable Terence Conboy was appointed a Federal Royal Commissioner to inquire into a range of matters relating to the Australian building and construction industry. This Royal Commission was the first national review of the conduct and practices in the Australian construction industry (Cole, 2003). One of the issues before the Commissioner was ‘security of payment’ in the Australian

construction industry. In February 2003, the Commissioner presented his Final Report (comprising 23 volumes) to the Australian Governor-General. The Commissioner gave his recommendations in respect of the security of payment issue in Volume 8 of the Final Report. In summary, the Commissioner considered the security of payment issue to be one of national relevance (Brand and Uher 2010).

In the present context ‘security of payment’ is a generic term used to describe “the entitlement of contractors, subcontractors, consultants or suppliers in the contractual chain to receive payment due under the terms of their contract from the party higher in the chain” (NSW Government 1996:41). Thus, the security of payment problem is the “consistent failure in the building and construction industry to ensure that participants are paid in full and on time for the work they have done, even though they have a contractual right to be paid” (Commonwealth of Australia 2002:7). In general, the security of payment problem relates to the arbitrary devaluation, late payment and/or non-payment of progress claims. It is a persistent problem for those who perform construction work, or supply goods and services, in the construction industry. In the second reading speech, the NSW Minister for Public Works and Services, Hon. Morris Iemma (1999:103), stated:

It is all too frequently the case that small subcontractors, such as bricklayers, carpenters, electricians and plumbers, do not get paid for their work. Many of them cannot survive financially when that occurs, with severe consequences to themselves and their families...The [New South Wales] Government is determined to rid the construction industry of such totally unacceptable practices.

The tactic of principals and contractors in delaying payments, or unduly reducing the value of payments, is designed to enhance their positive cash flow at the expense of subcontractors and suppliers. To recover payments due under a contract, subcontractors and suppliers have generally relied on one or more of traditional dispute resolution processes, such as arbitration (if available under the contract) and/or litigation. However, according to Cole (2003:10):

There is evidence based on comments made by subcontractors and their industry association representatives that some subcontractors fail to take the necessary actions and remedies available to them under the law (or in contract)...Reasons for this may include the high cost and time delays in taking legal action[...]

The prohibitive costs and time delays involved in recovering payment under traditional dispute resolution processes has often led subcontractors and suppliers to simply abandon their contractual right to payment and to move onto other construction projects in order to maintain cash flow. Cash flow is the “lifeblood of the construction industry”.¹

¹ Iemma M, Member for Lakemba, Legislative Assembly, *Second Reading Speech (Hansard)*, Building and Construction Industry Security of Payment Amendment Bill (NSW), 15 November 1999 at 81; see also *Dawnays Ltd v FG Minter Ltd* (1971) 1 WLR 1205 (CA) whereby it was held that an employer was obliged to pay a sum certified by an architect. The justification for this decision was said to be that cash flow was the life blood of the building trade.

The prohibitive costs and time delays involved in recovering payment under these processes has often led subcontractors and suppliers to simply abandon their right to payment and to move onto another projects in order to maintain cash flow (Commonwealth of Australia, 2002).

1.2 Legislative response to the security of payment problem

The *Building and Construction Industry Security of Payment Act 1999* (NSW) ('the NSW Act') was introduced in an attempt to counter the security of payment problem in New South Wales ('NSW'). New South Wales was the first Australian jurisdiction to introduce this form of legislation. The object of the NSW Act "is to ensure that any person who undertakes to carry out construction work (or who undertakes to supply related goods and services) under a construction contract is entitled to receive, and is able to recover, progress payments in relation to the carrying out of that work and the supplying of those goods and services".² To achieve this objective, the NSW Act introduced new statutory rights for claimants, such as: a right to progress payments;³ a right to interest on late payments;⁴ a right to suspend work;⁵ and a right of lien.⁶ The NSW Act also renders void 'pay-when-paid' clauses in construction contracts,⁷ and the parties cannot contract out of the NSW Act.⁸

In addition, the NSW Act introduced a unique form of 'rapid adjudication' of disputes over progress payment amounts whereby an independent adjudicator makes an interim determination as to the amount of progress payment to be paid to a claimant by a respondent. Only a claimant can initiate the adjudication process, however, both parties are entitled to make submissions to the adjudicator (subject to s. 20(2B) of the Act). An adjudicator can only be appointed by an Authorised Nominating Authority (ANA) chosen by the claimant.⁹

The procedures and timeframes in relation to the adjudication process are strict and governed solely by the NSW Act. An adjudicator's determination, while not final, is binding on the parties¹⁰ until the dispute is resolved by private agreement, a court, or some other process. If the respondent does not pay the adjudicated amount by the relevant date,¹¹ the adjudicator's determination is then capable of being registered as a judgement in a court of competent jurisdiction via a relatively straightforward administrative process prescribed under the NSW Act,¹² and is enforceable accordingly.¹³ If, subsequently, a respondent applies to the court to have the judgment set aside, the respondent will not be entitled to bring a cross-claim against the

² *Building and Construction Industry Security of Payment Act* (1999) NSW s. 3.

³ *Building and Construction Industry Security of Payment Act* (1999) NSW s. 8(1).

⁴ *Building and Construction Industry Security of Payment Act* (1999) NSW s. 11(2).

⁵ *Building and Construction Industry Security of Payment Act* (1999) NSW ss. 15(2), 16(2), 24(1).

⁶ *Building and Construction Industry Security of Payment Act* (1999) NSW s. 11(3).

⁷ *Building and Construction Industry Security of Payment Act* (1999) NSW s. 12(1).

⁸ *Building and Construction Industry Security of Payment Act* (1999) NSW s. 34.

⁹ *Building and Construction Industry Security of Payment Act* (1999) NSW s. 17(3)(b).

¹⁰ *Building and Construction Industry Security of Payment Act* (1999) NSW s. 23(2).

¹¹ *Building and Construction Industry Security of Payment Act* (1999) NSW s. 23(1).

¹² See *Building and Construction Industry Security of Payment Act* (1999) NSW ss. 24, 25.

¹³ *Building and Construction Industry Security of Payment Act* (1999) NSW s. 25(1).

claimant,¹⁴ or to raise any defence in relation to matters arising under the construction contract¹⁵ or to challenge the determination by the adjudicator,¹⁶ (other than on grounds of an adjudicator's jurisdiction). In addition, the respondent must pay into court as security the unpaid portion of the adjudicated amount pending the outcome of that proceeding.¹⁷

At the time the Final Report was presented, only New South Wales and Victoria had construction industry specific security of payment legislation in operation. The Victorian Act¹⁸ was closely modelled on the NSW Act in its unamended form. However, since that time both the NSW Act and the Victorian Act have undergone significant amendments.¹⁹ This has resulted in a significant divergence in the comparative operation and effect of these Acts, notwithstanding that both maintain virtually identical objectives.

Since the Final Report was presented, Queensland,²⁰ Western Australia²¹ and the Northern Territory²² have all enacted construction industry specific security of payment legislation (hereafter referred to as 'the Qld, WA and NT Acts', respectively). The Qld Act is closely modelled on the NSW Act (as amended). The NT Act, on the other hand, is largely modelled on the WA Act, which, in turn, is largely modelled on the UK²³ and New Zealand²⁴ 'construction contracts' legislation (hereafter referred to as the 'UK and NZ Acts', respectively).

The uptake of 'security of payment' legislation in Australia continues. At the time of writing, the South Australian²⁵, Tasmanian²⁶ and Australian Capital Territory²⁷ Parliaments have passed legislation, all of which are largely modelled on the NSW Act (hereafter referred to as the 'SA, Tas and ACT Acts', respectively). The similarity of the recent legislation introduced in South Australia, Tasmania and the Australian Capital Territory to the existing legislation in New South Wales, Queensland and Victoria clarifies that there are now two models in operation in Australia – the Western Australian/Northern Territory model (based closely on the UK legislation), and the uniquely 'Australian Model' employed by the remaining States and Territory.

In addition to the UK and NZ Acts previously mentioned, Singapore²⁸ ('the Singapore Act') has in operation security of payment legislation, which is largely modelled on the

¹⁴ *Building and Construction Industry Security of Payment Act* (1999) NSW s. 25(4)(a)(i).

¹⁵ *Building and Construction Industry Security of Payment Act* (1999) NSW s. 25(4)(a)(ii).

¹⁶ *Building and Construction Industry Security of Payment Act* (1999) NSW s. 25(4)(a)(iii).

¹⁷ *Building and Construction Industry Security of Payment Act* (1999) NSW s. 25(4)(b).

¹⁸ *Building and Construction Industry Security of Payment Act* 2002 (Vic).

¹⁹ See *Building and Construction Industry Security of Payment Amendment Act* 2002 (NSW); and *Building and Construction Industry Security of Payment (Amendment) Act* 2006 (Vic).

²⁰ *Building and Construction Industry Payments Act* 2004 (Qld).

²¹ *Construction Contracts Act* 2004 (WA).

²² *Construction Contracts (Security of Payments) Act* 2004 (NT).

²³ *Housing Grants, Construction and Regeneration Act* 1996 (UK) pt. 2.

²⁴ *Construction Contracts Act* 2002 (NZ).

²⁵ *Building and Construction Industry Security of Payment Act* 2009 (SA).

²⁶ *Building and Construction Industry Security of Payment Act* 2009 (Tas).

²⁷ *Building and Construction Industry (Security of Payment) Act* 2009 (ACT).

²⁸ *Building and Construction Security of Payments Act* 2004 (Singapore).

NSW Act (as amended). Consequently, the Singapore Act fits with in the Australian Model of adjudication.

2 ‘Defined Scheme’ and ‘Non-specific Scheme’ of adjudication

In Australasia there are two different models for adjudication in the building and construction industry. The first is the Defined Scheme. This is a term coined just for this paper. It is the scheme adopted under the Australian Model. The Defined Scheme was first created by the NSW Act, and is the scheme subsequently adopted for the Vic, Qld, ACT, Tas and SA Acts. It is a Defined Scheme because it only provides for claims by one party, and only for defined money claims, namely, progress claims. Only the person who carries out construction work or provides related goods or services can make claims under the Defined Scheme and only progress claims are adjudicated.

The second is the Non-specific Scheme. This is a term coined just for this paper. The Non-specific was first created by the UK Act, and is the scheme subsequently adopted for the NZ, WA and the NT Acts. It is a Non-specific Scheme because it does not specify that only one party can make claims and it is not limited to a specific type of money claim under a contract. For example, the Non-specific Scheme adopted for the NZ Act allows either party to a construction contract to refer to adjudication any ‘dispute’.²⁹ A dispute is defined as ‘a dispute or difference that arises under a construction contract’.³⁰ But only an adjudication determination that a party is liable to pay money is enforceable.³¹ Determinations on other questions of rights or liabilities are unenforceable. Furthermore, the WA Act allows either party to a construction contract to refer to adjudication a ‘payment dispute’.³² In this case, a payment dispute is defined as a claim for money or for release of any security due to be returned by a party.³³ The NT Act follows closely the WA Act in this regard.

The Defined Scheme has some important advantages in that, because it is limited in scope, it enables rapid and relatively inexpensive adjudication of progress claims. On that basis, it is not surprising that empirical research indicates significant use of the statutory adjudication process in the NSW construction industry (Brand and Uher 2007).

However, the courts, led by the Supreme Court of NSW, have allowed progress claims to include claims for delay costs and damages.³⁴ Instead of being merely able to use the legislation to recover a progress payment for work actually carried out, some claimants have used it to recover payments on account for damages for breach of contract and payment on account for amount claims for alleged delay costs. The authors submit that this was not the intention of the legislation. Nevertheless, this judicial interpretation results in a serious imbalance because, under the Defined Scheme, only one party can refer a claim to adjudication and recover money. The Victorian legislators attempted to

²⁹ Construction Contracts Act 2002 (NZ) s. 25.

³⁰ Construction Contracts Act 2002 (NZ) s. 5.

³¹ Construction Contracts Act 2002 (NZ) s. 58.

³² Construction Contracts Act 2004 (WA) s. 25.

³³ Construction Contracts Act 2004 (WA) s. 6.

³⁴ See, for example: *Trysams Pty Ltd v Club Constructions (NSW) Pty Ltd* [2007] NSWSC 941; and *Coordinated Construction Co Pty Ltd v J.M. Hargreaves (NSW) Pty Ltd & Ors* [2006] HCATrans 9 (3 February 2006).

solve the problem by making an adjudication determination void to the extent that it includes ‘excluded amounts’.³⁵ These are defined to be claims for delay costs and damages and some variation claims.³⁶ However, this amendment does not address the basic imbalance; only one party can take advantage of adjudication. The Vic Act is still the Defined Scheme.

3 The ‘Dual Scheme’ proposal

3.1 Introduction

The Dual Scheme proposal is to retain the Defined Scheme for purely progress payment claims, (i.e., for claims for the value of work, goods or services that have actually been provided), and to adopt the Non-specific Scheme for other payment disputes.

The Dual Scheme would continue the existing legislation for ‘Progress Claims’ (redefined to exclude the broader interpretation sometimes adopted by the Courts) and would allow *either* party to make other ‘Money Claims’ and have them adjudicated.

The process of adjudication for ‘Progress Claims’ would be basically the same as now exists. However, there are likely to be more adjudication applications made because the party now precluded from claiming under the Defined Scheme would be given the opportunity to make claims under the Dual Scheme via the Money Claim extension to the Defined Scheme.

It is proposed that the Dual Scheme would retain the objective of the Defined Scheme (i.e., “to ensure that any person who undertakes to carry out construction work ... is entitled to receive, and is able to recover, progress payments in relation to the carrying out of that work”).³⁷ But the Dual Scheme would also adopt the objective of the Non-specific Scheme. For example, the Non-specific Scheme adopted by the NZ Act provides “for the speedy resolution of disputes arising under a construction contract” and to provide “remedies for the recovery of payments under a construction contract”.³⁸ Similarly, in the Non-specific Scheme adopted by the WA Act, the objective is “to provide a means for adjudicating payment disputes arising under construction contracts”.³⁹

Under the Dual Scheme, the person who carries out construction work or provides related goods or services could continue to make progress claims and have them adjudicated in exactly the same way as now occurs, but progress claims would be redefined to exclude claims for other than the value of the work actually carried out or the goods and services actually provided. However, any claims for reimbursement of alleged extra costs incurred as a consequence of delay, and claims for damages, would not be Progress Claims under the Dual Scheme. If that person wants to recover damages, for example, for delay or other breach of contract or for release of a bank

³⁵ *Building and Construction Industry Security of Payment Act 2002* (Vic) s. 23(2B).

³⁶ *Building and Construction Industry Security of Payment Act 2002* (Vic) s. 10B.

³⁷ *Building and Construction Industry Security of Payment Act* (1999) NSW s. 3.

³⁸ *Construction Contracts Act 2002* (NZ) s. 3.

³⁹ *Construction Contracts Act 2004* (WA).

guarantee, the person would have to make a separate Money Claim. The other party to the construction contract would be equally entitled to make a Money Claim for damages for delay or other breach of contract.

The abuse which the Defined Scheme was established to abolish was the practice of respondent's to a payment claim withholding payment without a reason or on account of a cross-claim yet to be decided. In the absence of compulsory rapid adjudication, notwithstanding that the respondent has the benefit of the value of work, goods or services, the respondent can withhold payment and force the claimant (i.e., the person making the payment claim) to embark upon costly litigation or arbitration to recover payment. Often the amount involved would not justify the expense of litigation or arbitration and the claimant would be left with no effective remedy.

The Defined Scheme was designed to provide a more even playing field by allowing an adjudicator to decide that the claimant is entitled to a progress payment on account pending a final decision in arbitration or litigation. Through the authors' direct involvement in adjudication in Australia, experience indicates that, after adjudication, it is very rare for either party to embark on litigation or arbitration to obtain a final decision; it is common for adjudication to be used, in effect, to deal with the final account. This is consistent with the UK experience of adjudication (Kirkham 2004).

However, the problem with the Defined Scheme is that if, at the time for a progress payment, the respondent has not obtained a final determination that the respondent is entitled to an amount in respect of a cross-claim, the respondent cannot withhold payment on account of the cross-claim. Under the Defined Scheme there is no mechanism for the respondent to initiate an adjudication to obtain a determination, on an interim basis, of the respondent's claim. Consequently, when it comes time for a progress payment, the respondent is unable to set off against the progress claim amounts which the respondent claims. There are some exceptions, but there is a basic imbalance which arises from the fact that the claimant can claim damages and the respondent cannot. Examples below will illustrate how the Dual Scheme both protects the person who has carried out work or supplied goods or services while at the same time allowing the other party an equal right to adjudication of claims.

In the Appendix, there are a number of examples of how the Dual Scheme would work in practice.

3.2 Progress Claims (the 'Defined Scheme')

Under the Dual Scheme progress claims would be made and adjudicated just as they are now. The basic difference would be that progress claims would be restricted to claims for the value of work actually carried or goods or services actually provided, including work, goods or services comprised in variations, retention moneys and return of cash security. Reasons for withholding payment would be limited to arguments as to the value of work, goods or services, including whether they are defective and the estimated cost of rectifying defects, and the set off of agreed amounts and adjudicated amounts.

For example, assume that a construction contract is for the construction of a factory for the lump sum of \$1m. Assume that retention moneys are \$10,000 and there is also a

bank guarantee provided by the contractor for \$15,000. The contractor could make progress claims in the usual way for instalments of the contract price and, in due course, refund of retention moneys. The Defined Scheme presently does not encompass claims for release of bank guarantees. It is proposed that this be included as a Money Claim, as is the case under the WA Act. If the principal has directed extra work by way of a variation, the contractor could include the value of that work in a progress claim. The contractor's progress claims would be adjudicated in the same way as occurs at present and on the same timetable.

Instead of the current endorsement: *"This is a claim under the ... Act"*,⁴⁰ a progress claim would have to have the endorsement: *"This is a Progress Claim under the ... Act"*. A Progress Claim is for an amount that will be payable after the claim is made. It is payable on the due date decided by the adjudicator.

3.3 Other Money Claims (the 'Non-specific Scheme')

The Defined Scheme legislation would be amended to allow either party to the construction contract to make money claims. The claim would have to be endorsed, *"This is a Money Claim under the ... Act"*.

The claim would be for an amount allegedly already due at the date of the claim. For example, if a progress claim is not paid on the due date then the contractor could opt to make a Money Claim for the amount over due. The contractor could make a Money Claim at any time for damages for breach of contract. For example, if the contractor claims that the principal repudiated the contract, the claimant could make a progress claim in the usual way for the value of work actually performed but the progress claim could not include damages for the loss of the bargain – that would be a Money Claim.

The principal under the construction contract referred to in the example, could make a Money Claim at any time for liquidated damages for delay or any other damages allegedly due at the date of the money claim. For example, if the principal wants to set off against future progress payments an amount allegedly due to the principal for damages for breach of the contract by the contractor, the respondent can at any time make a Money Claim against the contractor.

To avoid confusion, it is proposed that the term 'claimant' presently used in the Defined Scheme to describe the person who contracts to provide work, goods or services, would be replaced with the term 'Supplier', and the term 'respondent' would be replaced by 'Purchaser'. The term 'Supplier' would describe a person who contracts to supply work, goods or services. The term 'Purchaser' would describe the person who is or may be liable to pay the Supplier for the work, goods or services. The term 'Claimant' would then be used to describe the person who makes a progress claim or a money claim. The term 'Respondent' would be used to describe the person against whom the claim is made.

Only a Supplier could make a Progress Claim, but a Money Claim could be made by a Supplier against a Purchaser and visa-a-versa.

⁴⁰ See, for example: *Building and Construction Industry Security of Payment Act* (1999) NSW s. 13(2)(c).

The process for a Money Claim would be that the person claiming, the Claimant, would make a claim against the other party to the construction contract for an amount which is allegedly due to the Claimant at the time of the claim. The claim would be served just as a Progress Claim is served but, unlike a progress claim, a Money Claim could be served at any time. This is because a Money Claim is for an amount [whether a debt or damages] allegedly due when the claim is made. The Respondent would have a prescribed period, e.g., 10 business days to provide a 'Defence'. The term 'Defence' is recommended to avoid confusion with a 'payment schedule'. A payment schedule is provided in answer to a Progress Claim.

If no Defence is provided within the prescribed time, the Claimant would have to give a reminder, similar to a s.17(2) notice under the NSW Act, and if no Defence was received within, say, another 10 business days, the Claimant could initiate an adjudication by making an adjudication application to an Authorised Nominating Authority. The claim would be adjudicated in the usual way. The adjudicator would decide whether the Claimant is entitled to the amount demanded, or any portion. The adjudicator would not decide a due date for payment, as happens with a Progress Claim. Either the amount is due at the date of the claim or it is not properly the subject of a Money Claim.

The Money Claim could include a claim for interest up to the date of the adjudicator's determination. In that event, to calculate the interest, the adjudicator would have to decide when, in the past, the amount claimed fell due. The Respondent who fails to serve a Defence within time would be precluded from making an adjudication response.

The Defence could be simply be a defence or it could be a cross-claim or both a defence and cross-claim. With the above timetable, the Respondent would have at least 21 business days to prepare and serve a Defence. If the Defence is simply a defence, the Claimant could immediately make an adjudication application. If the Defence includes a cross-claim and the Claimant wishes to defend the cross-claim, the Claimant would have a prescribed period in which to lodge a response to the cross-claim. It is suggested that this response be called a 'Rejoinder'. In an adjudication, the parties would be limited to submissions in support of their respective Money Claims, Defence (including cross-claim, if any) and Rejoinder. If the Claimant does not make an adjudication application, the Respondent would have the right to make an adjudication application in respect of the Respondent's cross-claim.

This may sound complicated but it simply mirrors the usual process of claim, defence, cross-claim and rejoinder familiar to arbitrators and courts. The Claimant is the equivalent of the plaintiff in litigation and the Respondent is the equivalent of the defendant. The adjudicator could decide that either party is indebted to the other. The adjudicated amount would have to be paid within a prescribed period. The successful party would be entitled to apply for an adjudication certificate and register it as a judgment. The amount paid would be a payment on account and in other proceedings it could be decided that the amount must be repaid with interest.

Here is a simple example of the Dual Scheme. Assume that a Purchaser claims that at 30 June 2010 the Supplier is late in achieving practical completion and liable for \$10,000 for liquidated damages. The Purchaser [now the Claimant] can make a Money Claim. The Supplier [now the Respondent to the Money Claim] could make a cross-claim for an amount that the Supplier claims is then due. The cross-claim could be for a progress payment that is already overdue but not for a Progress Claim. That is a different matter. Assume that the Respondent's Defence is simply that the Claimant is not entitled to any liquidated damages. Assume that the Defence is served within 10 business days. The Claimant can make an adjudication application immediately. The adjudicator would have a prescribed time to make a decision on the issue. Assume that the adjudicator decides on 31 July 2010 that the Claimant [Purchaser] is entitled to \$10,000.

Further assume that on 1 August 2010 the Supplier makes a Progress Claim for \$50,000. The Supplier would have to pay the Purchaser the \$10,000 within the time allowed by the Act or it would be set off against the amount of the progress payment. The Progress Claim would be decided by an adjudicator in the usual way but the adjudicator would not be able to override the previous adjudicator's decision on the \$10,000 (see further discussion on 'issue estoppel' below).

Money Claims should be confined to amounts due under or for breach of contract. It would not be appropriate to bring in claims under the *Trade Practices Act 1974* (Cth) (for example, a claim for misleading and deceptive conduct arising out of the formation of a construction contract),⁴¹ or claims in tort or restitution, since these would be claims made outside the construction contract.

Furthermore, because Money Claims may include cross-claims and are, therefore, likely to be more difficult to decide than Progress Claims, it is proposed that an adjudicator should have limited power to unilaterally extend the time for making a decision on a Money Claim.

4 Other issues relating to the 'Dual Scheme' proposal

4.1 Other Money Claims (the 'Non-specific Scheme')

The term 'issue estoppel'⁴² has been chosen to describe not only that a previous adjudicator's decision should not be overruled by a subsequent adjudication decision but the need for a more general bar on adjudicator shopping and a multiplicity of Money Claims.

It is proposed that in a Money Claim the Claimant (whether the Supplier or the Purchaser) should include all amounts then said to be due from the Claimant to the

⁴¹ See, for example: *Peninsula Balmain Pty Limited v Abigroup Contractors Pty Limited* [2002] NSWCA 211.

⁴² For a judicial discussion regarding Issue Estoppel and its application to adjudicator's determinations made under the *Building and Construction Industry Security of Payment Act* (1999) NSW see, for example: *Dualcorp Pty Ltd v Remo Constructions Pty Ltd* [2009] NSWCA 69; *Perform (NSW) Pty Ltd v Mev-Aus Pty Ltd t/as Novatec Construction Systems* [2009] NSWSC 416; *The University of Sydney v Cadence Australia Pty Limited* [2009] NSWSC 635; and *Watpac Constructions v Austin Corp* [2010] NSWSC 168.

Supplier under or for breach of contract. What is required is something along the lines of the rule in *Henderson v Henderson*,⁴³ which (Barnett 2001:24):

[...] requires the parties, when a matter becomes the subject of litigation between them in a court of competent jurisdiction, to bring their whole case before the court so that all aspects of it may be finally decided once and for all [subject to any appeal]. In the absence of special circumstances, the parties cannot return to the court to advance arguments, claims or defences which they could have put forward for a decision on the first occasion but failed to raise.

Therefore, applying the rule in *Henderson* to adjudication, a Claimant should not be able to include in any subsequent Money Claim any amount that was allegedly due to the Claimant at the date of an earlier Money Claim by the Claimant but was not included in that earlier Money Claim.

Take the earlier example; the Purchaser makes a Money Claim at 30 June 2010 for \$10,000 for liquidated damages. That is the only amount said to be due from the Supplier to the Purchaser at 30 June 2010. In a subsequent money claim, the Purchaser could not claim for other damages allegedly due at 30 June 2010. If an adjudicator has decided that the Purchaser was entitled to only \$7,500 for liquidated damages at 30 June 2010, any subsequent adjudicator could not decide differently. But a subsequent adjudicator could decide that additional liquidated damages were due for periods after 30 June 2010.

It is important that the Purchaser is not able to make separate Money Claims for every defective brick.⁴⁴ Consequently, if the Purchaser claims in a Money Claim dated 30 June 2010 that the Purchaser is entitled to \$17,000 for the cost of rectifying defects in the Supplier's work, that amount should cover all defects rectified at 30 June 2010. Similarly, if the Supplier makes a Money Claim at 30 June 2010 for \$18,000 for delay costs, that claim should cover all delays and all delay costs incurred at 30 June 2010.

While in litigation and arbitration the principle of issue estoppel is a bar to making further claims, in adjudication, the principle would be limited to making further Money Claims in respect of entitlements that existed at the date of an earlier Money Claim but were not included in the earlier Money Claim. Therefore, under the proposed scheme, when the first adjudicator decides that at a certain date the Respondent was not entitled to liquidated damages or an amount for defective work, the Respondent would not be able to again claim for amounts allegedly due for liquidated damages and defective work at the certain date. The Respondent could make a claim for liquidated damages allegedly accruing after the certain date or damages for defective work which did not exist until after the certain date.

⁴³ (1843) 3 Hare 100; 67 ER 313 cited Barnett P (2001) 'Res judicata, Estoppel and Foreign Judgement', Oxford University Press, New York, p. 24.

⁴⁴ *Conquer v Boot* [1928] 2 KB 336 per Sankey LJ at 342 cited in Davenport P (2006) 'Construction Claims' (2nd ed.), Federation Press, Sydney, p. 226.

4.2 Other Money Claims (the ‘Non-specific Scheme’)

It is important that any legislation allows a claimant to include in a Money Claim a claim for release of a bank guarantee or portion thereof. The construction contract should provide when the bank guarantee is due for release. Assume that it is 28 days after the end of the defects liability period and the guarantee is for \$15,000. If the Purchaser wants to claim an entitlement to withhold release of the guarantee on account of defective work for which the Purchaser claims \$25,000, the Purchaser should make a Money Claim for \$25,000 and have that adjudicated before the time for release of the bank guarantee. If the adjudicator decides that the Purchaser is only entitled to \$5,000, the adjudicator could decide that the Purchaser must release \$10,000 of the amount of the guarantee. A guarantee can be released in part.

There is need for legislation to bar the calling up of a bank guarantee before there is a final decision or a decision in an adjudication that there is a debt against which the proceeds of the bank guarantee can be applied.

5 Concluding remarks

The Dual Scheme would require legislation. It would provide for Progress Claims, Payment Schedules, liability in the absence of a payment schedule, and adjudication just as now occurs in the jurisdiction operating under the Australian Model.

The Dual Scheme would also provide that either party to a construction contract can make a Money Claim. That is a claim for amounts allegedly already due as a debt or damages under or for breach of the construction contract. That is the Non-specific Scheme that exists under the NZ, WA and NT Acts.

A Money Claim would be adjudicated similarly to Progress Claim but with a different timetable and allowance for a cross-claims and a rejoinder.

A Progress Claim could only be made on or from a reference date by a Supplier (i.e., the party to a construction contract who contracts to supply work, goods or services to another person, the Purchaser). A Money Claim could be made at any time by either party.

The essence of the Dual Scheme is that a Supplier should be able to recover progress payments for the value (taking into account defects) of work goods or services actually supplied without deduction of amounts for cross-claims which have not yet been quantified in adjudication or in final proceedings. Absent defective work, the only reason that a Purchaser could have for not paying the unpaid value of work, goods or services the benefit of which the Purchaser has received, is that the Purchaser is entitled to set off an amount already determined to be due from the Supplier to the Purchaser.

6 References

- Brand, M. C. and Uher, T. E. (2010) 'Follow-up empirical study of the performance of the New South Wales construction industry security of payment legislation', *International Journal of Law in the Built Environment*, **2**(1), pp. 7-25.
- Brand, M. C. and Uher, T. E. (2007) 'An examination of the performance of statutory adjudication in New South Wales after 2000 determinations', proceedings of *CME25 Construction Management and Economics: Past, present and future*, University of Reading, UK, 16-18 July, CD ROM, paper 0033.
- Barnett, P. (2001) *Res judicata, Estoppel and Foreign Judgement*, Oxford University Press, New York.
- Cole, T. R. H. (2003) 'Final Report of the Royal Commission into the Building and Construction Industry', vol. 1, p. 4.
www.royalcombeci.gov.au/docs/finalreport/V01Summary_PressFinal.pdf, viewed: 24 May 2008.
- Commonwealth of Australia (2002) Royal Commission into the Building and Construction Industry: Security of Payment in the Building and Construction Industry, Discussion Paper 12, Canberra.
- Davenport, P. (2006) *Construction Claims*, 2nd ed., Federation Press, Sydney.
- Iemma, M. (1999) Building and Construction Industry Security of Payment Bill (No. 2), *Second Reading Speech*, Parliamentary Hansard, Legislative Assembly, Parliament of New South Wales, 8 September, p. 103-8.
- Kirkham, F. (2004) 'Tomorrow and Tomorrow and Tomorrow: Keynote Address to the Adjudication Society Conference' *Construction Law Journal*, **20**(4), 189-96.
- Further reading:*
- Davenport, P. (2007) 'A Proposal for a Dual Process of Adjudication', *Australian Construction Law Newsletter*, (115), pp. 12-24.
- Davenport, P. (2004) *Adjudication in the building industry*, 2nd ed., The Federation Press, Sydney.
- Uher, T. E. and Brand M. C. (2008) 'Claimants' view of adjudication in New South Wales', *Journal of Engineering, Construction and Architectural Management*, **15**(5), pp. 470-484.
- Uher, T. E. and Brand M. C. (2007) 'A comparative analysis of the operation of compulsory rapid adjudication in New South Wales and New Zealand', *Construction Management and Economics — Special Issue on Construction Law and Management*, **25**(7-9), pp. 765-75.

Uher, T. E. and Brand M. C. (2005) 'Analysis of adjudication determinations made under security of payment legislation in New South Wales', *International Journal of Project Management*, **23**(6), pp. 474-82.

Appendix – Examples

Example 1: Progress claim and no payment schedule

This situation is very common. The Supplier claims an amount as a progress payment or a final payment or for return of retention. If the Purchaser fails to provide a payment schedule within 10 business days the Supplier can apply to a Court for summary judgment for the claimed amount⁴⁵ or elect to go to adjudication and give the Purchaser another opportunity to provide a payment schedule.

Assuming that the Supplier elects to go to adjudication and no payment schedule is provided within time, the adjudicator would decide the amount of the progress payment in the usual way.

The difference under the Dual Scheme would be that if the Progress Claim included a claim for damages, the adjudicator would not include damages in the amount of the progress payment. Similarly, if the Supplier applied to a Court for summary judgment, the Court could refuse to enter summary judgment (or summary judgment, if entered, could be set aside) if the Purchaser could show that the progress claim was not just a Progress Claim but includes a claim for damages. This would avoid the situation that occurred in the *Walter Construction* case where the claimant recovered a summary judgment which included amounts which could not properly be included in a progress claim.

It seems that in the majority of cases where there is no payment schedule Suppliers opt for adjudication rather than summary judgment. Consequently, the change which the dual scheme would make would have minimal impact. At the time of the *Walter Construction* case, the option to apply for adjudication rather than summary judgment did not exist.

2. Payment schedule disputes value of work

In many instances, the Purchaser's only reason for withholding payment is that the work does not have the value claimed (i.e., the Supplier has overvalued it having regard to the contract price and the extent of work carried out at the reference date) or that there is defective work and the Purchaser is entitled to have the estimated cost of rectifying defective work taken into account in calculating the amount of the progress payment.

In this situation there would be no difference between the Dual Scheme and the Defined Scheme. The difference between the Dual Scheme and the Non-specific scheme would

⁴⁵ See, for example: *Walter Construction Group v CPL (Surry Hills)* [2003] NSWSC 266 ('*Walter Construction* case')

be in detail not in substance; terminology and times for making claims would be different.

3. Payment schedule asserts that payment claim includes amounts that cannot be included in a progress payment

Sometimes in a payment schedule the Purchaser asserts that under the terms of the construction contract or the legislation the Supplier is not entitled to a progress payment on account of some or all of the amount claimed because the claim (or part) is for damages or for extras (particularly delay costs) for which there is no entitlement or the entitlement is barred.⁴⁶

This situation is common. Under the Defined Scheme it has given rise to the most problems and the most contentious litigation. It is the problem which the Victorian 2006 amendment attempts to avoid by creating the concept of “excluded amounts”. The term “ambush claim” has been used to describe claims for damages and delay that are included in progress claims. It is this situation which has generated, quite legitimately, the most criticism of the Defined Scheme.

Under the Dual Scheme the entitlement to a progress payment would be confined to a claim for the value of work actually carried out and goods and services actually provided. If, in the payment schedule, the Purchaser identifies other claims, then in the calculation of the progress payment due the adjudicator would not include any amount for these claims.

The Supplier could pursue by way of a Money Claim the delay claims or other damages claims which cannot be made in a Progress Claim.

4. Payment schedule asserts that a claim is barred

In a payment schedule, the Purchaser often relies upon a time bar clause or other clause in a contract to bar an entitlement. The authors recommend that, in a contract of adhesion, unfair time bar or other barring clauses be rendered void just as “pay when paid” clauses have been rendered void in each of the Acts referred to above.

However, assuming that such a barring clause is valid, then under the Dual Scheme the situation would be no different to that now existing under both the Defined Scheme and the Non-specific Scheme. An entitlement could be barred by a time bar clause or other barring clause that did not contravene the “no contracting out” provisions of the Act.

5. Payment schedule includes a cross-claim

It is common for a Purchaser to withhold payment on account of alleged set offs for “back charges”. These are claims by the Purchaser that the Supplier is liable to the purchaser for liquidated damages or other damages. The alleged damages may include the cost incurred by the Purchaser in allegedly rectifying defective or unfinished work.

⁴⁶ Examples include: *Minister for Commerce v Contrax Plumbing* [2005] NSWCA 142; *Coordinated Construction v Hargreaves* [2005] NSWCA 228; and *Coordinated Construction v Climatech* [2005] NSWCA 229.

They may include alleged damages consequent upon termination of the contract before completion. They may even include debts allegedly owed under another contract or set offs on account of claims by a third party against the Purchaser. What these back charges have in common is that they are all for amounts for which liability or quantum or both have yet to be finally decided by a court, tribunal or arbitrator or under a dispute resolution provision of the construction contract.

Where, in a Defined Scheme, the amount of a progress payment is to be calculated under the terms of the Act,⁴⁷ the Purchaser is not entitled to set off claims for back charges. The problem of deciding a set off claim for back charges only arises where the construction contract has a provision for set off against progress payments of amounts claimed by the Purchaser.

There are differences in the approach taken by adjudicators. Some take the approach that only a debt (owed to the Purchaser) that is admitted by the Supplier or has been decided by a court or tribunal or in arbitration or has been created under a dispute resolution clause in the contract can be set off against progress payments. Other adjudicators will decide disputed issues of liability for and quantum of the back charges. Sometimes this involves deciding what extensions of time the Supplier was entitled to and whether the superintendent acted fairly.⁴⁸

A legitimate criticism of the Defined Scheme is that the fast track system was designed for calculating the amount of a progress payment and was not intended to cover the more difficult task of deciding damages claims. It was not designed for deciding claims for extensions of time or for deciding whether breaches of contract had occurred and, if so, the quantum of damages. On the other hand, the Non-specific Scheme was designed for that purpose. That is the reason for creating a Dual Scheme.

The major difference between the Dual Scheme and the Defined Scheme would be that the Purchaser would not be able to raise a back charge as a reason for withholding payment of a progress payment, or set off against a progress payment an amount for a back charge, unless liability for and the quantum of the Purchaser's entitlement had been admitted by the Supplier or finally decided in litigation or arbitration or under a dispute resolution clause or had already been decided in an adjudication of a Money Claim.

This would simplify adjudication of a progress claim yet allow a Purchaser a right to pursue claims against the Supplier in adjudication of a Money Claim. Immediately a party to a construction contract considers that the other party has breached the construction contract, that party can make a Money Claim against the other party. The other party has an equal right to respond with a cross-claim. This is the Non-specific scheme. It is a fair scheme. The shortcoming in the Non-specific Scheme is that it does not provide a fast track process for recovery of a progress payment. The Dual Scheme combines the advantages of both Non-specific Scheme and the Defined Scheme.

⁴⁷ See, for example: *Building and Construction Industry Security of Payment Act 1999* (NSW) s. 9(b).

⁴⁸ The problem is illustrated by *John Holland v RTA* [2007] NSWCA 19.

To summarise, in an adjudication of a Progress Claim, the adjudicator would not decide any cross-claims but would decide issues about the value of work, goods and services, defects and the estimated cost of rectifying defects.

6. A Supplier claims delay costs

In larger construction contracts, the Supplier frequently claims that the Purchaser delayed the Supplier and the Supplier is entitled to delay costs. Under the Dual Scheme, the claim for delay costs could not legitimately be included in Progress Claim. At any time the Supplier could make a Money Claim for delay costs allegedly due up to the date of the Money Claim. The other party could raise a defence or cross-claim or both. The Supplier could concurrently make a Progress Claim and a Money Claim but only the latter could include a claim for delay costs. There is no necessity for the claims to be decided by the same adjudicator or at the same time.

At present, delay cost claims tend to be made as an ambit claim at the end of a project. This is because, as arbitration and litigation work, a party is effectively forced to leave all claims for damages until the end of the project. The right to have each claim for delay costs adjudicated immediately after the delay occurs would avoid the need for an ambit claim. Given this right, a clause in the contract could effectively and not unfairly bar the ambit claim for delay costs.

--End--

Comparative analysis of aspects of damages: consequential loss and classification

Maree Chetwin¹

¹ College of Business & Economics, University of Canterbury, Christchurch, New Zealand

Email: maree.chetwin@canterbury.ac.nz

Abstract

Purpose This paper critically reviews the judicial approach to aspects of contract damages in England and Wales, Australia and New Zealand.

Methodology The paper is confined to damages issues with particular emphasis on agreed remedies and will discuss leading judgments of the three jurisdictions and academic commentary.

Rationale There are a variety of agreed remedies, the main one being the agreed price. Others include liquidated damages and exception clauses which are the focus of this paper. The meaning of loss must be determined before any remoteness issues are addressed. In the case of liquidated damages, do the other limitations on damages play a part? What is the meaning of the term ‘consequential loss’ which frequently appears in exception clauses? This was at issue in *Croudace Construction Ltd v Cawoods Concrete Products Ltd*¹, *Environmental Systems Pty Ltd v Peerless Holdings Pty Ltd*² and *Oceania Furniture Ltd v Debonaire Products Ltd*³. Fuller and Perdue⁴ classified damages interests as expectation, reliance and restitution. While their terminology has had a substantial influence on contract damages, the classification is no longer appropriate. *Attorney General v Blake*⁵ has changed the face of contract damages and a claim is now possible on a restitutionary basis, but is it likely to result in a change to agreed damages clauses?

Conclusions The incorporation of agreed remedies facilitates the predictability of the outcomes of contracts. The assessments of damages in the jurisdictions discussed follow common precedents, but in some areas, such as the meaning of ‘consequential loss’ in exclusion clauses, the judgments indicate a difference in approach. It is vital that what is in the contract is clearly stated and intended by both parties.

Keywords Damages, Classification, Consequential, Liquidated, Exclusion Clauses.

1 Introduction

Parties entering a contract are entitled to create mutual rights and duties by agreement. “What the parties have agreed should normally be upheld. Any other approach will lead to undesirable uncertainty especially in commercial contracts”.⁶ If there is a dispute between the parties as to what is covered and the matter turns on the wording of the contract, there will be variations between jurisdictions depending on the issue in dispute. There is a wide variety of agreed remedies and the main types relate to financial payments. Despite the general principle of freedom of contract there are various rules that involve a balancing of values, and relief is given on the basis of fairness. This is evident in relation to the penalty clause jurisdiction. Of major importance in commercial contracts are liability provisions which can be contentious. An exclusion clause limits a party’s liability and for that reason it is likely to be carefully negotiated. However, in

¹ [1978] 2 Lloyd’s Rep 55 (CA).

² [2008] VCSA 26 (CA).

³ HC, Wellington, CIV-2008-485-001701, 27 August 2009, Clifford J.

⁴ Lon Fuller and William Perdue, “The Reliance Interest in Contract Damages” (1936 – 37) 46 Yale LJ 52 and 373.

⁵ [2001] 1 AC 268 (HL).

⁶ *Philips Hong Kong Ltd v The AG of Hong Kong* (1993) 61 BLR 49 (PC) 59 (Lord Woolf).

most commercial contracts “the value of protecting the weak, the foolish, and the thoughtless from imposition and oppression”⁷ is generally irrelevant. This paper considers aspects of the efficacy of these agreed remedies in protecting a party’s contract interests in England and Wales, Australia and New Zealand.

2 Protected interests in contract law and assessing loss

Fuller and Perdue’s article has been extremely influential and their terminology continues to be utilised by the Courts⁸ but it is not without its critics⁹. Three principal purposes which may be pursued in contract damages are distinguished:

Following a breach of contract the innocent party may have:

- (i) a *restitution interest*, namely the right to restoration of a valuable benefit conferred on the other party, the object being to prevent unjust enrichment;
- (ii) a *reliance interest*, namely the right to compensation for loss due to steps taken by the innocent party in reliance upon the existence of the contract, the object being to restore the innocent party to the position which he or she would have occupied had the contract not been made; and/or
- (iii) an *expectation interest*, namely the right to compensation for loss of the bargain, the object being to financially restore the innocent party to the position which he or she would have occupied had the contract been performed.¹⁰

In *C&P Haulage v Middleton*¹¹ the article was cited in relation to the reliance interest, which is applicable in limited situations. It is best viewed as protecting the expectation interest.¹² Rowan¹³ discussing the contractual performance interest calls it “a label now used in place of the more traditional ‘expectation interest’”. However, a search through the cases in the three jurisdictions does not confirm that the judiciary has replaced the label ‘expectation’ with ‘performance’. Certainly there has been much academic comment as to the lack of protection of the performance interest. The cases demonstrate that where appropriate the courts have paid due regard to the performance interest. It cannot be said that the performance interest has replaced the expectation interest.

The two main methods of assessing the plaintiff’s loss can be described as a cost of cure award and a difference or diminution of value award. They are particularly relevant for building contracts, but they will apply to any type of contract. In the case of cost of cure, the plaintiff is awarded the cost of curing the breach. Diminution of value is the difference in value between what was contracted and what was supplied. If there is no possible replacement, this may be the only

⁷ S M Waddams “Unconscionability in contracts” (1976) 39 (4) MLR 369.

⁸ See *Paper Reclaim Ltd v Aotearoa International Ltd* [2006] 3 NZLR 188 (CA) 223; *Bashall & Anor v Meihana* 10/09/2008, Gendall J, HC Wellington, Civ 2008-485-00742.

⁹ See D McLauchlan, “Reliance Damages for Breach of Contract” [2007] NZ Law Review 417; R Craswell “Against Fuller and Perdue” (2000) 67 U Chi L Rev 99; M B Kelly, “The Phantom Reliance in Contract Damages” (1992) Wis L R 1755 and D W Barnes, “The Net Expectation Interest in Contract Damages” (1999) 48 Emory L J 1137.

¹⁰ Fuller and Perdue (n 4).

¹¹ [1983] 1 WLR 1461, [1983] 3 All ER 94, [1983] EWCA Civ 5(CA).

¹² See Andrew Burrows, *Remedies for Torts and Breach of Contract* (3rd ed OUP, Great Britain 2004) 65.

¹³ Solene Rowan “For the recognition of remedial terms agreed inter partes” (2010) 126 LQR 448.

available measure. The recent Australian High Court case, *Tabcorp Holdings Ltd v Bowen Investments Pty Ltd*¹⁴ highlights the importance of performance of contracts, which is a feature of the expectation interest. In that case, a tenant with “contumelious disregard for the landlord’s rights”¹⁵ breached a leasehold covenant and destroyed the foyer and replaced it. The damages in the lower Court were AUS\$34,820, being the difference in value of the property with the old foyer and its value with the new foyer. On appeal, the damages were assessed on the cost of cure basis by the full Federal Court. This resulted in an increase to AUS\$1.38m being the estimated cost of reinstatement of the foyer to its former condition (\$580,000) and loss of rental income during the four month period in which such works would be carried out (\$800,000). The High Court of Australia dismissed the appeal and held that the appropriate measure was cost of cure. The Court reiterated the “ruling principle” with respect to damages as stated by Parke B in *Robinson v Harman*¹⁶. The Court expressed agreement with Oliver J’s statement in *Radford v De Froberville*¹⁷ that the words ‘the same situation, with respect to damages, as if the contract had been performed’ do not mean ‘as good a *financial* position as if the contract had been performed’.”¹⁸ *Ruxley Electronics and Construction Ltd v Forsyth*¹⁹ was distinguished on its exceptional facts. The High Court of Australia referred to the facts:

‘The House saw the following matters as indicating that the cost of reconstruction was not recoverable: ‘The trial judge made the following findings which are relevant to this appeal: (1) the pool as constructed was perfectly safe to dive into; (2) there was no evidence that the shortfall in depth had decreased the value of the pool; (3) the only practicable method of achieving a pool of the required depth would be to demolish the existing pool and reconstruct a new one at a cost of £21,560; (4) he was not satisfied that the respondent intended to build a new pool at such a cost; (5) in addition such cost would be wholly disproportionate to the disadvantage of having a pool of a depth of only 6ft as opposed to 7ft 6in and it would therefore be unreasonable to carry out the works; and (6) that the respondent was entitled to damages for loss of amenity in the sum of £2,500.’²⁰

In *Ruxley Electronics*, the plaintiff was awarded £2,500 amenity damages. The case has been criticised as it fails to protect the performance interest. However, the above facts support the view that in the particular circumstances it was not reasonable for the plaintiff to require complete replacement or reinstatement.

Rowan²¹ argues that the force and logic of freedom of contract should extend the available agreed remedies. The reason to advocate cost of cure damages clauses, restitution, penalty and specific relief clauses is to provide adequate protection to the performance interest. However, reasonableness is a vital ingredient of contract relief and it would be unlikely that the courts would agree that there was such a need. This

¹⁴ (2009) 83 A LJR 390 (HCA).

¹⁵ (2009) 83 A LJR 390 (HCA) [4].

¹⁶ (1848) 1 Exch 850 855.

¹⁷ [1977] 1 WLR 1262 1273.

¹⁸ (2009) 83 ALJR 390 (HCA) 394.

¹⁹ [1996] AC 344 (HL).

²⁰ (2009) 83 ALJR 390 [18] (French CJ, Gummow, Heydon, Crennan and Kiefel JJ).

²¹ Rowan (n 13) 448.

is evident from the New Zealand Court of Appeal discussion in relation to exemplary damages in contract law:

There is certainly no need for exemplary damages to fill any hole in the range of compensatory damages in the contract field. Contractual remedies now available in appropriate cases include expectation damages, reliance damages, and damages for non-pecuniary loss, mental distress, disappointment and loss of amenity. It has even been suggested that a Court could order an account of profits as a contractual remedy (*Attorney-General v Blake* [2001] 1 AC 268). In addition, in appropriate cases, indemnity costs may be available for improper conduct in the course of litigation. And, of course, also within the Court's armoury are the non-monetary remedies of injunction and specific performance. There is no reason in principle to add yet another remedy to the above list that would give a contracting party a windfall profit over and above that which it had bargained for.²²

The New Zealand Court of Appeal referred to overseas authority and stated that it was "particularly influenced by the detailed reports undertaken by the law commissions in the England and Wales and Ireland"²³.

In the House of Lords in *Attorney-General v Blake*²⁴ it was held that the Crown was not entitled to restitutionary damages but to an account of profits. It was clear that these are both exceptional remedies. In the case of a restitution or account of profits clause it would be impossible to assess the profit likely to be made and if all the profits were stipulated then that would be a penalty. It seems likely that the courts would regard a restitution clause as penal as "agreed remedies will to some extent be assessed in line with the judicial remedies available".²⁵

3 Agreed damages

In this section it is proposed to consider the approach of the three jurisdictions to the agreed remedies of liquidated damages and exclusion clauses. The interpretation of a contract will be particularly important when the contract excludes or limits liability. The courts have developed rules which at times may not reflect a possible objective interpretation but produce a fair result.

3.1 Liquidated damages

At the time of contracting, the parties may make provision in their contract as to what damages shall be recoverable in the event of breach. If there is a dispute and the damages are classified 'liquidated damages' then the clause will be enforced, but it will not be enforced if it is held to be a 'penalty clause'.

The applicable principles are those summarised by Lord Dunedin in *Dunlop Pneumatic Tyre Co Ltd v New Garage and Motor Co Ltd*:

²² [2006] 3 NZLR 188 (CA) 223 (Chambers J).

²³ [2006] 3 NZLR 188 (CA) 223 (Chambers J).

²⁴ [2001] 1 AC 268 (HL) 282.

²⁵ Burrows (n 12) 445.

- (a) It will be held to be a penalty if the sum stipulated for is extravagant and unconscionable in amount in comparison with the greatest loss that could conceivably be proved to have followed from the breach ...
- (b) It will be held to be a penalty if the breach consists only in not paying a sum of money, and the sum stipulated is a sum greater than the sum which ought to have been paid ...
- (c) There is a presumption (but no more) that it is a penalty when a single lump sum is made payable . . . on the occurrence of one or more or all of several events, some of which would occasion serious and others but trifling damage . . .
- (d) It is no obstacle to the sum stipulated being a genuine pre-estimate of damage, that the consequences of the breach are such as to make precise pre-estimation almost an impossibility.²⁶

The task of the court is one of construction and it may take into consideration the “inherent circumstances of each particular contract, judged as at the time of the making of the contract, not at the time of the breach...”²⁷ The law as to penalties is well settled but recent formulations have adopted different terminology “to avoid the phrase ‘in terrorem’:

...whether a provision is to be treated as a penalty is a matter of construction to be resolved by asking whether at the time the contract was entered into the predominant contractual function of the provision was to deter a party from breaking the contract or to compensate the innocent party for breach. That the contractual function is deterrent rather than compensatory can be deduced by comparing the amount that would be payable on breach with the loss that might be sustained if breach occurred.²⁸

There is a predisposition towards upholding commercial contracts to provide certainty, as was emphasised by Lady Justice Arden in *Murray v Leisureplay plc*²⁹. Arden LJ considered that “A pre-estimate is genuine if it is not unreasonable in all the circumstances of the case.”³⁰ Buxton LJ disagreed, as this overlooked one of Lord Dunedin’s principal tests³¹ and introduced a rigid and inflexible element into what should be a broad and general question”³². A clause is not a penalty unless it is ‘extravagant or unconscionable’ but neither the literal wording of that test nor the spirit of it applied to this case. Buxton LJ adopted *Dunlop’s* test as recast in modern terms by Colman J in *Lordsvale Finance plc v Bank of Zambia*³³ (see above quote) and approved by Mance LJ in *Cine Bes Filmcilik Ve Yapim Click v United International Pictures*³⁴. Clarke LJ preferred the broader approach of Buxton LJ. The English Court of Appeal held that the clause, which provided for a full year’s salary in the event of a wrongful

²⁶ [1915] AC 79 87.

²⁷ [1915] AC 79 (HL) 87.

²⁸ *Lordsvale Finance plc v Bank of Zambia* [1966] QB 752 762 (Colman J).

²⁹ [2005] EWCA Civ 963 (CA (Civ Div)).

³⁰ [2005] EWCA Civ 963 (CA (Civ Div)). [55] (Arden LJ).

³¹ *Dunlop Pneumatic Tyre Co Ltd v New Garage and Motor Co Ltd* [1915] A C 79 87. “It will be held to be a penalty if the sum stipulated for is extravagant and unconscionable in amount in comparison with the greatest loss that could conceivably be proved to have followed from the breach.” See also *Ringrow Pty Ltd v BP Australia Pty Ltd* (2005) 80 ALJR 219 (HCA) [32]: “It must be ‘out of all proportion’ and *Philips Hong Kong Ltd v Attorney-General of Hong Kong* (1993) 61 BLR 41 (PC) 57.

³² [2005] EWCA (CA (Civ Div)) 963 [114].

³³ [1996] Q B 752 (QBD) 762 .

³⁴ [2003] EWCA Civ 1669 (CA).

dismissal, was not a penalty. The amount payable greatly exceeded what would be payable at common law but other justifications such as market expectations, restraint of trade clause, a clean break, avoiding lengthy litigation and damaging publicity were relevant. The clause while “generous” was not ‘unconscionable’.

The courts have in recent times adopted a more liberal attitude to the enforcement of liquidated damages clauses where the test of commercial justification is satisfied. In *Alfred McAlpine Capital Projects Ltd v Tilebox Ltd*³⁵ Jackson J referred to the fact that the “rule about penalties is an anomaly within the law of contract” and that there is a strong predisposition to uphold the contractual term “in the case of commercial contracts freely entered into between parties of comparable bargaining power”.

The emphasis on ‘extravagant and unconscionable’ suggests the importance of the unfairness, which is also evident from the Australian and New Zealand decisions. The High Court of Australia in *Ringrow Pty Ltd v BP Australia Pty Ltd*³⁶ confirmed that the starting point of any discussion is Lord Dunedin’s speech in *Dunlop Pneumatic Tyre Co Ltd v New Garage and Motor Co Ltd*³⁷. The Court stated:

Exceptions from that freedom of contract require good reason to attract judicial intervention to set aside the bargains upon which parties of full capacity have agreed. That is why the law on penalties is, and is expressed to be, an exception from the general rule. It is why it is expressed in exceptional language. It explains why the propounded penalty must be judged ‘extravagant and unconscionable in amount’. It is not enough that it should be lacking in proportion. It must be ‘out of all proportion’. It would therefore be a reversal of longstanding authority to substitute a test expressed in terms of mere disproportionality.³⁸

Lord Woolf in *Philips Hong Kong Ltd v Attorney General of Hong Kong*³⁹, delivering the opinion of the Privy Council, approved the words of Dickson J in the Supreme Court of Canada in *Elsey v J G Collins Insurance Agencies Ltd*⁴⁰, namely:

It is now evident that the power to strike down a penalty clause is a blatant interference with freedom of contract and is designed for the sole purpose of providing relief against oppression for the party having to pay the stipulated sum. It has no place where there is no oppression.

The above words of Dickson J were also approved by the Australian High Court in *Esanda Finance Corporation Ltd v Plessnig*⁴¹. The New Zealand Court of Appeal in *Amaltal Corporation Ltd v Maruha (NZ) Corporation Ltd*⁴² also cited the dicta and concluded that the power to strike down a clause as a penalty, which was developed in the public interest, is “concerned with relief against oppression or unconscionable behaviour by a contracting party”⁴³.

The courts refuse to enforce penalty clauses as unlike liquidated damages they do not reflect the compensatory principle of contract law. Some argue that a penalty clause

³⁵ [2005] BLR 271 280.

³⁶ (2005) 80 ALJR 219 (HCA).

³⁷ [1915] AC 79 87.

³⁸ (2005) 80 ALJR 219 (HCA) [32] (Gleeson CJ, Gummow, Kirby, Hayne, Callinan and Heydon JJ).

³⁹ (1993) 61 BLR 49 (PC) 58.

⁴⁰ [1978] 83 DLR (SC) 15.

⁴¹ [1989] ALJ 238 (HCA).

⁴² [2004] 2 NZLR 614 (CA).

⁴³ [2004] 2 NZLR 614 (CA) [59] (Blanchard J).

would promote the fulfilment by the parties of the contractual obligations. For Rowan, “the most persuasive argument in support of penalty clauses is the paramount importance of freedom of contract”.⁴⁴ In *Alfred McAlpine Capital Projects Ltd v Tilebox Ltd*⁴⁵ Jackson J referred to the fact that in the four cases where his Honour found a clause which was struck down as a penalty; “there was in fact a very wide gulf between (a) the level of damages likely to be suffered and (b) the level of damages stipulated in the contract.”

Questions arise as to whether the normal limitations of remoteness and mitigation must be taken into account in a consideration of liquidated damages. Lord Justice Diplock in *Robophone Facilities Ltd v Blank*⁴⁶ was of the view that if a stipulated sum was a genuine pre-estimate, it would be liquidated damages even if it was beyond the defendant’s normal loss under the remoteness rules. In *Abrahams v Performing Right Society Ltd*⁴⁷ Hutchison J referred to the fact that the liquidated damages clause facilitates recovery of damages and stated obiter that mitigation was not relevant in the case of liquidated damages. In *Murray v Leisureplay plc*⁴⁸ Lord Justice Buxton also referred to the fact that “a liquidated damages clause ‘saves the expense and difficulty of bringing witnesses to that point’”.⁴⁹ He concluded:

And a clause that made reference to the duty to mitigate would also inevitably postpone payment under the clause well beyond the termination date: again, something that the inclusion of such a clause in the contract must have been intended to avoid. This last consideration strongly reinforces the general impression created by this case, that the traditional learning as to penalty clauses is very unlikely to fit into the dynamics of an employment contract, at least when the penalty is said to be imposed on the employer.⁵⁰

4 Exclusion clauses and ‘Consequential’ or ‘Indirect’ Loss

Clauses excluding liability have been a source of litigation and there are clear differences as to the meaning of consequential or indirect loss in England and Wales, Australia and New Zealand.

4.1 England and Wales

There is a long line of English case law⁵¹ that has consistently followed *Millar’s Machinery Co v David Way & Son*⁵² and *Croudace Construction v Cawoods Concrete Products*⁵³. The English courts have adopted a narrow view of the meaning of ‘indirect’

⁴⁴ Rowan (n 13) 460.

⁴⁵ [2005] BLR 271 280.

⁴⁶ [1966] 1 WLR 1428 (CA) 1448.

⁴⁷ [1995] ICR 1028 (CA).

⁴⁸ [2005] EWCA Civ 963(CA (Civ Div)) [115].

⁴⁹ *Kemble v Farren* (1829) 6 Bing 141 148 (Tindal CJ).

⁵⁰ [2005] EWCA Civ 963(CA (Civ Div)) [115].

⁵¹ See *British Sugar Plc v NEI Power Projects Ltd* (1997) 87 BLR 42 (CA); *Deepak Fertilisers and Petrochemicals Corp v ICI Chemicals & Polymers Ltd* [1999] 1 Lloyd’s Rep 387 403 (CA); *Pegler Ltd v Wang (UK) Ltd* [2000] BLR 218; 70 Con LR 68; *Addax Ltd v Arcadia Petroleum Ltd* [2000] 1 Lloyd’s Rep 493 496; *Hotel Services Ltd v Hilton International Hotels (UK) Ltd* [2000] 1 All ER (Comm) 750; [2000] EWCA Civ 7(CA), *Watford Electronics Ltd v Sanderson CFL Ltd* [2001] BWCA 5(CA); *Leicester Circuits Ltd v Coates Brothers plc* [2003] EWCA Civ 290 (CA); *GB Gas v Accenture(UK) Ltd and others* [2009] EWHC 2734 (Comm).

⁵² (1935) 40 Com Cas 204 (CA).

⁵³ [1978] 2 Lloyd’s Rep 55 (CA).

or ‘consequential’ loss. This view which has been consistently followed in England and Wales is that the words ‘directly’ and ‘indirectly’ refer respectively to the first and second limb in *Hadley v Baxendale*:

Where two parties have made a contract which one of them has broken, the damages which the other party ought to receive in respect of such breach of contract should be such as may fairly and reasonably be considered either arising naturally, i.e. according to the usual course of things, from such breach of contract itself, or such as may reasonably be supposed to have been in the contemplation of both parties, at the time they made the contract, as the probable result of it.⁵⁴

In *Hotel Services Ltd v Hilton International Hotels (UK) Ltd*⁵⁵, the Court was concerned with loss following the installation of defective minibars in hotel rooms. Hilton was awarded damages for the rental either overpaid or paid on a failed consideration, the cost of removal and storage of the chiller units and cabinets and the loss of profit on the minibars. An exemption clause excluded liability “for any indirect or consequential loss, damage or liability arising from any defect in or failure of the System or any part thereof or the performance of this Agreement or any breach hereof by the Company or its employees”. The Court concluded that both the cost of removal and the loss of any profits which it would otherwise have been earning were direct or natural consequences and were not covered by the exemption clause “which (since all recoverable loss is literally consequential) plainly uses ‘consequential’ as a synonym of ‘indirect’. But nothing, at least in this area of the law, is so simple.”⁵⁶

The Court considered four authorities beginning with *Croudace Construction Ltd* where the distinction drawn at first instance by Parker J between consequential loss and natural or direct loss for the purposes of an exclusion clause was held to be correct on appeal. In *British Sugar plc v Nei Power Projects Ltd*⁵⁷ the contract between the plaintiffs and defendants was for the design, supply, delivery, testing and commissioning of electrical equipment by the defendants. The plaintiffs allege that the equipment was poorly designed and badly installed, which resulted in breakdowns in the power supply. At issue on appeal was whether the words seeking to place a limitation on liability for damages in relation to ‘consequential loss’ did or did not apply to loss flowing directly and naturally from a breach. The Court of Appeal rejected the submission that ‘consequential’ loss, to a reasonable businessman, would include loss of profits. Waller J adopted *Croudace Construction* and *Millar's Machinery* in which closely similar exclusion clauses were at issue, and stated that:

.... once a phrase has been authoritatively construed by a court in a very similar context to that which exists in the case in point, it seems to me that a reasonable businessman must more naturally be taken to be having the intention that the phrase should bear the same meaning as construed in the case in point.⁵⁸

Loss of profits and indirect or consequential loss were excluded in *Deepak Fertilisers Ltd v ICI Chemicals and Polymers Ltd*⁵⁹. The Court considered that it was bound by *Croudace Construction* where a similar loss was held to be direct loss. The final

⁵⁴ (1854) 9 Exch 341.

⁵⁵ [2000] 1 All ER (Comm) 750 (CA).

⁵⁶ [2000] 1 All ER (Comm) 750 [8 -9] (CA) (Sedley LJ delivering the judgment of himself, Lady Justice Hale and Mr J Rattee).

⁵⁷ (1997) 87 BLR 42 (CA).

⁵⁸ (1997) 87 BLR 42 (CA) (Waller LJ with whom Evans and Aldous LJ agreed).

⁵⁹ [1999] 1 L L R 387 (CA).

authority considered in *Hotel Services Ltd v BHP Petroleum Ltd v British Steel plc*⁶⁰, where Rix J concluded:

... the parties are correct to agree that authority dictates that the line between direct and indirect or consequential losses is drawn along the boundary between the first and second limbs of *Hadley v Baxendale*.

Counsel for Hotel Services, the appellant, argued that McGregor's⁶¹ view should be adopted. McGregor defined 'normal' losses as those which every plaintiff in a like situation will suffer, whereas 'consequential' losses were those peculiar to the particular claimant's circumstances. "Consequential losses are anything beyond this normal measure, such as profits lost or expenses incurred through the breach, and are recoverable if not too remote. The distinction is not the same as that between the first and second rules in *Hadley v Baxendale*: a consequential loss may well be within the first rule." The Court of Appeal rejected this approach. Sedley LJ concluded that:

We prefer therefore to decide this case, much as *Victoria Laundry* was decided, on the direct ground that if equipment rented out for selling drinks without defalcations turns out to be unusable and possibly dangerous, it requires no special mutually known fact to establish the immediacy both of the consequent cost of putting it where it can do no harm and - if when in use it was showing a direct profit - of the consequent loss of profit. Such losses are not embraced by the exclusion clause, read in its documentary and commercial context⁶².

The recent High Court decision *GB Gas v Accenture*⁶³ strictly adhered to the *Croudace* approach. Accenture plc contracted to design, supply, install and maintain a new IT system (the Jupiter System), including an automated billing system based on pre-packaged SAP IS-U software. The exclusion clause read⁶⁴:

16.2 Consequential Loss

Subject to Clause 16.7 or as otherwise expressly provided in this Agreement, in no event shall either Party be liable whether in contract, tort (including negligence) or otherwise in respect of any of the following losses or damages:

16.2.1 loss of profits or of contracts arising directly or indirectly;

16.2.2 loss of business or of revenues arising directly or indirectly;

16.2.3 any losses, damages, costs or expenses whatsoever to the extent that these are indirect or consequential or punitive

It was common ground that the words 'directly' and 'indirectly' in Clause 16.2 referred respectively to the first and second limb of the rule in *Hadley v Baxendale*⁶⁵.

Lord Hoffmann in *Caledonia North Sea v British Telecommunications plc*⁶⁶ reserved his view as to whether the interpretation adopted in the Hotel Services line of cases was correct.

4.2 Australia

Consequential loss was at issue in the decision of the Supreme Court of Victoria, *Environmental Systems Pty Ltd v Peerless Holdings Pty Ltd*⁶⁷. Peerless, an organic

⁶⁰ [1999] 1 L L R 583.

⁶¹ H McGregor, *McGregor on Damages* (18th ed), London, Sweet & Maxwell, 2009, [1-036-1-040].

⁶² [2000] 1 All E R (Comm) 750 [20].

⁶³ [2009] EWHC 2734 (Comm).

⁶⁴ [2009] EWHC 2734 (Comm) [77].

⁶⁵ (1854) 9 Exch 341

⁶⁶ [2002] 1 Lloyd's Rep 553 (HL).

recycling company, required a replacement afterburner. It accepted a written proposal from *Environmental Systems Pty Ltd* (ES) to replace the afterburner with a Regenerative Thermal Oxidiser (RTO). The system was installed and commissioned in October 1998. Despite considerable efforts to make the RTO work it was eventually shut down by Peerless in late July or August. Peerless sued ES for damages for breach of contract, among other actions. The following exclusion clause was in the agreement: Cl 8.9 “As a matter of policy, ES does not accept liquidated damages or consequential loss...”

The judge at first instance held that the expression ‘consequential loss’ in the exception clause referred to losses recoverable only under the second limb of *Hadley v Baxendale*. The appellant contended that upon its proper construction ‘consequential loss’ in clause 8.9 meant any losses that a business person would understand that term to embrace (in other words, any losses consequent upon a breach other than ‘liquidated damages’). Nettle J A (with whom Ashley and Dodds-Streeton JJA agreed) considered that the English authority appeared to be flawed. Reference was made to the view expressed in the earlier editions of McGregor on Damages that “the true distinction is between ‘normal loss’, which is loss that every plaintiff in a like situation will suffer, and ‘consequential losses’, which are anything beyond the normal measure, such as profits lost or expenses incurred through breach”⁶⁸. This distinction, McGregor explained, was evident in ss 50 and 51 of the Sale of Goods Act 1893 (Eng), which dealt with damages for non-acceptance and non-delivery of goods. The *Croudace* line of cases was considered and Nettle AJ agreed with McGregor that it was illogical to confine consequential loss to the second limb of *Hadley v Baxendale*⁶⁹ and then exclude it under the contract.⁷⁰ His Honour utilised Sedley LJ’s words in the *Hotel Services* case to describe McGregor’s approach, “that the conception of consequential loss should be restored to ‘the natural meaning of which commercial and legal usage in exclusion clauses has long since robbed it’”⁷¹.

Nettle JA was of the view that ordinary reasonable business persons would naturally conceive of ‘consequential loss’ in contract as everything beyond the normal measure of damages, such as profits lost or expenses incurred through breach. It was not correct to construe ‘consequential loss’ as limited to the second rule in *Hadley v Baxendale*. It was held that the ordinary and natural meaning of “consequential loss” intended by the parties in cl 8.9 was everything beyond the normal measure of damages, which is loss that every plaintiff in a like situation will suffer such as profits lost or expenses incurred through breach. Nettle JA concluded:

ordinary reasonable business persons would naturally conceive of ‘consequential loss’ in contract as everything beyond the normal measure of damages, such as profits lost or expenses incurred through breach⁷².

The New South Wales Court of Appeal in *Allianz Australia Insurance Ltd v Waterbrook at Yowie Bay Pty Ltd*⁷³ affirmed the Victorian Supreme Court’s decision in *Peerless*: the term “consequential loss” may include losses under the first limb of *Hadley v Baxendale* that arise “naturally” and in the ordinary context of the breach. Allianz was

⁶⁷ [2008] VSCA 26.

⁶⁸ *McGregor on Damages* (15th ed) London Sweet & Maxwell 1988 [25] et seq.

⁶⁹ (1854) 9 Ex 341; 156 ER 145.

⁷⁰ [2008] VSCA 26 [90].

⁷¹ [2008] VSCA 26 [90].

⁷² [2008] VSCA 26 [93].

⁷³ [2009] NSWSCA 224.

granted leave to appeal the Supreme Court decision, *Waterbrook at Yowie Bay Pty Limited v Allianz*⁷⁴. Under section 18B of the Home Building Act 1989 (NSW) every residential contract has implied in it certain statutory warranties, and a contract insuring residential building work must insure against the risk of loss that may arise out of a breach of the statutory warranty. Neither the statutory warranties nor the statutory insurance rights can be excluded or restricted. At issue in *Waterbrook v Yowie Bay Pty Limited v Allianz* in this context was the exclusion clause in Allianz's insurance policy covering the building work done by Waterbrook. The clause stated in part: "The insurer should not be liable for any claim for loss or damage: (10) For loss of use and/or consequential loss of any other kind arising directly or indirectly out of any event listed in the Building Owner's indemnity, except as provided in subclause (5)." In the Supreme Court, McDougall J held that the relevant measure of damages reflected the first limb in *Hadley v Baxendale*. His Honour adopted the approach in *Environmental Systems Pty Ltd v Peerless Holdings Pty Ltd* that some 'consequential loss' may well fall within the first limb. A majority of the new South Wales Court of Appeal agreed with McDougall J that limitations on the entitlement to indemnity in the policy were void due to inconsistencies with the Home Building Act.

4.3 New Zealand

In the High Court decision, *Oceania Furniture Ltd v Debonaire Products Ltd*⁷⁵ there was no binding authority in New Zealand on the meaning of 'consequential' losses. Clifford J discussed the two views and noted that the highest courts in England and Wales and Australia have not yet considered the issue. His Honour preferred the Victorian Court of Appeal's approach in *Peerless*, which was to consider the ordinary commercial meaning of 'consequential losses' in the context of the particular contract. The meaning of the terms 'indirect' and 'consequential loss' was in Clifford J's view clear from looking at those terms within the context of the particular contract entered into by Oceania and Debonaire. On the basis of the Australian view, 'consequential loss' would include everything beyond normal loss, and would include lost profits and expenses. In this case, the ordinary and natural meaning of the exclusion of liability for 'indirect and consequential losses' covered all losses other than those arising directly from, or immediately associated with, Oceania's obligations.

4.4 The desirable approach?

In the three jurisdictions discussed above the approach that will be adopted by the courts is uncertain as in all cases there has not been a definitive decision from a higher court. The meaning of the words is not clear and as a matter of construction it will vary from contract to contract. Carter's convincingly rejects both the *Coudace* and the McGregor approach⁷⁶ and argues, inter alia, that on a consideration of the dictionary meaning⁷⁷, the words are concerned with causation and have nothing to do with remoteness or the measure of damages. The courts in the discussed cases have determined the meaning of the words on a 'once and for all' basis. However, the function of a court is to construe

⁷⁴ [2008] NSWSC 1451.

⁷⁵ 27/8/ 2009, Clifford J, HC Wellington CIV 2008 485 1701 [121].

⁷⁶ 27/8/ 2009, Clifford J, HC Wellington CIV 2008 485 1701 [121].

⁷⁷ The Oxford English Dictionary's first two meanings of 'consequential' are 'of the nature of a consequence or sequel' and 'of the nature of a consequence merely, not direct or immediate'.

each particular contract and a presumptive approach that the words have a particular meaning is not appropriate. The difficulties with these words highlight the importance of strict attention to the drafting of contract provisions to ensure that they say what is intended.

5 Conclusion

Fuller and Perdue's classification is no longer appropriate when viewed in light of the changed available monetary relief for contract damages since *Attorney-General v Blake*⁷⁸. A more appropriate classification is that of Tipping J in the New Zealand Supreme Court, where his Honour described the types of monetary relief in response to civil wrongs as:

compensatory damages, disgorgement damages and restorative damages. Compensatory damages are loss based. Disgorgement damages are based on giving up a gain. Restorative damages are based on restoring to the plaintiff value transferred.⁷⁹

Compensatory damages are “the governing principle in contract”⁸⁰ and these are the category of damages which the courts will generally award. In some cases the courts will emphasise the performance aspect of expectation damages but this should not be regarded as a different measure of contract damages. The principle of freedom of contract is subordinate to this governing principle and it is unlikely that there will be a radical change to the availability of agreed damages.

The three jurisdictions have adopted a similar approach to liquidated damages clauses and have readily adopted new terminology in their approach to liquidated damages. There are opposing views in relation to the meaning of ‘consequential’ loss in an exception clause. Parties do not enter their own contracts on the basis of what the words mean in other contracts, and each particular contract should be viewed on its facts. The meaning attributed in previous contracts should be irrelevant in deciding what an exception clause means. What is required is a ruling from the courts at the highest levels to ensure that the meaning is ascertained by construction of the contract rather than a presumptive approach.

⁷⁸ [2001] 1 AC 268 (HL).

⁷⁹ *Premium Real Estate Ltd v Stevens* [2009] 2 NZLR 384 (SC) 418 (Tipping J).

⁸⁰ *Golden Strait Corp v Nippon Yusen Kubishika Kaisha (The Golden Victory)* [2007] UKHL 12; [2007] 2 AC 353 (HL) [9] (Lord Bingham).

The capitalization rates for the valuation of the land uses in the Spanish law context.

Vicente. Caballer¹, Natividad Guadalajara¹, Elena De la Poza¹

¹ INECO. Faculty of Business Administration.

Universidad Politécnica de Valencia Building 7J. Camino de Vera s/n. 46022. Valencia. Spain.

Email: vcaballer@esp.upv.es ; nguadala@omp.upv.es; elpopla@esp.upv.es

Abstract:

Purpose - In 2008, a Royal Decree was passed in Spain regulating the real estate valuation processes in cases of expropriation, forced sales, urban planning and development, and Public liability of personal assets. In valuation terms, the Law differentiates between rural and urban land. In cases of rural lands, the law only considers as an appraisal methodology, the capitalization of the real annual incomes or the potential incomes of the farm, including the subsidies of the crops and land uses. The study develops a critic of the methodology applied by law, especially the rates of capitalization used (3-year bonds).

Design/methodology/approach - In particular, the Spanish law specifies that the rate of capitalization to apply will be the last data published by the Spanish Central Bank about the return of the Spanish Public Debt (3-year bond) in the secondary markets. However, the law of General State Budget could modify these rates of capitalization and fix minimum values according to the kind of crops and uses of land. Those values are estimated by the division between the return of the land and the land values, according to the crops and Autonomous Communities in Spain. Finally, it is shown a graphic of the development of the return of the land and the rate return of the Spanish public debt that shows the differences between both

Findings - As a result of the application of the law, the estimated values are always lower than the market values, as the land has a lower rate of return than rates of return of the public debt. Capitalizing the rents of the land at the going rates of return of the public debt, land values are lower than the market values. As consequence, the owners become impoverished. The behavior of the rates can't be generalized to other countries; for instance in USA, the development was completely the opposite.

Originality/value – This paper can be interesting for developing the Regulations of the Land Law. Its main value is to invite the reader to think about rates of capitalization more adjusted to the market.

Key words: Land, royal decree 2008, rate of capitalization, rate of return, 3-year bonds, valuation.

1 Introduction

In Spain, there exists three important laws that condition the regulated real estate valuation; two of them of private interest and third one of public interest. Regarding to the private interest are the Accounting Plan Real Decree 1514/2007, of 16th of November, that dictates norms for the valuation (accounting effects), and Order ECO/805/2003, BOE n 85 of 9/4/2003, that regulates the valuation related to:

- The Mortgages' credits or loans warranties
- Coverage of the technical provisions of the insurance firms
- Definition of the net assets of the real estate funds investment firms
- Definition of the net real estate assets of the Retirement Funds

The legislation of valuation of public interest is Real Legislative Decree 2/2008, BOE n 154 of 26/6/2008, by which the rewritten text of the Land Law is approved. This new replaces:

- Law 8/2007, of 28th of May, related to land, which as well, came to replace previous Law 6/1998, of 13 of April, Land systems and Valuations.
- Real Legislative Decree 1/1992, of 26th of June, which approves the rewritten text of the Law that regulates Land systems and Urban Development.

The law 8/2007 and the Real Decree 2/2008, were passed with the aim of restraining the urban development corruption and ending a stage in which the price of the housing grew sharply. In the introduction of the Law is said *"..... the classification (of the land) has contributed historically to the inflation of the values of the land, having incorporated expectations of revaluation long before than any action necessary to materialize the urban development were taken by public authorities and, therefore, it has also fomented the speculative practices, against which we must fight by constitutional imperative."*

According to the Spanish Ministry of Housing, www.mviv.es the Spanish average value of housing was 2,071 €/sqmetre in 2008, after undergoing an accumulated increase from the year 2000 of 143.64%, (and a annual average increase of 15,96%). From 2008 the average price falls down due to the current economic crisis, which affects of special way to the construction sector in Spain.

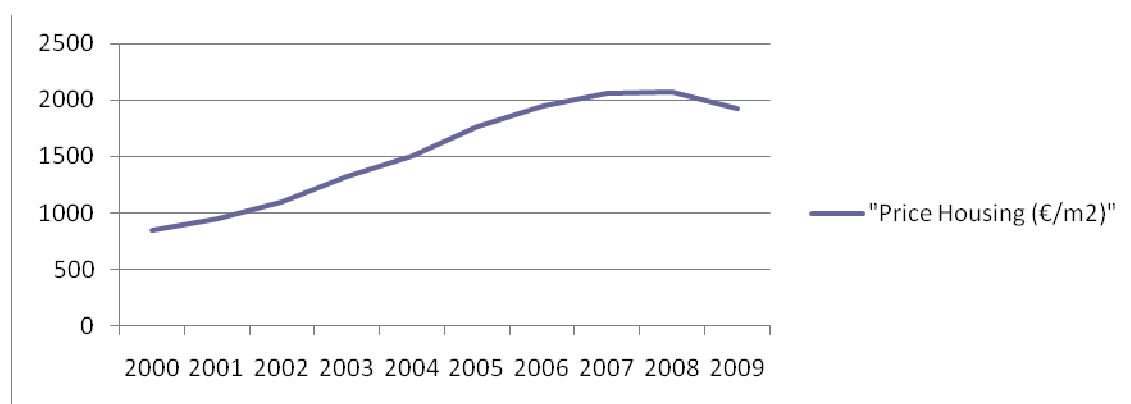


Figure 1. Development of the average housing price in Spain (€/sqmetre)
(Source: Own Performance)

The average annual increase of 15.96% during the period 2000-2008 was the highest ever happened in Spain. In fact, the average annual growth during the decade 1998-2008 was much lower, of 10.3% (as figure 2 shows).

However, this increase has not taken place equally in all the country. In particular, the Autonomous Communities (AC) that experienced a higher growth were the coastal ones, as it is the case of Andalusia, Murcia and the Balearic Islands with a rate of growth superior to the national average (superior to 11.9%).

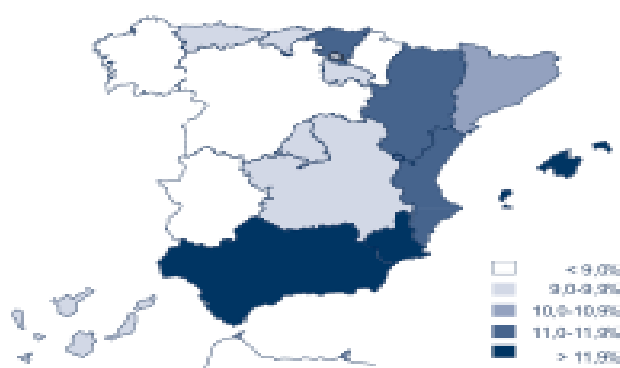


Figure 2. Average increase of the housing prices by Autonomous Communities (AC) (1998-2008).
(Source: Own performance from data provided by the Spanish Ministry of Housing).

One of the issues that mainly influences the urban development of Spain, it's the property of a second house, in the beach or the mountain, to spend the weekends, or holidays. Figure 3 shows the proportion of first houses by AC. Between the 26 and 40% of the total volume of residential housing is a holiday house.

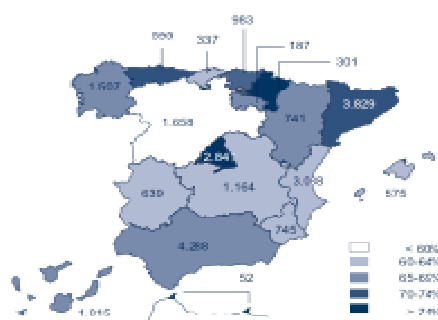


Figure 3. Proportion of main houses by Acin 2007.
(Source: Own performance from data provided by the Spanish Ministry of Housing).

This important urban development and the increase of the price of the residential real estate sector produces a high level of pressure over the rural land (Berry and Plaut, 1978; Levia, 1998); and therefore an increase of the price of the land with agricultural use (Shi, *et al*, 1997; Plantinga *et al*, 2002). This fact justifies the development of the Law.

2 The method of valuation according to the Law.

The title III of the Law¹, in its article 21 focused on the objects of valuation of the land, the facilities, constructions and buildings regulated by the mentioned Law, which in particular are:

- Territorial and urban city development
- Expropriation
- Sale or compulsory replacement
- Net assets responsibility of the Public Administration.

Also, in the same title III, the valuation procedures are settled down, for both uses of the land: rural and urban ones, whose value is estimated by its real situation and not by its expectations. This way, the purely speculative practices of classification of the uses of the land are discouraged.

According to the Law the uses of the land are: rural or urban. Particularly, section 3 of the Law says: *“the land with urban use is the one which is integrated legally and effectively in the net of facilities and services of the city centre. It will be this will also apply in case the plots of land are or not built; when the plots have access to the facilities and services required by the urban legislation or can be gotten connecting the plots to the facilities already operating.”*

Therefore, by exclusion the non-urban land will be defined as rural land, that is approximately a 2% of all Spain.

For each type of land different methods of valuation are applied. With regard to the rural land, the Law says in its article 23: *“The land will be appraised by capitalization of the real or potential annual rent (the one that is higher) of the farm according to its state at the moment the valuation is produced.....”*

“...In any case the appraisal will take in account the expectations of building or the urban development plan.”

The analytical method or the capitalization postulates that the value of a building is closely related to the *rent* that produces or that is able to produce in the future. For its application it is demanded to have simple knowledge about accounting and finances, in such a way that the value will be estimated by the divisor between the real or potential rent R and the rate of actualization (r).

$$V = R/r \quad [1]$$

Clark *et al* (1993) and Goodwin *et al* (2003) criticize the capitalization method as a suitable method of valuation to evaluate the land of agrarian use, because the rents often

¹ Law (Land Law) 8/2007, May 28th.

do not reflect the value of the land (Engsted, 1998), and exist numerous drivers that condition their value (Xu, *et al*, 1993).

The valuation of the rural land according to the Law, (in accordance with the principles of the Spanish Constitution), supposes an enormous difficulty for two reasons: on one hand because the Constitution recognizes the property right of all Spaniards and the value of replacement. On the other hand, also in the Spanish Constitution, the fight against the land speculation is defended, since it is a scarce good.

These two principles of the Constitution lead to an enormous contradiction for two reasons: the right to the property supposes the compensation by the value of replacement of the land, that is not other than the market value; and also, the market value favors the speculation due to the shortage of the land and the increase of the demand by the speculative bubble (Tegene and Kuchler, 1993).

Consequently, the Law does not apply the market value estimated by comparative methods (based on using recent values of trades of similar buildings), it demands the valuation by the method of capitalization of the real or potential rent of the farm. This supposes to link its value to the highest and better use, but without urban developments expectations, which is extremely restrictive (Shi *et al*, 1997; Cavailhès and Wavresky, 2003).

Besides defining the valuation method to apply, the Law says in section 1 of the additional seventh regulation: “... *the last reference published by the Spanish Central Bank of the rate of return of the 3-year bond public debt in the financial (secondary) market will be used as rate of capitalization.*”

Next, the section 2 of the regulation says: “*In the Law of General Budgets of the State it will be possible to modify the rate of capitalization established in the previous section and to determine the minimum values according to types of crops and land advantages, when the development observed in the prices of the land or the rates of interest risks to move away from the result of the valuations with respect to the prices of market of the rural land without consideration of urban development expectations.*”

Therefore, besides the fact the rate of capitalization does not have to consider the potential urban development of the land, it is possible to think that the trend of the rates of capitalization does not have to coincide with the rates of interest (Daouli and Demoussis, 1992), and in particular with the rate of return of the public debt, because the market of the land can have a different development from the financial markets.

At the present time, the regulations of application of the Law is being developed, a extremely complex document by all the features that can be found in the real estate sector in Spain.

Consequently, the main objective of the present work is to analyze the possible correspondence between the yield of the 3-year Spanish public debt in the secondary markets and the observed land rate of return. The purpose is to act as a consultant to set the more suitable rates of capitalization, fulfilling the principle of the right to the property of the Spanish Constitution.

Also, the study tries to analyze if these differences are only in Spain, or if it is a common fact in other countries of the world, as United States of America.

3 Method y Data

A graphical method is used to compare the development of the rates of interest, for both the Spanish public debt and also the land rates of return during the last years.

Also, for the estimation of the rates of return of the land, expression [1] is applied, as result of dividing the rent of the land and its market value.

The three year public debt (in the secondary market) data were obtained in Web (www.tesoro.es) whereas the market values and the rent of the land were provided by the Ministry of the Environment and Rural and Marine (MERM) (www.marm.es).

The MERM shows the average prices of the land by AC and main crops from 1983. Between the most important crops are: dry lands, irrigated land, irrigation-transformed vineyard, irrigation-transformed olive groves, pastures, and meadowlands.

The values of the rent are also obtained for each AC and crop, but only for the period of time (1998 to the 2008), both inclusive. Therefore, this period of 10 years will be our object of study.

4 Results

The trend of average prices of the agrarian land and its rent or cash rent, in €/ha is showed in figure 4a and 4b.

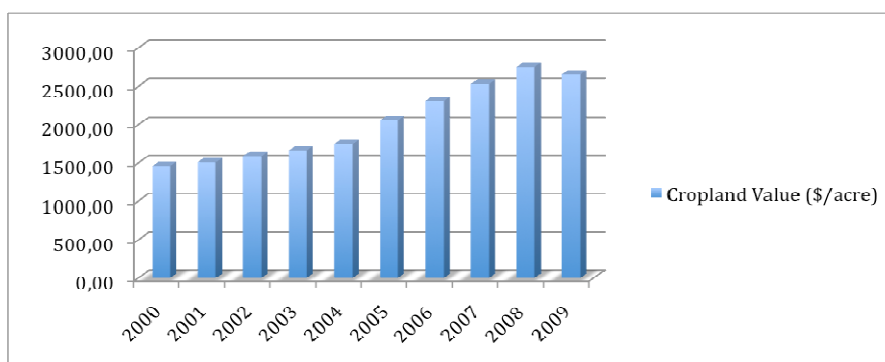


Figure 4a. Development of the average prices of the land of agrarian use in Spain.
(Source: Own Performance from data published by MERM.)

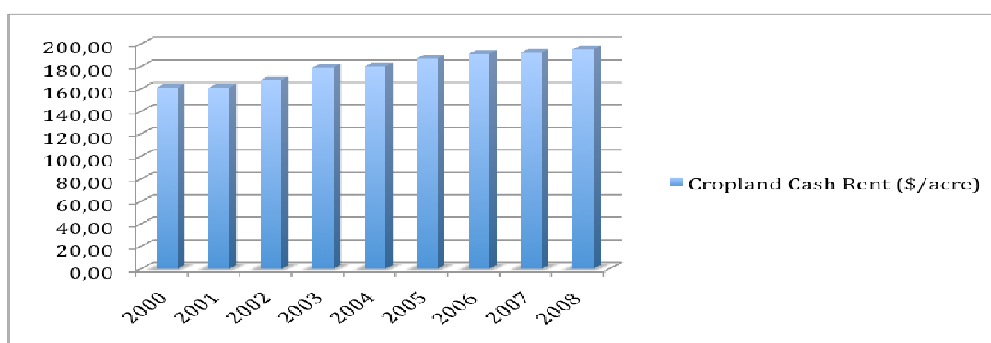


Figure 4b. Development of the average cash rents of the land of agrarian use in Spain.
(Source: Own Performance from data published by MERM.)

It is observed how the rent has practically remained constant during the period 2000-2008, fluctuating between 161 €/ha in the year 2000 and 195 €/ha in 2008, which has supposed an accumulated increase of 21.11% during the 9 years considered (what is equivalent to an average annual increase of the 2.34%).

The growth of the average land prices has duplicated the increase of the rent, concretely 49.39% during the same period, equivalent to 5.48% annual. These increases have been more noticeable in pastures, in which the rent has increased an accumulate rate of 25.86% (equivalent to an average annual increase of 2.87%); and the prices have experienced an accumulated increase of 53.38% (5.93% annual average increase) (See figure 5a, 5b)

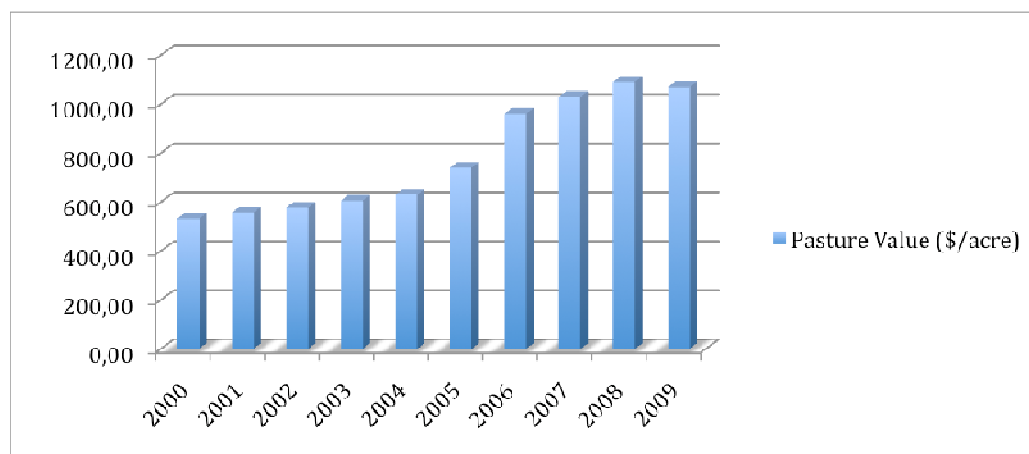


Figure 5a. Development of the average prices of the pastures in Spain.
(Source: Own Performance)

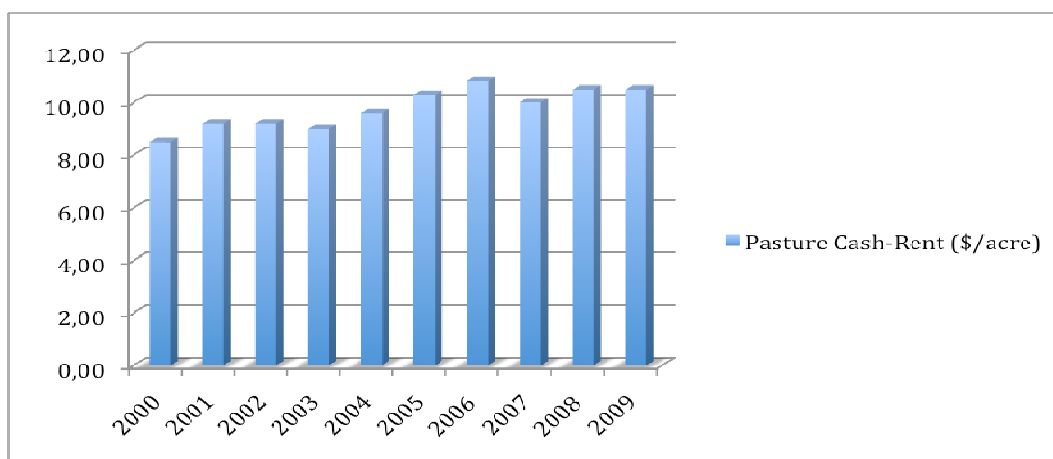


Figure 5b. Development of the average cash rents of the pastures in Spain.

(Source: Own Performance)

The trend of land prices has produced a continued decrease of the rates of return of the land (both kind of land: cropland and pastures) during the period (2000-2008). Figure 6 shows this development and also the development of the rate of return of a 3-year bond Spanish public debt.

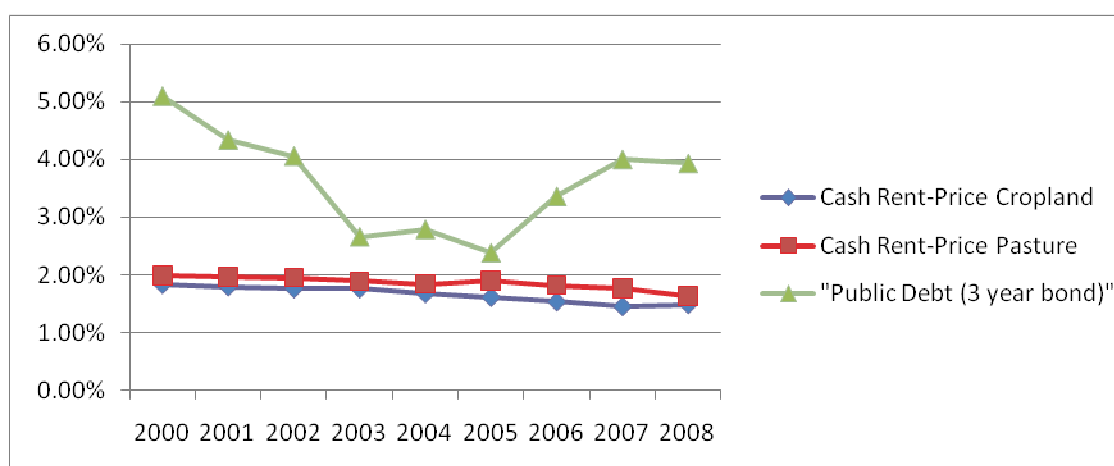


Figure 6. Development of the rates of return of the Public Debt and the rate of return of the cropland and pastures in Spain.

(Source: Own performance)

The rates of return of the two kinds of land, (cropland and pastures), are very similar and always lower to 2%. The rate of return of cropland was 1.98% in the year 2000 and 1.63% in 2008, whereas the pastures rates of return were 1.83% and 1.49%, in the year 2000 and 2008 respectively.

On the contrary, the rate of return of the Spanish public debt is higher every year, the average rate was 3.62% in the period considered, with increases till rate of 5.10% at the beginning of the period and dropped to 2.38% in 2005. As figure 7 shows, it does not follow a constant tendency, since it has cyclical movements.

This indicates that the behavior of the financial markets free of risk is very different from the land market; and in addition, the financial markets are not good indicators of the land markets.

Capitalizing the rents to high rates, (more than double), leads to much lower land values, practically more than half according to the considered years.

The values obtained, are not representative of the market value of the land of agrarian use, as cropland land or pastures.

Consequently, the application of the rates established by the Land Law constitutes a strong offense for the proprietors, not only as land with possible urban expectations, but also as exclusively agrarian land. A more exhaustive analysis of the returns of the land by crops is shown in figure 7.

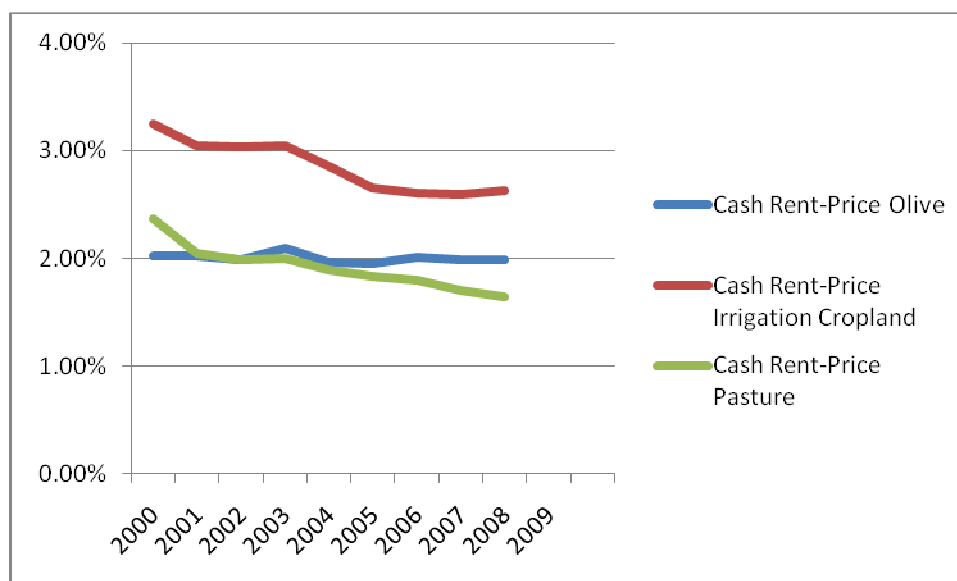


Figure 7. Development of the returns of the land by crops.

(Source: Own performance)

The land yield, and therefore the rates of capitalization to apply in each crop to obtain the market value, is different in each case, (concerning to the value in every year, and with regard to the tendency). The higher returns are obtained in the irrigated land, although it shows a decreasing tendency, as it happens in pastures. On the contrary, the yield of the irrigated-transformed olive groves land practically stays constant in all the period and around 2%.

5 A comparative analysis of the public debt and the return of the land in USA.

The information published annually by United States Department Agricultural (USDA) about the development of the average prices of the land and the rents shows the value of the land of agricultural use has experienced an accumulated increase of the 81.5% and the 101.5% for crops and pastures respectively. This implied an annual increase of 8.1% and 10.15% for crops and pastures respectively.

On the other hand, the rents experienced during the same period, an accumulated increase of the 28.57% and 23.52% in lands of agricultural use and pastures

respectively, whereas the average annual increase reached the value of 2.85% and 2.3% respectively.

In particular, figure 8 shows the development of prices of the cropland and the average rents for period 2000-2009. It is observed that in average, the value of the cropland decreased in 3.2% in 2009 from 2008.

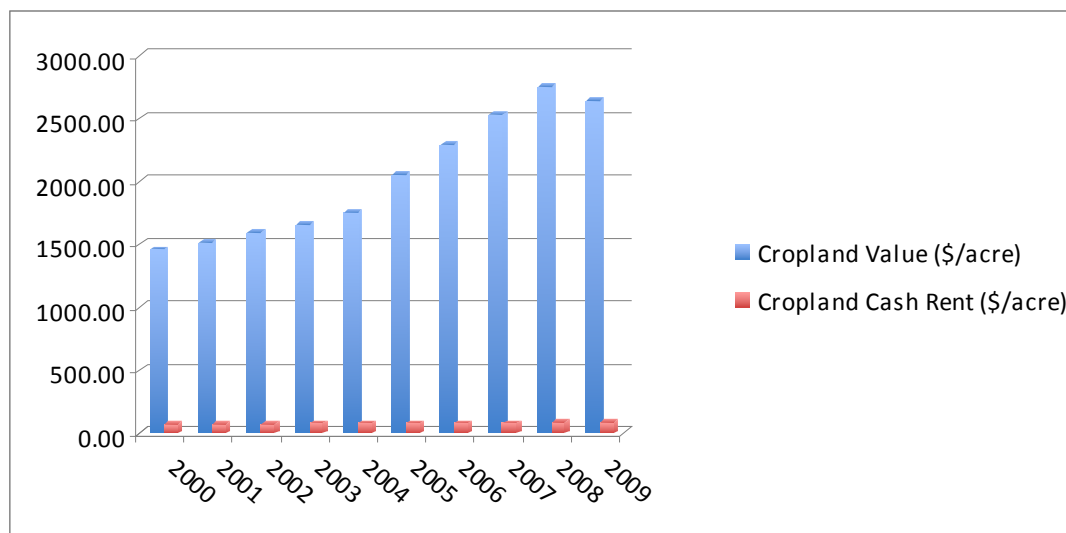


Figure 8. Development of the annual average prices and cash rents of the pastures and crop lands in USA.

(Source: Own Performance from data provided by USDA)

Also, figure 9 shows the development on average of the pastures for the same period 2000-2009, as well as the average rents produced by the pastures. It is noticed as the value of the pastures dropped in 2009, in a 1.8%. (Figure 9).

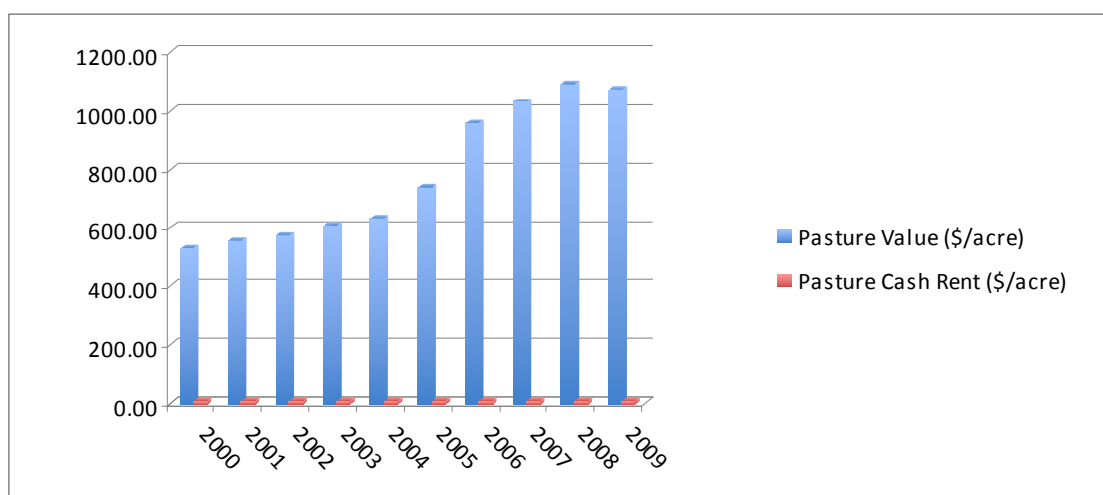


Figure 9. Development of the average prices and cash rents of the pastures in USA.

(Source: Own Performance from data provided by USDA)

From this information, figure 10 shows the rates of return of the cropland and pastures. The rates of return of the cropland and pastures and also the rates of interest of the three years public debt (three year bonds) (Federal Reserve the United States) are graphed.

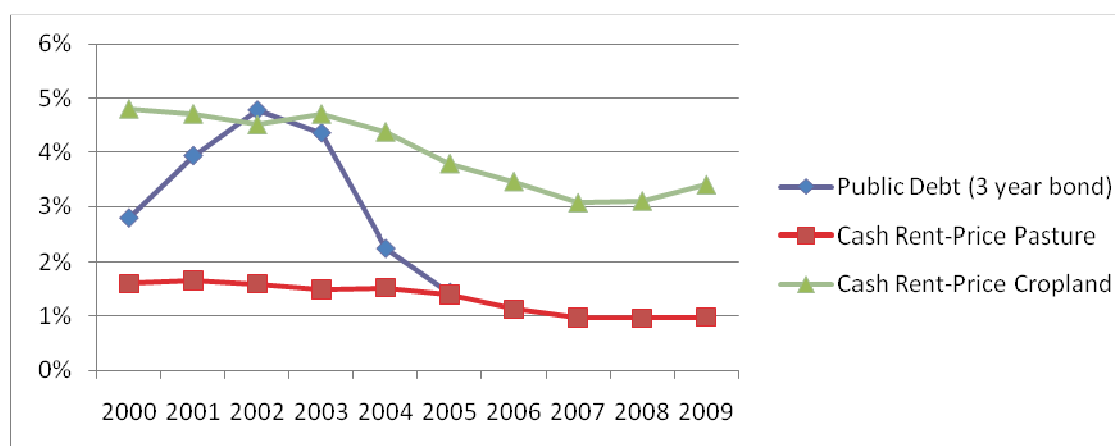


Figure 10. Development of the rates of return of the 3-year public debt and the rates of return of the land. (Cash Rent-Price Cropland) in USA.

(Source: Own performance)

As in Spain, the rates of return of the cropland and the pastures follow parallel tendencies (but very different values) in the USA during the period of time considered. The average rate of return of the land in the United States is 3.99%, and 1.33% for pastures. One of the factors that can influence in these differences between countries, are the subsidies provided by the public institutions. In Spain, in particular, the new Community Agrarian Policy (CAP) is applied from 2006, and implies the decoupled of the production from the subsidies; this model was already applied in EE.UU under the name of the Farm Bill, which started up, in 1960.

On the other hand the national debt follows a behavior cyclical, just as in Spain, as a result of the periods of expansion and economic contraction (crisis), but of different form. That is to say, it shows a similar behavior in both countries until the year 2005, and then it is observed how the rate of return of the public debt in Spain increased, as a result of the policy of the Government to catch bottoms, whereas in USA remained at levels of 1%. An application of the method of capitalization in the United States, according to the Spanish Law would not suppose any offense for the proprietors, but rather on the contrary.

6 Conclusion

In Spain has taken place during the period 2000-2008 a strong increase of the price of the housing and an urban development, which led to a consistent pressure on the rural land.

With the intention of guaranteeing the principle of the Spanish Constitution to restrain the speculation of the land, the Law of the land was promulgated in 2007 and its Real Legislative Decree in 2008. Its Regulation of application is pending to publish.

The purposes of the Law are of public utility: expropriation, territorial and urban city development, sale or compulsory replacement, net assets responsibility of the Public

Administration. Finally the Law determines the methods of valuation for these aims and each land class.

For the rural land the method of capitalization of the rent is defined as unique useful method, which is of doubtful validity, because the rent is not the unique determinant of the prices of the land of agricultural use. In addition it is set a rate of capitalization from the financial market of public debt, whose behavior in Spain is particularly different from the Land market. An application of the rates defined by Law can reduce the prices of the Land to half of its market value, which supposes an offense for the proprietor.

Also, in the United States the land markets are different with respect to the public debt, but in this case the rates of return of the public debt are higher than the rates of return of the land, consequently, the application of the Law would favor the proprietor.

Therefore, the rates of the public debt should not be used as rates of capitalization in all the cases. This rate of capitalization should be fixed annually and for each crop with the purpose of obtaining a value adapted to estimate the value of the rural land.

7 References

- Berry, D. y Plaut, T. (1978), "Retaining Agricultural Activities under Urban Pressures: A Review of Land Use Conflicts and Policies", *Policy Sciences*, Volume (9), pp 153-178.
- Burger, A. (1998), "Land Valuation and land rents in Hungary". *Land Use Policy*, Volume (15,3), pp 191-201.
- Levia, D.F. (1998), "Farmland conversion and residential development in North Central Massachusetts", *Land Degradation & Development*, Volume (9), pp 123-130.
- Plantinga, A.J.; Lubowski, R.N. y Stavins, R.N. (2002), "The effects of potential land development on agricultural land prices", *European Review of Agricultural Economics* Volume (52), pp 561-581.
- Cavailhès, J. y Wavresky, P. (2003), "Urban influences on periurban farmland prices", *European Review of Agricultural Economics*, Volumen (30;3), pp 333-357.
- Clark, J. S.; Fulton, M. y Scott, J.T. (1993), "The Inconsistency of Land Values, Land Rents, and Capitalization Formulas", *American Journal of Agricultural Economics*, Volume (75), pp 147-155.

Daouli, J. y Demoussis, M. (1992), "Rents, Interest Rates and Real Agricultural Land Prices: An Application to a Greek Province", *European Review of Agricultural Economics*, Volume (19;4), pp 417-425.

Engsted, T. (1998), "Do farmland prices reflects reflect rationally expected future rents?", *Applied Economics Letters*, Volume (5), pp 75-79.

Goodwin, B.K.; Mishra, A.K. y Ortalo-Magne, F.N. (2003), "What's wrong our models or agricultural land values?", *American Journal of Agricultural Economics*, Volume(85 ;3), pp 744-752.

Shi, Y.J.; Phipps, T.T. y Colyer, D. (1997), "Agricultural Land Values under urbanizing influences", *Land Economics*, Volume (73;1), pp 90-100.

Tegene, A. y Kuchler, F. (1993), "Evidence on the Existence of Speculative Bubbles in Farmland Prices", *Journal of Real Estate Finance and Economics*, Volume (6), pp 223-236.

Xu, F.; Mittelhammer, R.C. y Barkley, P.W. (1993), "Measuring the Contributions or Site Characteristics to the Value of Agricultural Land", *Land Economics*, Volume (69;4), pp 356-369.

International contracts and the choice of law in New Zealand

Dr N Richardson

Senior Lecturer in Law

University of Canterbury

New Zealand

Introduction

An international contract is a contract that has contacts with more than one jurisdiction. Parties to such a contract may actually incorporate foreign law by writing terms of that foreign law into their contract, or they may also choose a particular jurisdiction in which to litigate, if the need arises, by putting a choice of forum clause into their contract. This can be a simple clause stating that the forum of a certain country or state is to apply if there is a dispute. Arbitration clauses may also be written into an international contract in the same way. In New Zealand the proper law of the contract determines which law is to apply if a dispute arises.¹

The Proper Law Defined

The term “proper law of a contract” means the system of law by which the parties intended the contract to be governed or, where their intention is neither expressed nor to be inferred from the circumstances, the system of law with which the transaction has its closest and most real connection.²

This definition, given by Dicey & Morris, includes the three rules on choice of law for contracts, because it is sufficiently widely drafted to incorporate:

- the parties’ actual express choice;

¹ *Vita Food Products Inc v Unus Shipping Co Ltd* [1939] AC 277 (PC); *Club Mediterranee NZ v Wendell* [1989] 1 NZLR 216 (CA).

² L Collins, Dicey & Morris *The Conflict of Laws* (11th ed), London, Stevens & Sons Limited, 1987, Rule 180, 1161-1162 (hereinafter referred to as Dicey & Morris). This is the last edition where the law of New Zealand coincided with that of England. The EEC Convention on the Law Applicable to Contractual Obligations 1980 also called the Rome Convention is now in force in England by virtue of The Contracts (Applicable Law) Act 1990). See generally Dicey & Morris, 1161; D Goddard & H McQueen, *Private International Law in New Zealand*, Wellington, New Zealand Law Society, Repr. 2001, 131. See also *Mount Albert Borough Council v Australasian Temperance and General Mutual Life Assurance Society* (1938) AC 224, at 240 per Lord Wright (PC); *James Miller and Partners Ltd v Whitworth Street Estates (Manchester) Ltd* [1970] AC 583, at 603 per Lord Reid (HL).

- the parties' choice which although not express can be inferred;
- the situation where the parties have not made a choice of law decision. Here the court applies the law most closely connected to the contract to govern.

Express Choice

Parties to an international contract can expressly state that they wish the law of New Zealand, or wherever, to govern their contract and then the court will, subject to certain qualifications discussed below, honour such choice. This freedom to choose is called party autonomy. Cases applying party autonomy have existed since 1760³. Dicey & Morris state the rule in the following words:

When the intention of the parties to a contract, as to the law governing the contract, is expressed in words, this expressed intention in general determines the proper law of the contract.⁴

All the parties need to do is declare their common intention by a simple statement that the contract shall be governed by the law of a particular jurisdiction. In an early New Zealand case it was specified that, "the contract evidenced by this bill of lading shall be governed by the law of England."⁵ The case, which was a decision of the then New Zealand Supreme Court, concerned a contract involving shipment of goods from England to New Zealand. The court held that whilst English law governed the contract as the law of the place of contracting (*lex loci contractus*) the matter was completely settled by the incorporation of this clause and thus New Zealand legislation did not apply to the contract.⁶

The leading case is the Privy Council decision of *Vita Food Products Inc v Unus Shipping Co Ltd*.⁷ where a cargo of herrings was shipped from Newfoundland for New York on board a Nova Scotian vessel. Owing to negligent navigation the ship came to grief off the coast of Nova Scotia and the herrings were damaged. The bill of lading stated that the contract was governed by English law and exempted the ship-owners from liability for negligence. A Newfoundland Statute provided that every outward bill of lading

³ *Robinson v Bland* (1760) 2 Burr 1077. See also *Lloyd v Guibert* (1865) LR 1 QB 115; *Jacobs v Credit Lyonnais* (1884) 12 QBD 589.

⁴ See Dicey & Morris, above n2, Rule 180, Sub Rule 1, 1168. See also *Gienar v Meyer* (1796) 2 Hy Bl 603.

⁵ *New Zealand Shipping Co Ltd v Tyree* (1912) 31 NZLR 825.

⁶ Thus also illustrating the New Zealand development away from the mechanical rule of the law of the place of contracting (*lex loci contractus*).

⁷ [1939] AC 277 (PC), referred to or applied in New Zealand in other contexts (mostly jurisdiction) but not discussed in any detail. See for example, *Cornwall Properties v King* [1966] NZLR 239; *Campbell Motors Ltd v Storey* [1966] NZLR 584 (CA); *Graham v Attorney-General* [1966] NZLR 937 (CA); *Carey v Hastie* [1968] NZLR 276 (CA); *Air New Zealand Ltd v The Ship "Contship America"* [1992] 1 NZLR 425 (HC); *Knyvett v Christchurch Casinos Ltd* [1999] 2 NZLR 559 (CA).

from Newfoundland must contain the Hague Rule. However, because out of date forms were used the bill of lading in question did not contain these relevant provisions. The buyers sued the ship-owners. Both parties assumed that Newfoundland law was the proper law of the contract.⁸ The Privy Council, however, held that English law was the proper law and consequently, as the choice of it was valid, there was no liability by reason of the contract.

Limits on Party Autonomy

Bona Fide, Legal and not Contrary to Public Policy

In the *Vita Foods* decision Lord Wright also dealt with the question of the limits to be assigned to the doctrine of party autonomy. He said that, “it is difficult to see what qualifications are possible, provided the intention expressed is bona fide and legal and there is no reason for avoiding the choice on the ground of public policy.”⁹

This statement is not free from ambiguity as it is difficult to know by which law Lord Wright intended the legality to be tested. The only safe conclusion would appear to be that a choice is not bona fide or legal if it can only be explained as having been made to escape some otherwise applicable law. Dicey & Morris¹⁰ suggest that a foreign law would thus not be applied if it contravened the public policy of the forum or the mandatory provisions of a proper law. A choice of law unconnected with the contract could be seen as an evasive attempt and therefore not bona fide.

There would appear to be only two Commonwealth decisions, both Australian decisions, where the parties’ choice was set aside as being either unconnected with the realities of the contract or an attempt to evade a statute in force in the forum. The first case was *Golden Acre Ltd v Queensland Estates Ltd*.¹¹ Where a contract was made in Queensland between a Queensland undertaking and a Hong Kong company. The latter was doing business in Queensland as a real estate agent and the contract involved the sale of some Queensland land to, amongst others, persons from Hong Kong. Queensland statute law

⁸ *Vita Food Products Inc v Unus Shipping Co Ltd* [1939] AC 277 (PC).

⁹ See J H C Morris & G C Cheshire, “The Proper Law of a Contract” (1940) 56 LQR 320, 325.

¹⁰ See Dicey & Morris, above n 2, at 1172.

¹¹ [1969] Qd R 378 (QSC), affirmed on other grounds by the High Court of Australia (1970) 123 CLR 418, Sub nom; *Freehold Land Investments Ltd v Queensland Estates Pty Ltd*.

required that real estate agents be licensed and provided for maximum rates of commission chargeable by real estate agents. The Hong Kong company had charged excessive commission and was unlicensed. It was considered that in order to circumvent the Queensland law the parties had made Hong Kong law the proper law of the contract. Hoare J held that their choice was not bona fide and that the law of Queensland was the proper law of the contract.

I am satisfied that the attempted selection of this [Hong Kong] law was for no other purpose than to avoid the operation of the Queensland law. Under all the circumstances, I conclude that the purported selection of the Hong Kong law was not a *bona fide* selection.¹²

In the second case¹³ Campbell J said, obiter:¹⁴

The question is not one of jurisdiction which is not raised but what law should be applied. If the parties in this clause have indicated an intention that the law of Hong Kong is to govern the whole of their contractual relations they have expressed a choice unconnected with the realities of the contract ...

It has been pointed out¹⁵ that whilst the decision reached in *Golden Acres* was not questionable the method of reaching it perhaps was. It could be argued that even if *Hoare J* had treated the parties' choice as effective the Queensland Act would not necessarily have been made inapplicable. The Queensland Act was obviously in force in Queensland and could have been applied notwithstanding the parties' choice of law.¹⁶

The use of the word "legal" by Lord Wright is ambiguous and a bona fide limitation has its difficulties. It may well have been a bona fide choice in a case such as *Golden Acre Ltd v Queensland Estates Ltd*.¹⁷ If the result of the parties' actions is to avoid some consequence that would otherwise apply to their detriment it is easy to suggest that their actions are not bona fide. If the court considers that the parties' choice is not bona fide it may, in a *Golden Acres* situation, decide that the issue is adequately covered by the public policy limitation.

Public Policy

Public policy, whether the forum's public policy or that of another jurisdiction, may overrule a choice of

¹² Ibid, at 385.

¹³ *Queensland Estates Pty Ltd v Collas* [1971] Qd R 75 (QSC).

¹⁴ Ibid, at 80.

¹⁵ D St L Kelly, "International Contracts and Party Autonomy" (1970) 19 ICLQ 701, 702.

¹⁶ See J L R Davis, "A Note" (1970) 44 ALJ 80.

¹⁷ [1969] Qd R 378 (QSC).

law clause.¹⁸ For example, in *Mitsubishi Corp v Alafouzos*¹⁹ Steyn J stated that “contractual disputes must be resolved in accordance with settled principles and rules of contract law but that exceptionally, the dictates of a recognised head of public policy, grounded on incontestable and fundamental moral considerations, operate to render defensible what would otherwise be enforceable contractual rights.”

In *Knyvett v Christchurch Casinos Ltd*²⁰ the New Zealand Court of Appeal noted that public policy, understood in a wider sense, may at times be better served by refusing to nullify a bargain save on serious and sufficient grounds.

Choice of Law unconnected with the Contract, Capricious Choices

Another possible limitation on party autonomy is the question of whether the parties may choose a governing law which is unconnected with their contract, in other words, may the parties make a capricious choice or alternatively, if not capricious, one totally unconnected with the contract? The matter was considered by Lord Wright²¹ in the *Vita Foods* decision. His Lordship said “connection with English law is not, as a matter of principle, essential.” There can be sound financial or commercial reasons for choosing a governing law that appears otherwise unrelated to the contract.²² A contrary view has been taken in other decisions.²³

From a practical point of view it may be noted that there are no reported decisions where the parties have made a wholly capricious choice as defined by the judge deciding the case. Possibly the answer lies in the fact that persons involved in international contracts, simply do not make such choices.

Overriding or mandatory laws

Mandatory rules, whether of the forum or the proper law, are rules that are considered sufficiently important to override choice of law clauses.²⁴ It has been suggested that Lord Wright in the *Vita Foods*²⁵

¹⁸ *Regazzoni v K C Sethia Ltd* [1958] AC 301 (HL).

¹⁹ [1988] 1 Lloyd's Rep 191, at 192.

²⁰ [1999] 2 NZLR 559, at 565 per Blanchard J (CA).

²¹ *Vita Food Products Inc v Unus Shipping Co Ltd* [1939] AC 277, at 290 (PC).

²² See for example *British South Africa Co v De Beers Consolidated Mines Co Ltd* [1910] 1 Ch 354, at 381 per Swinfen Eady J; *Rex v International Trustee for the Protection of Bondholders Aktiengesellschaft* [1937] AC 500, at 529 per Lord Atkin (HL); *British Controlled Oil Fields v Stagg* [1921] 127 LT 209; *Tzortzis v Monarch Line A/B* [1968] 1 All ER 949 (CA); *Akai Pty Ltd v The People's Insurance Co Ltd* (1996) 141 ALR 374 (HCA); cf *Re Claim by Helbert Wagg* [1956] Ch 323.

²³ In *Re Claim by Helbert Wagg*, *ibid*, at 341, Upjohn J said that the court will not necessarily regard the intention expressed by the parties “as being the governing consideration where a system of law is chosen which has no real or substantial connection with the contract looked upon as a whole.”

²⁴ See *Boissevain v Weil* [1950] AC 327 (HL).

²⁵ [1939] AC 277 (PC); Davies, Bell and Brereton, Nygh's Conflict of Laws in Australia (8th ed), LexisNexis 2010, at 402.

decision had this limitation in mind when he said that the choice must be legal.²⁶ In *Akai Pty Ltd v The People's Insurance Co Ltd*²⁷ a New South Wales corporation (Akai) entered into a contract of insurance with a Singapore Corporation (PIC) to protect Akai from bad debts. New South Wales had in force the Insurance Contracts Act 1984 (Cth), section 8 of which overrode any choice of law clause. Section 52 of the Act stated that where a provision was unhelpful to anyone (except the insurer) it was void. The contract stated that the policy was to be governed by the laws of England and any dispute arising from the policy was to be referred to the courts of England. A majority of the High Court of Australia held that section 8 required the court to ignore any of the evidence of the parties' choice of a governing law, including the reference to English Courts. The majority also held that section 52 precluded the evasion of the protection given by the Act through a choice of a foreign forum.²⁸ The forum is not concerned with the mandatory law of another jurisdiction unless its law is the proper law of the contract.²⁹

Generally a choice of law clause only binds the parties to the contract and not third parties

Following general principles of contract law in New Zealand a choice of law clause only binds the parties to the contract and not third parties.

Meaningless choice of law clauses will be disregarded or altered

In *Compagnie D'Armement Maritime SA v Compagnie Tunisienne de Navigation SA*³⁰ the contract contained a clause holding that the governing law would be the law of the flag of the vessel carrying the goods. A number of boats were used, involving the flags of France, Italy, Liberia, Sweden, Bulgaria and Norway. Lord Reid said the clause had failed and therefore the third rule of the proper law had to be used.³¹ The parties' intention was no longer relevant. It was decided that French law was the system of law with which the contract was most closely connected.

The law chosen must refer to a municipal system of law

After a period of uncertainty it seems clear that the courts require the parties to choose a definite municipal system of law to govern their contract. Thus parties may not choose public international law to govern. In *Amin Rasheed Shipping Corp v Kuwait Insurance Co*,³² Lord Diplock said that contracts were

²⁶ *Akai Pty Ltd v The People's Insurance Co Ltd* (1996) 141 ALR 374 (HCA).

²⁷ Ibid. See also F M B Reynolds, "Overriding Policy of the Forum: The other side of the Coin" [1997] LMCLQ 177; J Greene, "Party Autonomy in Choice of Law in Contract: Through the Lens of *Akai Pty Ltd v The People's Insurance Company*" (1997) 25 ABLR 330.

²⁸ *Akai Pty Ltd v The People's Insurance Co Ltd* (1996) 141 ALR 374, at 447-448 (HCA).

²⁹ *Vita Food Products Inc v Unus Shipping Co Ltd* [1939] AC 277 (PC); *Akai Pty Ltd v The People's Insurance Co Ltd* (1996) 141 ALR 374 (HCA).

³⁰ [1971] AC 572 (HL).

³¹ As discussed below.

³² [1984] AC 50, at 56 (HL).

incapable of existing in a legal vacuum:

They are mere pieces of paper devoid of all legal effect unless they are made by reference to some legal system of private law which defines the obligations...

The House of Lords explicitly rejected the notion of the internationalised contract governed by a set of general principles derived from public international law. It has also been suggested that parties are not at liberty to subject their contract to a legal system which is no longer in force, or to a draft of a foreign code or to a system which they have freely invented.

Parties are limited to the substantive law of the chosen law

In other words there is no renvoi.³³ This means that the law refers to the domestic law of a jurisdiction and not its conflict of laws rules which might refer the matter back or onto another jurisdiction.

The choice of law limited to certain issues

Party autonomy does not apply in the field of capacity to contract. Parties may not arrange their affairs in such a way as to confer capacity upon themselves by a choice of law clause. Dicey and Morris³⁴ state that an individual's capacity to contract is governed by the system of law with which the contract is most closely connected or by the law of his or her domicile or residence.

Parties may not choose a law to govern procedure

Parties are not allowed to choose a governing law to govern procedure. These questions are governed by the law of the forum (the *lex fori*). It would be quite impractical to allow otherwise.

Both parties must agree on the choice of law clause

If one party is in a weaker position there may be no true agreement on the choice of law.³⁵ This is a problem with standard form contracts, where the weaker party has no real choice in negotiating the contract.

Limitations concerning time of choice

Courts will not enforce a governing law clause whereby the law chosen is made to depend upon some future event.³⁶ In *The Iran Vojdan*³⁷ the choice of law clause stated that, at the option of the carrier, either

³³ This was affirmed in the *Amin Rasheed* decision, *ibid*.

³⁴ See Dicey & Morris, above n 2

³⁵ *Fricke v Isbrandtsen* 151 F Supp 465 (1957).

Iranian law or German law or English law could govern the contract. Bingham J held that as a matter of English law it was “clear and not disputed that this clause in the bill of lading is bad insofar as it envisages what may be called a ‘floating proper law’.” There must be a choice of law clause at the time of contracting if it is to be a valid clause that the courts will uphold. If a so-called floating proper law clause is made the courts will ignore it and apply the system of law which is most closely connected to the contract.

The parties may, however, provide for two proper laws, one to be applied in one event and the other if that event was negated. This has been considered to be good commercial sense. In *The Mariannina*,³⁸ for example, the parties had chosen arbitration in England pursuant to English arbitration law but if this was unenforceable then Greek law was to apply. The Court held that English law would apply. If for any reason it could not have applied then the Greek choice would have prevailed.

Finally, it is acceptable to choose one law to govern one part of the contract and another law to govern another aspect of it. This is called *depeçage*.

Inferred Choice

The intention of the parties may be inferred from the terms of the contract and the surrounding circumstances.³⁹

This second limb or rule applies again where the parties have actually made a choice. The difference between the first and second rule is that in the first situation the parties have expressly stipulated a choice of law to govern their contract and have inserted such a clause in their written agreement. In this second situation the parties, while having made a choice, have failed to articulate it. In the words of Dicey & Morris:⁴⁰

When the intention of the parties to a contract with regard to the law governing the contract is not expressed in words, their intention is to be inferred from the terms and the nature of the contract, and from the general circumstances of the case, and such inferred intention determines the proper law of the contract.⁴¹

³⁶ *The Iran Vojdan* [1984] 2 Lloyd’s Rep 380.

³⁷ *Ibid.*

³⁸ *Astro Venturoso Compania Naviera v Hellenic Shipyard SA (“The Mariannina”)* [1983] 1 Lloyd’s Rep 12.

³⁹ *R v International Trustee for the Protection of Bondholders Akt* [1937] AC 500, at 529 per Lord Atkin (HL).

⁴⁰ See Dicey & Morris, above n 2, Rule 180, Sub-Rule 2, 1182.

⁴¹ *Ibid.* See also *Re a Mortgage* [1933] NZLR 1512.

At one time it was thought that the problem of inferred choice could be solved by means of an implied term in the contract. However, it has been pointed out⁴² that this creates difficulties and Lord Reid held in *James Miller & Partners v Whitworth Street Estates (Manchester) Ltd*⁴³ that “the better view is now to apply a more objective test.” Similarly, Lord Wilberforce⁴⁴ in the same decision said:

In my opinion, once it was seen that the parties had made no express choice of law, the correct course was to ascertain from all relevant contemporary circumstances⁴⁵ including, but not limited to, what the parties said or did at the time, what intention ought to be imputed to them on the formation of the contract.⁴⁶

Thus, the rule here is that the parties’ choice still prevails if it can be discovered. The choice may be said to be implicit. The difficulty is to work out what law the parties had in fact had in mind when they entered their contract. As the litigants have not articulated their choice the court is in a difficult position. Traditionally, a number of presumptions existed which, if applicable to the contract under consideration, were used to assist the court in inferring a choice. Today, presumptions in general are now out of favour and indeed the development of some of them in recent years, as will be shown, has reduced their usefulness in the search for the parties’ choice.

Aids in Establishing Party Choice

Arbitration Clauses

Until the end of the nineteen sixties⁴⁷ the English courts had generally held that if the parties had agreed that arbitration should take place in a given country then the law of that country was intended to be the proper law of the contract.⁴⁸ Today an arbitration clause is just one factor to take into account when considering the issue.⁴⁹

Particular language as an aid to inferred choice

It has been suggested⁵⁰ that the use of terms such as “Act of God” or “Queen’s Enemies” in a contract may suggest that English law was intended to govern. These two phrases are, for example, meaningless in

⁴² *James Miller & Partners v Whitworth Street Estates (Manchester) Ltd* [1970] AC 583, at 603 per Lord Reid (HL).

⁴³ Ibid.

⁴⁴ Ibid, at 614.

⁴⁵ This test was applied in *JMJ Contractors v Marples Ridgway* (1985) 31 BLR 100 by Davies J.

⁴⁶ See also [1970] AC 583, at 611 per Viscount Dilhorne (HL).

⁴⁷ Until *Compagnie d’Armement Maritime SA v Compagnie Tunisienne de Navigation SA* [1971] AC 572 (HL).

⁴⁸ See eg *Hamlyn v Talisker Distillery* [1894] AC 202 (HL); *Spurrier v la Cloche* [1902] AC 466 (PC); *N.V. Kwik Hoo Tong Handel Maats-Chappij v James Findlay & Co* [1927] AC 604 (HL); *Tzortzis v Monarch Line A/B* [1968] 1 WLR 406 (CA); *Compagnie d’Armement Maritime SA*, ibid.

⁴⁹ *Astro Venturoso Compania Naviera v Hellenic Shipyard SA (“The Mariannina”)* [1983] 1 Lloyd’s Rep 12.

⁵⁰ See Dicey & Morris, above n 2, at 1184.

German law.

The use of an English standard form is another factor that may be taken into account. For example, in *Amin Rasheed Shipping Corporation v Kuwait Insurance Co* Bingham J, at first instance, said:

The use of an English standard form may be a powerful, even conclusive, indication that the parties intended to contract with reference to English law [...] But there is, in my judgment a factor here which cannot be ignored in assessing whether, and if so, how strongly, that inference should be drawn. The evidence in this case plainly establishes that this form of ruling policy, produced and developed on the London Insurance market, has achieved worldwide currency. Partly this is due to the long history, the great experience, the professional expertise and the high standing of that market, combined with a traditional dominance of London as a Commercial centre. Partly it is due to the process of imperial fertilization which has led to the reproduction of the *Marine Insurance Act 1906* in far corners of the globe [...] the more international and generally used the reference (i.e. the standard form) is the less specifically English it becomes.⁵¹

In *NZI Insurance New Zealand Ltd v Hinton Hill & Coles Ltd*⁵² Smellie J, referring to *Amin*, said that “it was not unusual for a ‘foreign’ Court to apply English law when interpreting particular kinds of insurance contracts which are couched in arcane language, derived from long usage in the London market, and for the interpretation of which reference must be made to established lines of English authority.” Dicey & Morris suggest that a contract making reference to, for example, a Scottish Statute, would presumably be taken to intend that it be governed by Scottish law even if its subject matter was situated in England and the person executing it be domiciled there.⁵³

Other matters have included the place where the contract is made⁵⁴ or the choice by the parties of a method of negotiating and financing a loan.⁵⁵ The fact that a form or wording of a contract has been approved or prescribed by the authorities of a given country or by the Head Office of a commercial undertaking with branches in a number of countries may be a pointer to the proper law.⁵⁶ The residence of the parties, occasionally the nationality of the parties, the nature and location of the subject matter of the contract and the currency of payment have likewise been considered. In *Club Mediterranee v Wendell*,⁵⁷ Cooke P said that, “Reading through the terms and conditions, one cannot avoid the conclusion that, taken

⁵¹ [1982] 1 WLR 961, at 968. See also *Armar Shipping Co v Caisse Algerienne d’Assurance et de Reassurance* [1981] 1 All ER 498 (CA); *Monterosso Shipping Co v International Transport Workers Federation* [1982] 3 All ER 841 (CA).

⁵² [1996] 1 NZLR 203, at 208 (HC). This was said in the context of staying proceedings.

⁵³ See Dicey & Morris, above n 2, at 1184, footnote 45 and cases cited eg *Re Pilkington’s Will Trusts* [1937] Ch 574; cf *Lindsay v Miller* [1949] VLR 13 (VSC).

⁵⁴ *Club Mediterranee v Wendell* [1989] 1 NZLR 216, at 218-219 (CA).

⁵⁵ *Re United Railways of Havana etc Warehouses Ltd* [1960] Ch 52 (CA) affirmed sub nom *Tomkinson v first Pennsylvania Banking & Trust Co* [1961] AC 1007 (HL).

⁵⁶ *Anspach & Co Ltd v CNR* [1950] 3 DLR 26, and other cases cited by Dicey & Morris, above n 2, at 1184, footnote 48.

⁵⁷ [1989] 1 NZLR 216, at 218-219 (CA).

as a whole this was a contract made by parties in New Zealand, to be paid for in New Zealand in New Zealand currency, and that the primary obligations under it must have been intended by the parties to have been governed by New Zealand law.”

Finally, the parties’ choice could be inferred from the theory that parties intend their contract to be valid. If there are two possible laws to choose from, then if a particular choice would render the contract void this choice is to be avoided. However, as with arbitration clauses all these considerations are merely indicators and must be considered in the light of all relevant circumstances. At the end of the day the assistance offered a judge is of little help. It has been said that “The only certain guide is to be found in applying sound ideas of business, convenience and sense to the language of the contract itself, with a view to discovering from it the true intention of the parties.”⁵⁸

The difficulties in establishing the parties’ choice when they have failed to articulate their wishes are obviously considerable. The lack of presumptions leaves the court without any really useful aids. However, the present English law is that the parties’ choice will be upheld even though it is not explicitly expressed. The court will consider all relevant circumstances but will not rely on any predetermined presumptions.

Law of Closest Connection

Definition

When the intention of the parties to a contract with regard to the law governing it is not expressed and cannot be inferred by the circumstances, the contract is governed by the system of law with which the transaction has its closest and most real connection.⁵⁹

This is sometimes referred to as the *Bonython* test after the case of that name. The judge puts him or herself in the place of the reasonable person and determines the proper law for the parties. (S)he does not attempt to attain the actual intention of the contracting parties because that is non-existent but “how a just and reasonable person would have regarded the problem.”⁶⁰

⁵⁸ *Jacobs v Credit Lyonnais* (1884) 12 QBD 589, at 601 per Bowen LJ (CA).

⁵⁹ See Dicey & Morris, above n 2, Rule 180, Sub Rule 3, 1190-1191. *Compagnie d’Armement Maritime SA v Compagnie Tunisienne de Navigation SA* [1971] AC 572, at 603 (HL); *Connell Dowell Constructors Ltd v Lloyd’s Syndicate* 396 [1988] 2 NZLR 256 (CA); *Club Mediterranee NZ v Wendell* [1989] 1 NZLR 216 (CA).

⁶⁰ *The Assunzione* [1954] P 150, at 176 per Singleton LJ.

Application of the Test

When applying this third test the judge may consider all the relevant circumstances. Two text book examples illustrate this rule. In *Rossano v Manufacturers' Life Insurance Co*⁶¹ the court held that Ontario law applied as the proper law to the transaction which involved three endowment policies, all of which were associated with Ontario. It was noted that the defendant company had its head office in Ontario, that the form of policy was based on the law of Ontario, that despite *Rossano*, an Egyptian national resident in Egypt, negotiating the contracts initially in Cairo, he would never have got the policies if the Ontario head office had not approved. Finally, the Court considered that generally a person contracting with a foreign insurance company does so because he has faith in it and the system of law under which it operates.

The *Assunzione*⁶² is the perfect example of a court adding up all relevant factors to determine the applicable law. A contract was made for the carriage of wheat from Dunkirk to Venice on board an Italian ship. The charterers were an organization of French grain merchants. The wheat was shipped under an exchange agreement between the French and Italian Governments (but the Italian ship owners did not know this). The contract was negotiated by correspondence between brokers in France and brokers in Italy. It was formally concluded in Paris in the English language and on an English standard form. Freight and demurrage was payable in Italian currency in Italy. The points in favour of applying French law were:

- The contract was made in France. It was headed "Paris, 1949";
- It was written in English but had a French supplement;
- The Bills of Lading were in French Standard form;
- The Charterers were a French firm acting for the French Government (although the Italians did not know this).

Matters favouring the application of Italian law were:

- The ship was Italian, flying an Italian flag;
- Italy was the place of performance where delivery was to be made;
- The Bills of Lading were assigned to Italian Consignees;
- The freight expenses were payable in Naples in Italian currency.

The English Court of Appeal had to use "a very delicate pair of scales"⁶³ in order to determine the proper law. It was unanimously held that Italian law was the proper law of the contract. The decisive factor was

⁶¹ [1963] 2 QB 352. Another example is *Offshore International SA v Banco Central SA* [1977] 1 WLR 399.

⁶² [1954] P 150.

⁶³ J H C Morris & D McClean, *The Conflict of Laws* (5th ed), London, Sweet & Maxwell, 2000, 324.

that both parties had to perform in Italy.⁶⁴

Dicey & Morris⁶⁵ use the term “system of law”. Generally the country with the closest connection will be the same as the system of law with the closest and most real connection but cases do exist where the difference could be important.⁶⁶ Country is also inappropriate, obviously, where there may be a number of different states with different laws.

It must be noted that the test can be difficult to apply.⁶⁷ This, in part, can be attributed to the fact that an international contract can be so international as to not have a close connection with any particular system of law.

Scope of the Proper Law

Capacity to contract is governed by the system of law with which the contract is most closely connected or by the law of the person’s domicile and residence.⁶⁸ The proper law governs issues of offer, acceptance, consideration⁶⁹ and consent.⁷⁰ Formal validity of an international contract is governed by the law of the jurisdiction where the contract is made or by the proper law of the contract. If it complies with either of those systems of laws it will be formally valid.⁷¹ Material or essential validity is also determined by the proper law.⁷²

Conclusion

The difficulties with the New Zealand approach would suggest that this is a suitable area of law for legislative reform which if it did eventuate would hopefully follow European law to make for uniform rules for conflict of laws issues.

⁶⁴ See also *Sayers v International Drilling Co* [1971] 3 All ER 163 (CA); *Brodin v A/R Seljar* [1973] SC 213 and *Coupland v Arabian Gulf Petroleum Co* [1983] 3 All ER 226 (CA).

⁶⁵ See Dicey & Morris, above n 2, Rule 180, Sub Rule 3, 1190, 1191.

⁶⁶ *James Miller & Partners v Whitworth Street Estates (Manchester) Ltd* [1970] AC 583 (HL).

⁶⁷ In *Sayers v International Drilling Co.* [1971] All ER 163, at 168 (CA), for example, Salmon LJ said “Sometimes it is said that the test – and it is a very useful test – is: what system of law has the closest and most real connection with the contract? My difficulty is that I can find very little clue in the contract as to what the parties intended, and very little indication that the contract has a very real or close connection with any particular system of law.”

⁶⁸ See Dicey and Morris, above n 2, at 1203; See also Nygh and Davies, above n 25 at 372-375.

⁶⁹ *Re Bonacina* [1912] 2 Ch 394 (CA).

⁷⁰ See Dicey and Morris, above n 2, at 1199.

⁷¹ *Ibid*, Rule 183, 1207.

⁷² *Ibid*, Rule 184, 1213.

The legal consequences of breach of government undertakings to investors

Asanga Gunawansa

National University of Singapore, 4 Architecture Drive, 117566

Singapore

Email: bdgasan@nus.edu.sg

Abstract:

For all countries, attracting foreign and private sector investment for development projects is a key strategy for nation building. As far as developing countries are concerned, there is high competition to attract investors. In order to encourage investors to choose their jurisdictions as investment destinations, developing country governments provide various investment protection undertakings. Legal consequences of breaching such undertakings can be severe as it is well established in international investment law that States should not breach promises made concerning the stability of investments. However, sometimes developing country governments take decisions that adversely affect the interests of the investors, without due consideration of the legal consequences. This paper examines the undertakings given by Sri Lanka to investors and the legal consequences of the breach of such undertakings.

Keywords:

Government undertakings, government guarantees, foreign investors, stabilization clauses, arbitration.

1. Introduction

It is generally accepted that attracting foreign investment is an important part of nation building. Given that of the approximately 195 countries in the world, more than two thirds are considered to be developing countries (IMF, 2009), it could be said that foreign investment is an integral part of the open and effective international economic system in which we live today and that it is a major catalyst to development. As far as developing countries are concerned, foreign investment is used mainly as a means of overcoming the shortage of domestic funds for development activities, especially in the provision of infrastructure facilities to people.

The benefits of FDI do not accrue automatically and evenly across countries, sectors and local communities (OECD, 2002). The popular belief that 'foreign direct investment is developed countries investing in developing countries' is unsubstantiated. The evidence indicates that more than 60 per cent of the FDI flows amongst the developed countries (Tesfachev and Sauv ant, 2005). Thus, there is immense competition among the large number of developing countries to attract investors. As a result, sound national policies and good international investment architecture matter for attracting FDI to a larger number of developing countries and for reaping the full benefits of FDI for development.

No investor would be interested in taking high risks and developing projects in countries which do not offer a safe investment environment. Therefore, although in theory project finance mechanisms widely available for developing projects are supposed to be non-recourse in nature because project revenue is the only source of repayment for investors, many developing countries continue to be burdened with government pay-back guarantees and other investment protection obligations to attract investors.

This paper examines the validity of government undertakings and guarantees given to investors to attract them. In particular, the paper looks at the investment protection guarantees given by the Government of Sri Lanka and the legal consequences of the breach of such guarantees.

2. Investment Guarantees

There are several ways in which countries offer investment guarantees to investors. The most straightforward manner is to provide investment protection in the investment agreement itself. However, it is common for countries to offer investment protection by specific statutory provisions and general undertakings, in investment policy statements and propaganda material used to attract investors. Some may argue that propaganda materials usually falls under the realm of “trade puff” in contract law, but there is a big difference between trade puffs of private companies promoting their goods and services, and investment guarantees made by public sector entities which are backed by government policy.

As far as Sri Lanka is concerned, it is common knowledge that the Board of Investment (BOI), which was established by Law No. 4 of 1978, is the “one stop shop” of the Sri Lankan government for attracting investors and facilitating investment projects. According to BOI, it performs its duties and functions as the “central facilitation point for investors” and it is the “utmost statutory body” which promotes Sri Lanka as a safe destination for investors (BOI, 2009). To encourage investors to invest in projects in Sri Lanka, the BOI makes the following promises to all investors (BOI, 2009a):

- i. “When you sign an agreement with the BOI, the provisions embodied in the agreement remain valued for the life of the enterprise. Successive Governments cannot change these provisions, ensuring a degree of stability that few other countries can offer or match.”
- ii. “The government has never defaulted nor requested rescheduling of any of its international obligations. Significantly this protection extends to foreign investors.”
- iii. “The safety of foreign investment is guaranteed through the acceptance by two third majority of Parliament of the Constitutional Guarantee of Investment Protection Agreements.”
- iv. “Sri Lanka is also a founder member of the Multilateral Investment Guarantee Agency (MIGA). This provides further safeguards against expropriation and non-commercial risk.”
- v. “Bilateral investment agreements supported by a constitutional guarantee provide strong protection for foreign investment for Sri Lanka.”
- vi. “Under article 157 of the country's constitution, the agreement enjoys the force of law and no legislative, executive or administrative action can be taken to contravene the provisions of a bilateral investment agreement otherwise than in the interests of national security.”
- vii. “A clause in the Sri Lankan constitution ensures the sanctity of the agreements. These agreements provide the following:
 - a. Protection against nationalisation.
 - b. Prompt and adequate compensation if required.
 - c. Free remittance of earnings, capital and business fees.
 - d. Settlement of disputes under the International Convention for the Settlement of Investment Disputes (ICSID).”

In addition to the above general representations, it should be noted that Section 26(1) of the BOI Law of Sri Lanka provides *inter alia* that any dispute relating to the interpretation of the provisions of an agreement entered into with the BOI for the purpose

of investing in Sri Lanka, or any rights the contracting parties will have under such agreement, shall, unless otherwise agreed to between the parties, be referred for settlement to the ICSID under the Convention on Settlement of Investment Disputes (CSID) between States and Nationals of other States of 1965¹. As the said section provides that it is applicable only when the parties have “unless otherwise agreed”, it is clear that BOI agreements can have dispute settlement clauses which allow parties to submit their disputes to other forums.

The provision made for resolution of disputes by submitting them to other forums than ICSID becomes relevant when an investor who has a dispute with the BOI is from a country that is not a party to the CSID. Further, as Section 17 of the BOI Law enables local investors to enjoy the same concessions granted under the BOI Law to foreign investors, the room for resolving disputes outside ICSID becomes important. The common practice among investors who invest in Sri Lanka is to insist on the inclusion in the BOI agreement of a dispute resolution clause that calls for the submission of disputes to international arbitration by a Panel of Arbitrators appointed under the arbitration rules of the International Chamber of Commerce (ICC)², or that the dispute be resolved by arbitration under the rules of the United Nations Commission on International Trade Law (UNCITRAL)³.

In addition to the aforesaid representations by the BOI and the provisions of the BOI Law, Article 157 of the Constitution of Sri Lanka - which is also referred to in the aforesaid representations of the BOI - provides as follows:

157. Where Parliament by resolution passed by not less than two-thirds of the whole number of Members of Parliament (including those not present) voting in its favour, approves as being essential for the development of the national economy, any Treaty or Agreement between the Government of Sri Lanka and the Government of any foreign State for the promotion and protection of the investments in Sri Lanka of such foreign State, its nationals, or of corporations, companies and other associations incorporated or constituted under its laws, such Treaty or Agreement shall have the force of law in Sri Lanka and otherwise than in the interests of national security no written law shall be enacted or made,

1 The government has ratified the Convention on Settlement of Investment Disputes, which provides the mechanism and facilities for international arbitration through the World Bank's International Centre for the Settlement of Investment Disputes (ICSID).

2 For ICC Rules (1998) see: http://www.iccwbo.org/uploadedFiles/Court/Arbitration/other/rules_arb_english.pdf

3 For UNCITRAL Rules (1976)

see: http://www.uncitral.org/uncitral/en/uncitral_texts/arbitration/1976Arbitration_rules.html

and no executive or administrative action shall be taken, in contravention of the provisions of such Treaty or Agreement.

The effect of Article 157 of the Sri Lankan Constitution is that it pledges constitutional protection to investors coming from countries which have entered into bilateral investment treaties with Sri Lanka⁴. The cumulative effect of this Article, the aforesaid representations by the BOI and the provisions of the BOI Law is that investors are given the assurance that Sri Lanka is a very safe investment destination and that the government facilitates and protects investments. Further, there is assurance that investments will not be taken away without adequate compensation.

3. Breach of Investment Guarantees

The legal consequences of the breach of government guarantees to investors can arise in three situations. The first is the government entity that has entered into an investment agreement with the investor breaching one or more of the guarantees given. The second is the government directly or indirectly expropriating the investment project. The third is the government annuls the contract on the basis that it is against public policy or any other ground valid under the applicable law of the land.

In the case of Sri Lanka, in all three situations above, the investor concerned can submit the dispute for arbitration by ICSID under Section 26 of the BOI Law, provided that no other mode of dispute resolution has been agreed to in the contract and that the investor is from one of the member countries of the CSID. If any other mode of dispute resolution has been agreed between the parties, for example submission of the dispute to arbitration under the ICC or UNCITRAL rules, then the investor can take the necessary initiative under such other mode.

It should be noted that the settlement of investment disputes and the rights of parties, especially of the investor, to submit the dispute for arbitration is not always clear cut. For example, when the validity of an investment contract is in doubt and the government has annulled the contract on the basis that it is against public policy or any other grounds applicable under the law of the land, there might be doubt as to the rights of an investor to seek relief against the government. This is because the agreement to arbitrate is usually found in the investment agreement itself.

⁴ Sri Lanka has entered into a number of Investment Protection Agreements with several countries including Belgium, China, Denmark, Egypt, France, Finland, Germany, Italy, Indonesia, India, Iran, Japan, Korea, Luxembourg, Malaysia, The Netherlands, Norway, Romania, Singapore, Sweden, Switzerland, Thailand, U.K. and U.S.A.

The recent decision by the Supreme Court of Sri Lanka in the case involving the “Waters Edge” Project⁵ is ideal to examine. The petitioners representing public interest complained of violation of their fundamental rights, due to the alienation of a land by the Urban Development Authority of Sri Lanka (UDA) to an investment project approved by the BOI under Section 17 of the BOI Law. The basis of the complaint was that the said alienation of land to a private investor was against public policy, as the land had been acquired for a “public purpose”.

Having examined the material submitted for their consideration, their lordships of the Supreme Court of Sri Lanka concluded that the fundamental rights of the Petitioners had been violated by several of the Respondents, including the former President, Chandrika Bandaranaike Kumaratunge. In the circumstances, their Lordships decreed that the entire Waters Edge Project be reverted to the ownership of the UDA. The Supreme Court also ordered the UDA to pay the investor a sum representing the cost of construction of the buildings as at the date of construction, excluding all other development on the said land, within 4 months of the judgment⁶.

Given that the Supreme Court Judgment was delivered approximately 7 years after the land had been alienated, the question that arises in the context of this paper is whether the investor of the Waters Edge Project has a legal right to claim compensation for the development it has done over the said period of seven years, on the basis that the investment protection guarantees promised by the government, in this case by the BOI, had been breached.

It is true that in the said case the Supreme Court ordered some compensation as noted above. However, under law of contract, an innocent investor is entitled to not only the cost of construction of buildings on the alienated land as at the date of construction, but the market value of the entire project at the time it reverted to the UDA, as well as expected future profits from the project. In the circumstances, it is useful to examine the law applicable to investment protection guarantees and the legal consequences of the breach of such guarantees.

5 SC/FR Application 352/2007.

6 According to the information available to the author, this payment is outstanding to date.

4. Applicable Law

4.1. *Validity of the Agreement to Arbitrate*

It is settled law, both under domestic common law and international law that an arbitration agreement survives the termination of the main contract (which contains an arbitration clause). In the English case of *Heyman v Darwins*⁷, the House of Lords, in laying down the scope of applicability of the arbitration clause in relation to the contract containing the clause, took the view that an arbitration clause is a written submission which is agreed to by the parties to the contract and should be treated as a standalone agreement to which the parties have submitted themselves to.

In India, in the case of *The Branch Manager, MAGMA Leasing & Finance Ltd v Potluri Madhavilata*⁸ the Supreme Court held that an arbitration clause is a collateral term in the contract relating to the resolution of disputes. Thus, even if the performance of the contract comes to an end on account of repudiation, frustration or breach of contract, the arbitration agreement would survive for the purpose of resolving disputes arising from or in connection with the contract. The US Supreme Court in *Buckeye Check Cashing Inc v Cardegna*⁹ acknowledged that the severability rule permits a court "to enforce an arbitration agreement in a contract which the arbitrator later finds to be void".

In international law too, it is an established rule that arbitration clauses, comparable to the stabilization clauses¹⁰ in their investment protection intention, are separate from the main agreement and cannot be unilaterally revoked by a state once it has submitted to arbitration (Ball, 1986; Bourquin, 1964; Mehren and Ko urides, 1981). In the *Libyan American Oil Company (LIAMCO) vs The Libyan Arab Republic* arbitration¹¹, the Arbitrator Mahmassani summed up the position in the following words:

"It is widely accepted in international law and practice that an arbitration clause survives the unilateral termination by the state of the contract in which it is inserted and continues in force after that termination. This is a logical consequence of the interpretation of the intention of the contracting parties, and appears to be one of the basic conditions for creating a favorable climate for foreign investment."

⁷ (1942) 1 ALL ER 337.

⁸ Civil Appeal 6399/2009, September 18 2009.

⁹ 546 US 460 (2005).

¹⁰ Stabilization clauses are contractual clauses that aim to guarantee that domestic laws with respect to investments will remain unchanged. They are used in investment agreements as a means of ensuring that the agreement concluded between a host state, or its agent and the investor will not be altered and its terms remain stable.

¹¹ Award of 12 April 1977, 62 I.L.R 140 (1977).

4.2. *Internationalization of Investment Agreements*

International investment law provides external legal protection in the event of ensuing conflict between a foreign investor and the host state. However, international law is not per se applicable to state/investor relations (Crawford and Johnson, 1986). Thus, in order to benefit from such protection, investors will often seek to internationalize the agreement. This is done primarily through the medium of a choice-of-law provision which places the relationship under the aegis of international law as the governing law of the agreement. When such provisions are combined with the choice of international arbitration in a neutral forum as the principal mechanism of dispute resolution, the investors have a form of binding dispute settlement which is not under the control of the state party¹².

There are academic writings to support the thesis that even in situations in which local law of a host state has been declared as the law applicable to a contract, general principles of law could be used to fill the lacunae (McNair, 1957; Schreuer, 2001). At its most extreme, this theory holds that the very nature of the foreign investment contract gives rise to the inference that it is subject to a supranational system of law¹³. It was concluded in the case of *Aminoil vs Kuwait*¹⁴ that even if the host state's law is the relevant law, it is still subject to the standards of international law, so that, if it conflicts with the standards of international law, it must give way to these standards.

It should be noted that the result of internationalization of investment contracts, even in circumstances where the contract specifies a local host state law as the applicable law, has significant consequences. For example, the host state cannot rely on its own law and argue that the agreement is a nullity. This rule was clearly stated in the *Sapphire International Petroleum Ltd. v. National Iranian Oil Company ("NIOC")* arbitration by Judge Cavin¹⁵.

As far as the application of legal principles relating to international investment law in Sri Lanka is concerned (in addition to what has been explained above concerning the application of such principles based on the theory of internationalization of contracts) it should be pointed out that Article 157 of the Constitution of Sri Lanka and the

12 In the arbitration of *Texaco Overseas Petroleum Co./California Asiatic Oil Co. v. Libya*, 17 I.L.M. 1, 8 (1977), the arbitration tribunal expressed the view that "the application of principles of Libyan law does not have the effect of ruling out the application of principles of international law, but quite the contrary: it simply requires us to combine the two in verifying the conformity of the first to the second."

13 *Revere Copper & Brass Inc. vs. OPIC* (1978) 17 ILM 1321.

14 (1982) 21 ILM 976. Also see, *SPP vs Egypt* (1982) 22 ILM 752.

15 (1963) 35 ILR 136.

undertakings given by the BOI to investors, make it clear that the general principles of law relating to protection of investment and compensation have become part and parcel of the law of Sri Lanka. It should also be pointed out that the purpose of Section 16 of the BOI Law is to provide all investors the same status when they enter into contracts under the law. Thus, there is no distinction between foreign investors and local investors. All investors entering into BOI agreements are entitled to equal protection of the law and the benefits and protection guaranteed by the government.

4.3. Consequences of Breach of Government Undertakings

The relative stability of key investment conditions responsible for the economic and financial performance of an investment venture is at the heart of investor concerns when negotiating an investment contract with a public sector entity. It is particularly so for projects where duration and risk exposure are long, capital investment is intensive, and project risk is acute. Therefore the extent to which the relevant public sector entity undertakes to protect the investment would be at the center of both the negotiation of a specific contractual regime and the proper design of a project structure. Other relevant factors include investment recovery, ability to maintain contractual and proprietary rights and titles (in particular after discovery and development), ability to sell, ability to retain and repatriate foreign exchange earned, and ability to operate the project under reasonably foreseeable conditions, consistent with international standards. In the circumstances, stabilization guarantees provided by the government by direct contractual undertakings or by way of legislation and policy statements are a key feature of government-investor relationship.

It is settled law that where governments or state entities unduly interfere with the contractual rights of the investors, the latter should be adequately compensated. This principle flows from the most important United Nations General Assembly (UNGA) resolution on permanent sovereignty, UNGA Resolution 1803 of 1962. This resolution adopted language declaring that investors "shall be paid appropriate compensation ... in accordance with international law" where their property, including by inference their contract rights, has been taken. Further it adopted that disputes over compensation "should be made through arbitration or international adjudication" where the investor and state have so agreed¹⁶. Furthermore, Article 3 of the said resolution affirmed that host state promises to investors were legally binding.

In the circumstances, it is important to note that as long as an investor has access to a neutral, authoritative adjudicator, he would almost certainly be awarded meaningful damages in the event that a host country breached its promise to protect the investment. This is because of the almost universal recognition of the principle of "*pacta sunt*

16 G.A. Res. 1803 (XVII), U.N. Doc. A/5217 (Dec. 14, 1962), Article 4.

servanda”, which in its most common sense refers to private contracts, stressing that clauses contained in them are law between the parties, and implying that non-fulfillment of respective obligations is a breach of the pact. With reference to international agreements, the principle provides that “every treaty in force is binding upon the parties to it and must be performed by them in good faith”¹⁷.

When an investment agreement has been negotiated and agreed by a host state that guarantees contractual stability under some general legislation, the promised stability is a factor taken into consideration in investment dispute arbitrations in ascertaining when compensation is due, and in determining the quantum of compensation, when that state subsequently revokes or ignores its legislative promise¹⁸. It follows that the unilateral amendment by the state of the contract and/or the guaranteed regulatory and legislative protection would immediately and directly engage the responsibility of the state to compensate the investor. This is because such an act would be considered illicit by its very nature, without further need for the state to be in breach of the customary international law conditions relating to public purpose, non-discrimination, and payment of adequate compensation (Wehberg, 1959; Walde, 1994).

In these circumstances, in practical terms, arbitral tribunals would give full effect to the undertakings given to an investor and, at a minimum, order full compensation to investors injured, due to the breach of such undertakings. In this connection it is worth to note that in the *Texaco Overseas Oil Petroleum Co./California Asiatic Oil Co. v. Libyan Arab Republic* oil industry arbitration¹⁹, which involved an issue concerning the validity of a stabilization clause, the sole arbitrator stated that:

“There is no doubt that in the exercise of its sovereignty a State has the power to make international commitments....There is no value to dwell at any great length on the existence and value of the principle under which a State may within the framework of its sovereignty, undertake international commitments with respect to a private party. This rule results from the discretionary competence of the State in this area....The result is that a State cannot invoke its sovereignty to disregard commitments freely undertaken through the exercise of this same sovereignty, and cannot through measures belonging to its internal order make null and void the rights of the contracting party which has performed its various obligations under the contract.”

17 Vienna Convention on the Law of the Treaties (1980), Article 26.

18 See The “Pyramids” Case, 8 ICSID-Rev.-F.I.L.J. at 231. Also see *Nykomb Synergetics Tech. Holding AB v. Latvia*, Award, § 4.3.1 (Stockholm Chamber of Com. Inst. Dec. 16, 2003), 2003 WL 24045555 (APPAWD).

19 17 I.L.M. 1, 8 (1977). See also *RCA v. China*, 30 Am. J. Int’l L. 535, 540 (1936).

In another oil industry arbitration, namely *Aramco v. Saudi Arabia*²⁰, the tribunal came to a similar conclusion with regard to the binding force of the stabilization clause under international law. The decision concluded that:

“By reason of its very sovereignty within its territorial domain, the State possesses the legal powers to grant rights [by] which it forbids itself to withdraw before the end of the concession, with the reservation of the Clauses of the Concession Agreement relating to its revocation. Nothing can prevent a State, in the exercise of its sovereignty, from binding itself irrevocably by the provisions of a concession and from granting to the concessionaire irrevocable rights. Such rights have the character of acquired rights.”

This trend toward arbitral enforcement of state contracts was further confirmed by the 1963 award in *Sapphire International Petroleum Ltd. v. National Iranian Oil Company* (“*NIOC*”). On the issue of whether state promises to investors should be recognized as binding, with breaches subject to a duty to compensate, the sole arbitrator left no doubts²¹:

“It is a fundamental principle of law, which is constantly being proclaimed by international courts, that contractual undertakings must be respected. The rule pacta sunt servanda is the basis of every contractual relationship.”

In *Libyan Am. Oil Co. v. Gov't of the Libyan Arab Republic*²², where the claim alleged breaches of a concession contract by the Libyan government without payment of compensation, the arbitrator concluded that the action of the state party amounted to repudiation of the contract in apparent violation of so-called “stabilization” clauses through which Libya had expressly promised not to change the legal framework in ways adverse to the investors' interests.

In the *El Triunfo* case²³ the majority of the arbitrators found that the investor concerned was “entitled to compensation for the result of the destruction of the concession and for appropriation of the property belonging to the company”. In another arbitration dealing with the issue of investment protection, the arbitrator held that the government taking over the investment project, “even on the best of grounds, works injustice” to the

20 27 I.L.R. 117, 168 (1963); see also *Mobil Oil Iran Inc. v. Islamic Republic of Iran*, 16 Iran-U.S. Cl. Trib. Rep. 3, 64-65 (1987), and *Phillips Petroleum Co. Iran v. Islamic Republic of Iran*, 21 Iran-U.S. Cl. Trib. Rep. 79, 88 (1989).

21 35 I.L.R. 136 (1963) at 172 – 173.

22 20 I.L.M. 1 (1981). Also see *Texaco Overseas Petroleum Co./California Asiatic Oil Co. v. Gov't of the Libyan Arab Republic*, 17 I.L.M. 1 (1978), and *BP Exploration Co. (Libya) Ltd. v. Gov't of the Libyan Arab Republic*, 53 I.L.R. 297 (1979).

23 *Salvador Commercial Company v. Salvador*, 1902 Foreign Relations of the United States 838.

investor, and that the government ought to make compensation for the injury inflicted²⁴. The permanent Court of International Justice took a similar view in the *Chorzow Factory case*²⁵. In the circumstances, it could be said that it is well established by international dispute resolution forums that investors who have suffered due to breach of investment protection guarantees given by governments are entitled to just compensation.

The concept of right to compensation of investors, in the event of undertakings given to them in the form of guarantees being breached by governments, is recognized also by domestic laws. For example, the provisions of the BOI Law in Sri Lanka and Article 157 of the Constitution which have been considered earlier evidence the fact that there is statutory recognition of this concept in Sri Lanka. In addition to statutory recognition, there is also judicial recognition of this concept in most common law countries. For example, the Supreme Court of India in several cases has recognized the right to property of investors, wherein it has held that just and reasonable compensation should be provided in light of the market value of the property in case of acquisitions by the state²⁶. This is in consonance with the fair and equitable standard in international investment law which has been dealt with above, and thus assures the same protection to the investor as expected in International law.

As far as the quantum of compensation is concerned, in the arbitration between *Karaha Bodas Company and Perusahaan Pertambangan Minyak dan Gas Bumi Negara and PT PLN*²⁷, an arbitration held under the UNCITRAL Rules, the arbitration tribunal concluded that the goal of compensation ought to be to leave the investor in the same position it would have been in had the property not been taken, and that the principal guideline for calculating compensation is fair market value (FMV) of the property. The Iran-United States Claims Tribunal in the case concerning the *American International Group, Inc./American Life Insurance Company and the Islamic Republic of Iran/Central Insurance of Iran*²⁸, calculated the award based on FMV to include the net present value of future earnings under the conditions at the time of the taking. In *Asian Agricultural Products v. Sri Lanka*²⁹ the arbitration tribunal held that “interest becomes an integral part of the compensation itself, and should run consequently from the date when the State’s international responsibility became engaged”.

24 Shufeldt Claim, USA v. Guatemala, 2 UN RIAA 1079.

25 Judgment No. 13, September 13, 1928, P.C.I.J. Publ., Ser. A. No. 17, p 46.

26 See *Bihar v. Kameshwar Singh*, AIR 1952 SC 252; *K Kochhuni v. State of Madras*, AIR 1960 SC 1080; *Sajjan Singh v. State of Rajasthan*, AIR 1965 SC 845.

27 Award made on 18 December 2000. For a summary of the award see *Karaha Bodas Company L.L.C. vs Perusahaan Pertambangan Minyak dan Gas Bumi Negara and PT PLN* 335 F.3d 357.

28 19 December 1983.

29 Final award of 27 June 1990 in ICSID case no. ARB/87/3.

Since the case involving the Waters Edge project was mentioned earlier, one more issue that needs addressing here is whether the investor has a right for compensation when an investment contract has been held to be void by a court of law. In this connection, it should be noted that ICSID has held that a court of a country should be considered as an organ of the State and thus the State is responsible for actions of the court if such actions result in an unfair outcome affecting an investor³⁰. Thus, one could argue that in the said Waters Edge case, even though court has concluded the alienation of the land by the UDA is bad in law, the investor, if innocent of any participation in the *ultra vires* process by which the land is alleged to have been alienated, is entitled for just compensation for the losses it has suffered as a result of losing its investment after approximately seven years.

5. Conclusion

Recent trends in the provision of infrastructure development indicate that the private sector is playing an increasingly important role in the development of projects. This trend has mainly arisen out of a necessity for the development of infrastructure to be undertaken at a rate that responds to the modernization needs of countries and growing demand for infrastructure services. In order to attract investors countries, specially developing nations, are in the habit of offering strong investment protection guarantees. Some such guarantees offered by the government of Sri Lanka have been examined in this paper.

Governments may sometimes consider it within their powers to terminate investment agreements even if such actions result in the breach of investment protection guarantees. Mostly, such decisions may be taken by the public sector entities who have given the undertakings to protect the investments. Sometimes, some other organ of the state may take the decision which adversely affects the investor. Sometimes, such decisions may be taken for reasons that are justifiable in the eyes of the public. Sometimes they may be taken on the basis that an investment contract is bad in law or that the concessions given to an investor are contrary to public policy. It is important in such situations for the relevant parties to understand the legal consequences of the decision they are making. As explained in this paper, it is well established in law that investors have a right to compensation in the event of breach of investment guarantees given to them.

³⁰ The *Loewen Group, Inc and Raymond L. Loewen v. United States of America*, Case No. ARB (AF)/98/3 award of June 26, 2003.

References

- Ball, M., (1986), 'Structuring the Arbitration in Advance: The Arbitration Clause in an International Development Agreement', In: Julian Lew, *Contemporary Problems in International Arbitration*, Martinus Nijhoff Publishers, pp 309-11.
- Board of Investment (BOI), (2009), 'About us', http://www.boi.lk/2009/about_us.asp, viewed: 30/06/2010.
- Board of Investment, (2009a), 'Why Invest in Sri Lanka', <http://www.boi.lk>, viewed: 30/06/2010.
- Bourquin, M. (1964), 'Arbitration and Economic Development Agreements', In: Southwestern Legal Found, *Selected Readings on Protection by Law of Private Foreign Investments*, pp 99-104.
- Crawford, J., and Johnson, W. (1986), 'Arbitrating with Foreign States and Their Instrumentalities', *International Financial Law Review*, Volume 5, p 11
- McNair, L. (1957), 'The General Principles of Law Recognized by Civilised Nations', *British Year Book of International Law*, Volume 33, pp 7 - 10.
- Mehren, R.B., and Kourides, P.N. (1981), 'International Arbitration Between States and Foreign Private Parties: The Libyan Oil Nationalization Cases', *American Journal of International Law*, Volume 75, p 476.
- International Monetary Fund (IMF). (2009), 'World Economic and Financial Surveys', <http://www.imf.org/external/pubs/ft/weo/2009/02/weodata/groups.htm#ae>, viewed: 05/07/2010.
- Organisation for Economic Development and Co-operation (OECD). (2002), 'Foreign Direct Investment for Development, Maximizing Benefits, Minimizing Costs', <http://www.oecd.org/dataoecd/47/51/1959815.pdf>, viewed: 05/07/2010.
- Schreuer, C. (2001), *The ICSID Convention: A Commentary*, Cambridge.
- Tesfachew, T., and Sauvart, K.P. (2005), 'Making FDI Work for the Poor', *Development and Cooperation*, Volume 32, p 198.
- Wehberg, H. (1959), 'Pacta Sunt Servanda', *American Journal of International Law*, Volume 53, pp 775-786.
- Walde, T. W. (1994), *Stabilising International Investment Commitments: International law Versus Contract Interpretation*, Centre for Petroleum and Mineral Law and Policy, Scotland, pp 23-26.

Children's citizenship and housing: the dilemma of property law

Dr Simon Hoffman

Swansea University

s.hoffman@swan.ac.uk

Abstract: *Housing is fundamental to citizenship providing a base for individuals to develop attachment to community, and to engage with others and with community life. It is accompanied by a host of citizenship rights and obligations. This paper will discuss the housing rights of independent minor children in the UK from the perspective of children's citizenship. It is argued that housing law supports children's citizenship but that this is undermined by principles of property law affecting minors.*

Introduction

The notion of citizenship is 'variable, contextual and often contested'.¹ As a concept citizenship is complex and multi-dimensional, made up of 'building blocks' which include: rights; responsibilities; status; competence; dignity; social membership; independence; dependence; and, participation.² According to Marshall, the status of citizen guarantees equality with respect to rights and duties: 'he saw progress toward full citizenship as a measure of society's aspirations for social justice'.⁴ Marshall analysed citizenship as a triad of civil, political and social rights – the latter taking place as entitlements within state welfare regimes.⁵

¹ Doek, J.E., Foreword, in Antonella Invernizzi, and Jane Williams (eds.), (2008), *Children and Citizenship*, (London: Sage), xii. Most straightforward is formal legal citizenship which denotes membership of a nation-state.

² Lister, R., Unpacking Children's Citizenship, in Invernizzi and Williams, *ibid.*

³ Marshall, T.H. (1950), *Citizenship and Social Class*, (Cambridge: Cambridge University Press), p.18.

⁴ *Ibid.* Discussed by, Powell, M., 'Equality and Marshallian citizenship: why E does not equal MC', in Scott Greer (ed.), (2009), *Devolution and social citizenship in the UK*, (Bristol: Policy Press), pp.23-38.

⁵ Dean, H., (2002), *Welfare Rights and Social Policy*, (Harlow: Prentice Hall). Marshall's notion of social citizenship is often argued as the 'core' of the welfare state: Esping-Anderson, G., (1990), *The three worlds of welfare capitalism*, (Cambridge: Polity Press), p.21. Marshall also identified obligation as an aspect of citizenship: the benefits of citizenship should be accompanied by 'a lively sense of responsibility toward the welfare of the community', see, above n.4, p.41 (in the context of labour disputes). What Marshall meant by duty is unclear, perhaps an 'obligation to give' (see: Marshall, T.H., (1981), *The Right to Welfare*, (London: Heinemann), p.92), although others make connections between entitlements and 'corresponding' duties, see: Roche, M., (1992), *Rethinking citizenship: Welfare, ideology and change in modern society*, (Cambridge: Polity Press), p.117. This aspect of citizenship is not discussed in this paper but will be developed in future drafts.

Citizenship may be seen as comprising a 'set of rights that come with belonging to a political community', including social benefits.⁶ In contemporary analysis new ways of thinking have emerged to enrich the debate on the substantive content of citizenship. Demands for recognition – distinct from redistribution – have drawn attention to participation rights,⁷ but also to the notion of citizenship as a practice, to draw attention to issues of inclusion and participation.⁸ Housing is a key-building block in the construction of citizenship, as an aspect of the practice of citizenship. In addition to its function as a place of refuge and safety ('well-being'⁹) housing provides a centre for independent living to facilitate individuals in the exercise of social rights such as gaining and maintaining employment; it provides social status; and, enables individuals to put down roots, to develop a sense of belonging to community, and, to engage with and participate in society.¹⁰

Citizenship provides a lens through which to examine the position of children in society. The concept of children's citizenship has emerged as a way of thinking about children as rights-holders and as participants in society and community.¹¹ In the UK statute address the housing needs of minor children¹² providing a number of safeguards against homelessness: children are entitled to have their housing needs taken into account by statutory welfare service providers, and in some cases to be provided with accommodation. The Children Act 1989 (CA 1989) and the Housing Act 1996 (HA 1996) both provide a route to housing for independent minor children (IMC) – those aged 16 or 17 years-old whose housing needs are not met by family resources. This paper will first discuss the issue of children's citizenship before considering the housing and property rights of IMC from a perspective which pays attention to these issues.

⁶ Scott L. Greer and Margitta Matzke, 'Introduction: devolution and citizenship rights', in Greer, S.L., (ed.), (2009), *Devolution and social citizenship in the UK* (Bristol: Policy Press), p.2.

⁷ Honneth, A., (1995), *The Struggle for Recognition: The Moral Grammar of Social Conflicts*, (Cambridge: Polity); Fraser, N. and Honneth, A., (2003), *Redistribution or Recognition? A Political-Philosophical Exchange*, (London: Verso).

⁸ Oldfield, A., (1990), *Citizenship and Community, Civic Republicanism and the Modern World*, (London/New York: Routledge).

⁹ Buyse, A., (2006), 'Strings attached: the concept of "home" in the case law of the European Court of Human Rights', 3 *European Human Rights Law Review*, pp.294-307; Lawrence, R.L. (ed.), (2004), 'Special Issue: Housing, Health and Well-Being', *Reviews on Environmental Health*, Vol.19, Nos.3-4.

¹⁰ Fox, 2002; DCLG, 2007. As an example responsabilization (above n.5) housing is accompanied by obligations: to pay rent; and, for conduct.

¹¹ Above n.1, Invernizzi, and Williams (eds.), x.

¹² In the UK the age of majority is 18 years:

Children and citizenship

The (contested) nature of citizenship is made more complex in the case of children as any discussion will need to take into account the developmental and progressive nature of childhood.¹³ Issues of capacity and competence arise; in particular as the status of dependent child has the consequence of exclusion from certain citizenship rights.¹⁴ Children are often denied direct access to the benefits of citizenship as a consequence of legal structures for delivery of social entitlements. For minor children, in the UK those aged under-18, social citizenship rights are often 'by-proxy'; secured through the agency of parents (or those with parental responsibility).¹⁵ Age criteria are applied to determine when children are permitted access social security and welfare benefits as individual claimants.¹⁶ It has been observed that there has been a 'continual extension of the period of economic dependence previously associated with childhood',¹⁷ suggesting a limited view of children's citizenship in public policy.¹⁸ However, children are different from adults in a variety of ways, physically and intellectually, which may provide an explanation for the by-proxy realization of children's citizenship rights via the agency of the family.¹⁹

The United Nations Convention on the Rights of the Child (UNCRC) sets out a number of rights guaranteed for children aged under-18 years old.²⁰ These extend to all children irrespective of capacity; what remains at issue is the age at which a child is able to exercise

¹³ Marshall's account provides a limited view of children's citizenship: children are seen as 'citizens in the making', see: above n.3, p.16.

¹⁴ Cohen, E.F., (2005), 'Neither seen nor heard: children's citizenship in contemporary democracies', *Citizenship Studies*, 9(2), pp.221-40; Alderson, P., 'Rights of children and young people', in Anna Coote (ed.), (1992), *The Welfare of Citizens*, (London: IPPR/Rivers Oram Press), pp.153-80; Hill, M. and Tisdall, K., (1997), *Children and Society*, (Harlow: Prentice Hall).

¹⁵ Jones, G., 'Youth, Citizenship and the Problem of Dependence', Invernizzi and Williams, pp.97-107, p.99, above n.1; Jones, G. and Wallace, C., (1992), *Youth, Family and Citizenship*, (Buckingham: Open University Press). In the UK assumptions of child dependency on parents were re-enforced following a report into social security in the mid-1980s: Fowler Report, (1985), *The Reform of Social Security*, Cmd. 9518. Although the same report found that some under-18s were not dependent on their parents.

¹⁶ Age is used as a criterion for determining rights as well as duties: at times these appear arbitrary. For example: a 10 year-old may be held liable for a criminal offence; 16 years-olds may marry, smoke, and leave school, but cannot vote or buy alcohol; a 17 year-old may drive a car but cannot make a will in case they are killed in a car accident.

¹⁷ Above n.15, p.97.

¹⁸ Ibid, Jones also points out that at times policy has seen the removal of entitlement re-enforcing dependence, as in the case of removal of welfare benefit entitlement to 16/17 year-olds in 1989, pp.102-103.

¹⁹ Fox Harding, L., (1997), *Perspectives in Child Care Policy*, (London: Longman), pp.123-127.

²⁰ This upper age-limit is itself subject to the age of majority in signatory states where majority is set at a younger age (Article 1).

autonomous decision-making in pursuit of the rights guaranteed.²¹ In this respect Article 5 of the UNCRC requires States Parties to ‘respect the rights and duties of parents ... to provide, in a manner consistent with the *evolving capacities of the child*, appropriate direction and guidance *in the exercise by the child of the rights*’ (my emphasis). Article 12 of the UNCRC requires that children should not be treated as ‘mere passive recipients of adult protective care’, but rather as active participants in decision-making processes that affect their lives. Whilst this guarantees participation it does not secure autonomy of decision-making.²² What is suggested by the notion of ‘evolving capacities’ is an individuated approach to determine the extent of children’s participation and autonomy in pursuit of rights guaranteed by the UNCRC. This is significant to any discussion on children’s citizenship as the assumption and exercise of rights is a key component in the construction of citizenship. In this context, the obligation to take account of child’s evolving capacity is a ‘participatory or emancipatory concept’ which has the effect of ‘transferring rights from adults to the child in accordance with their level of competence’.²³ It is an approach which contrasts with the system of statutory age-limits to determine a child’s competence to exercise rights.²⁴ The concept of a child’s evolving capacity for self-determination is one which the judiciary has shown itself willing to embrace. The case of *Gillick* provides an example: the House of Lords held that competent children, those with sufficient intelligence and understanding, should be permitted responsibility for decisions which affect their own lives.²⁵

Children’s housing rights

The United Nations Declaration of Fundamental Rights,²⁶ and the International Covenant on Economic Social and Cultural Rights²⁷ both refer to ‘adequate’ housing as a fundamental human right without age restriction. Principle 4 of the United Nations Declaration on the Rights of the Child proclaims that the child shall have the right to adequate housing.²⁸ Article 27(1) of the UNCRC recognizes the right of every child to a ‘standard of living adequate for the child’s physical, mental, spiritual, moral and social development’. Article 27(3) provides that ‘States Parties...shall take appropriate measures to assist parents and others responsible for the child to *implement this right* and shall in case of need provide material assistance and

²¹ Lansdown, G., (2005), *The Evolving Capacities of the Child*, (Florence: UNICEF Innocenti Research Centre), p.ix.

²² Lockyer, A., ‘Education for Citizenship: Children as Citizens and Political Literacy’, in Invernizzi and Williams, pp.20-31, pp21-22, above n1.

²³ Above n.21, x.

²⁴ *Ibid*, xi.

²⁵ *Gillick v West Norfolk and Wisbech Health Authority* [1985] 3 WLR 830.

²⁶ UDHR, Article 25.

²⁷ ICESCR, Article 11.

²⁸ Declaration of the Rights of the Child, proclaimed by General Assembly resolution 1386(XIV) of 20 Nov. 1959, <http://www.cirp.org/library/ethics/UN-declaration/>

support programmes, particularly with regard to... *housing*' (my emphasis).²⁹ In broad terms Article 27 of the UNCRC provides recognition of every child's right to adequate housing but also draws attention to the particular role of the family in implementing the right, and the need for State support for parents to provide housing for their children. Whilst in the UK housing will ordinarily be delivered via their parents (or guardian) not all children will be adequately housed in the parental home which may not be available due to family economic circumstance; or there may be good reason for a child to leave home; or, parents may refuse to act as agent for a child's housing rights. It is apparent that whilst the UNCRC places some emphasis on the family as the primary provider of housing for children, Article 27(1) provides recognition of the individual child's right to an adequate standard of living, including housing, which extends to IMC.

The UK government has sought to address the issue of children's housing entitlements through statute. S17 of the CA 1989 places a duty on a local authority to safeguard and promote the welfare of children in need within their area.³⁰ Responsibility for dealing with a child in need falls on the social services department (SSD) of a local authority.³¹ An SSD will meet its s17 duty through the provision of services, which may include accommodation.³² S17 does not provide a child with a specific right to be accommodated, rather it is a 'target duty owed to the whole community rather than to individual children.'³³ S20 of the CA 1989 provides a more explicit housing duty. An SSD is under a duty to provide accommodation for any child in need within their area who appears to them to require accommodation as a result of: [1] there being no person with parental responsibility; [2] the child being lost or abandoned; or, [3] the person who has been caring for the child being prevented from providing suitable accommodation or care.³⁴ A further housing duty arises a child aged

²⁹ The UNCRC contains a number of provisions setting out rights having a direct bearing on housing, these include: the right to protection from all forms of physical or mental violence, injury or abuse, neglect or negligent treatment, maltreatment or exploitation, including sexual abuse (Article 19), all of which may result from homelessness; the right to special protection and assistance in cases where a child is temporarily or permanently deprived of his or her family environment (Art. 20); the right to the highest attainable standard of health (Art. 24) – health is intimately related to good housing. These rights are indivisible from each other and provide a useful tool for advocating children's housing rights: COHRE, (2006), *Defending the Housing Rights of Children*, (COHRE: Geneva), pp.17-18.

³⁰ S17(1). A child is a child in need if disabled or unlikely to achieve or maintain, or to have the opportunity of achieving or maintaining, a reasonable standard of health or development or their health or development is likely to be significantly impaired, or further impaired, without the provision of services by a local authority.

³¹ Local Authority Social Services Act 1970, Schedule 1.

³² S17(6).

³³ *R (on the application of M)(FC) v London Borough of Hammersmith and Fulham* [2008] UKHL 14, para.17.

³⁴ [1]-[3], s20(1).

sixteen or older if the SSD considers that the child's welfare will be seriously prejudiced if does not provide accommodation.³⁵

In addition to its duties under the CA 1989 a local authority, acting as a local housing authority (LHA), may be required to provide accommodation to an IMC under Part VII of the HA 1996. This places a duty on an LHA to assess eligibility for permanent housing in cases where a person is deemed homeless or threatened with homelessness.³⁶ Where it is determined that a person is eligible for permanent housing this is usually provided in the form of social housing (LHA or Registered Social Landlord).³⁷ Social housing is rationed amongst homeless households by reference to need, and it is only those who fall into a 'priority need' category that may be entitled to permanent housing under the HA 1996.³⁸ Priority need categories under the 1996 Act include homeless 16 and 17 year olds.³⁹

Housing and children's citizenship

It is significant that under s20 of the CA 1989 and Part VII of the HA 1996 the link between the implementation of rights and the agency of the family is severed. Legislation which provides an individual IMC with a right to be considered for housing, or a right to housing, secures for that child a social benefit as an aspect of citizenship. This is underpinned by the assumption of evolving capacity but undermined in practice as LHAs are often reluctant to provide tenancies to young people under the age of 18. The reasons for this include an ongoing debate over the capacity of minors to hold a legal tenancy,⁴⁰ an issue considered in the case of *Re Alexander-David*,⁴¹ but also because of assumptions that inform the delivery of housing services to IMC.

The dilemma of property law

In *Re Alexander-David* a child aged 16 years had been provided with temporary accommodation by the LHA to allow her application for permanent housing under the HA 1996 to be determined. The LHA purported to grant a legal tenancy which, in the circumstances of the case, took effect as a trust of land with the tenant as beneficiary and the local authority as trustee. This situation was brought about by operation of s1(6) of the Law of Property Act 1925 (the 'LPA 1925') which prevents a minor from holding a legal estate in

³⁵ S20(3).

³⁶ S184.

³⁷ The duty arises under s193 of the HA 1996.

³⁸ S190 and s193.

³⁹ Homelessness (Priority Need for Accommodation) (England) Order 2002/2051, and, Homeless Persons (Priority Need) (Wales) Order 2001/607.

⁴⁰ <http://www.yjb.gov.uk/en-gb/practitioners/Accommodation/Housingoptionsforyoungpeople/>

⁴¹ [2009] EWCA Civ.259.

land (in this case a tenancy), and paragraph 1(1) of Schedule 1 of the Trusts of Land and Appointment of Trustees Act 1996 (TOLATA 1996), which imposes a trust of land in circumstances where it is impossible to grant a legal estate. When the local authority served a notice to quit terminating the tenancy and then sought possession the child tenant successfully argued in her defence that the NTQ would have the effect of destroying the subject-matter of the trust, and was therefore an unlawful breach of trust by the trustee.⁴²

The problem for a local authority is that where it purports to grant a tenancy to a minor this will take effect as a trust of land, with the authority as trustee. The authority will then be required to protect the trust property but is simultaneously under a duty to act in the public interest by protecting its assets, which may necessitate possession proceedings. As noted in *Re Alexander-David*, the solution to this dilemma is for the local authority to divest itself of its duty to act as a trustee by way of application to the county court for the appointment of a new trustee.⁴³ This response to the problem created by s1(6) LPA 1925 is primarily one aimed at resolving the difficulty faced by the local authority without undermining the general purpose of Part VII of the HA 1996, which requires an LHA to provide housing (temporary or otherwise) for homeless persons aged 16 or 17 years. What is overlooked – understandably having regard to the issue before the court – is the significance of the decision for children's citizenship, and the rights which contribute to full citizenship.

The introduction of a trust by operation of law undermines the child's interest in the property by establishing their status as beneficiary rather than holder of a legal estate. The general scheme of trustee responsibilities under the TOLATA 1996 has the effect of disempowering IMC as tenants. By virtue of s6 of the TOLATA 1996 the trustee is given the powers of absolute owner in respect of the trust property. Whilst this power must be exercised having 'regard to the rights of beneficiaries'⁴⁴ statute prevents a child from giving their consent to the exercise of certain functions by the trustee;⁴⁵ a trustee is unable, even if so inclined, to

⁴² In the course of the judgment that Court of Appeal confirmed that the LPA 1925 Act did not prohibit the grant of a legal estate to a minor, instead it merely prevented a minor from holding such an estate and accepted that a local authority might grant tenancies to minors that were effective in equity. Referring to the case of *Kingston upon Thames RLBC v Prince* (1999) 31 HLR 794, in which it was held that a minor could succeed to a secure tenancy (applicable in the case of LHA) under the Housing Act 1985. Hale J. said: "A minor can hold an equitable tenancy of any property, including a council house." See also: Law Commission's Report on Minors' Contracts (Law Commission No.124, 1984).

⁴³ Under the Trustee Act 1925. A local authority cannot seek to absolve itself of its duty to act in the public interest.

⁴⁴ S6(5), TOLATA 1996.

⁴⁵ S.10, TOLATA 1996, i.e. those powers required by the trust disposition to be exercised with consent

delegate their powers to be exercised by a child;⁴⁶ and, in any event, a trustee is not required to consult with a minor beneficiary in the exercise of their functions as a trustee.⁴⁷ In this respect, the use of a trust as a solution to the dilemma caused by property law undermines the notion of children's citizenship by: failing to provide IMC with the rights of absolute owner which would normally accompany tenure; limiting IMC participation in decision making to less than that anticipated for adult beneficiaries.

In the course of judgment in *Re Alexander-David* the Court of Appeal considered an alternative to a trust which would be to treat the occupation agreement as a licence rather than a tenancy. As a licence is not a legal estate s1(6) LPA 1925 would not apply, and there would be no need for a trustee to be appointed. However, as the Court of Appeal recognised, merely describing an agreement as a licence will not suffice to avoid grant of a tenancy.⁴⁸ On this issue Sullivan LJ put forward a solution which, in effect, is to deprive a child occupier of their entitlement to a tenancy. Noting that in the case of 16 or 17 years olds housed independently of their family there would often be a need for the local authority to arrange support for the child to maintain themselves, the Judge commented: 'it should not be too difficult' for the local authority in discharging both its local housing authority and social services functions to co-ordinate matters so that, whether by making provision in the agreement for some attention or services, or by permitting inspection of the premises by those charged with the child's welfare, and thereby preventing the grant of exclusive possession, any agreement with a 16–17 year old is not merely expressed to be, but is the grant of a licence rather than a tenancy.'⁴⁹

In *Villiers v Antoniadou*⁵⁰ the Court emphasized the importance of distinguishing genuine arrangements for a licence from 'sham' agreements designed to relieve occupiers of their rights. Whilst it may be the case that some IMC provided with social housing may require additional support to maintain their tenancy, the introduction of support to undermine the right – which would otherwise be established – to a tenancy cannot be justified. Recognition that an IMC may require support services incorporates notions of evolving capacity, as it pays attention to the child's developmental progress toward independence. The purpose of support is to facilitate a child's development in a direction which will enable them to live independently, not to undermine the acquisition of rights which contribute toward independence. In the case of a tenancy, this will include security of tenure. Whilst it is correct that s79(3) of the Housing Act 1985 ensures that the same security of tenure is given to a

⁴⁶ S9(1), TOLATA 1996.

⁴⁷ S11(1), TOLATA 1996.

⁴⁸ See: *Street v Mountford* [1985] AC 809.

⁴⁹ Above n.41, para.37.

⁵⁰ [1988] 3 WLR 139.

licensee of social housing as to a tenant,⁵¹ there will only be security where the occupier has exclusive possession of the premises.⁵² In any event, other interests would be affected if a licence were the form of occupation agreement used. A licensee cannot take advantage of certain rights afforded to local authority tenant such as availing themselves of the right-to-buy.⁵³ In this situation an IMC would be denied the opportunity to enhance their status in society - as home-ownership is valued in the UK above all other forms of tenure. It is perhaps unlikely (but not inconceivable) that an IMC would have the means to be able to exercise the right-to-buy, however, what may be more pernicious is that during the period of occupation as a licensee the child will not acquire any accumulated discount as an occupier for the purposes of any future exercise of the right.⁵⁴ Finally, statutory provisions which require a social landlord to consult on housing management issues apply only to tenants, therefore denying licensees the opportunity to participate in decision-making processes which have consequences for their occupation.⁵⁵

As has been noted the HA 1996 is not the only route into housing for 16 or 17 year old IMC. S20 of the CA 1989 has a wide compass and it is likely that many homeless 16 or 17 year olds will be more appropriately dealt with under this section.⁵⁶ Examples of cases where an application under the HA 1996 may be more appropriate include: where a child becomes homeless after a period of independent living and is not deemed a child in need; or, where a child does not want to be accommodated under s20.⁵⁷ If an SSD is under a duty to accommodate an IMC it may discharge its responsibilities by making use of a range of housing options. These include: placing the child with its family, relative or any other person; or in a children's home; or, by making other arrangements.⁵⁸ The SSD is required to ascertain the child's wishes and feelings, and give due consideration to those wishes and feelings, before accommodation is provided.⁵⁹ The duty to accommodate under s20 of the CA 1989 triggers a much broader range of duties to 'looked after children',⁶⁰ and 'eligible' and 'relevant' children.⁶¹ These include providing assistance to enable 'young people who have

⁵¹ Noted in, above n.42.

⁵² *Westminster City Council v Clarke* [1992] 2 AC 288. In any event the decision overlooks the position of occupiers who are not housed by the local authority where s79(3) will not apply.

⁵³ Part V, HA 1985, the right to buy is only available to secure tenants.

⁵⁴ Schedule 4, HA 1985.

⁵⁵ S105 and s106, HA 1985.

⁵⁶ This is irrespective of whether or not a duty may be found under Part VII of the HA 1996: *R (on the application of G)(FC) v London Borough of Southwark* [2009] UKHL 26.

⁵⁷ Both examples given by the House of Lords in *R (on the application of G)*, *ibid*.

⁵⁸ S23, CA 1989.

⁵⁹ S20(6), CA 1989.

⁶⁰ S22, CA 1989.

⁶¹ S1 and s2, Children (Leaving Care) Act 2000.

been looked after by a local authority move [to live] independently in as stable a fashion as possible.’⁶²

The Court of Appeal has pointed out that the child's wishes and feelings are not determinative and has stated that ‘children are often not good judges of what is in their best interests’ and that ‘there may be cases where the child's wishes are decisive...[but] a local authority should reach the conclusion that the child's wishes are decisive only as part of its overall judgment including an assessment of the child's welfare needs and the type and location of accommodation that will meet those needs.’⁶³ The extent to which a social services authority is able to give effect to a child's expressed wishes concerning the type of accommodation provided, where these point at independent housing, will depend on the resources available. Whilst data shows that the majority of children aged 16 or 17 years who are looked after by a local authority are not provided with independent accommodation, but are instead housed in other accommodation (for example, with foster parents or in secure accommodation),⁶⁴ this may reflect a genuine assessment of need rather than any lack of independent housing options. However, where it is determined that the most suitable accommodation is independent housing SSDs may call upon LHAs to assist with the discharge of their functions, including asking an LHA to make accommodation available.⁶⁵ In cases where an LHA provides accommodation the dilemma caused by s1(6) LPA 1925 is again relevant. The legal issue may be resolved by the SSD acting as trustee, however, the same problems for children's citizenship arise and may be exacerbated as the provision of support services to an IMC by an SSD may itself provide justification for the use of a licence by the LHA.

Conclusion

Sullivan LJ, in *Re Alexander-David*, recognized the tension between housing law and property law,⁶⁶ noting that the interrelationship between the two statutory schemes raises a dilemma.⁶⁷ The tension is brought about as a consequence of the contrast between the

⁶² Explanatory Notes to Children (Leaving Care) Act 2000, para.5.

⁶³ *The Queen (on the application of Liverpool City Council) v The London Borough of Hillingdon* [2009] EWCA Civ. 43, para.32, the first part of which was cited with approval in *R(on the application of G)*, above n.56, para.7.

⁶⁴ ONS, Children looked after in England (including adoption and care leavers) year ending 31 March 2008, Table F1.

⁶⁵ S27 CA 1989. Joint working is encouraged between the various departments involved with children and meeting the needs of children who are in need of housing. For 16 and 17 year olds the objectives of joint working include providing accommodation and 'support pathways towards independent living', and the provision of a personalised support package: DCSF (2008) Joint working between Housing and Children's services preventing homelessness and tackling its effects on children and young people, p.29.

⁶⁶ Above n.41, para.12.

⁶⁷ *Ibid*, para 13.

welfarist approach which underpins housing law and which seeks to give effect to social rights, and the principles of property law. In seeking to resolve the problem the Court of Appeal in *Re Alexander-David* adopted a solution to accommodate property law principles without compromising children's rights to occupation. Understandably, having regard to the narrow issue raised for adjudication, the Court's reasoning took place without reference to rights issues having a bearing on children's citizenship.⁶⁸ Notwithstanding, citizenship rights are weakened by the accommodation arrived at in the case, either as a consequence of the operation of a trust, or through the use of licence agreements. This will have an effect on IMC who apply for housing under Part VII of the HA 1996, or under s20 of the CA 1989. The scheme of the CA 1989 suggests a pathway into housing which recognizes the child's capacity to contribute to decisions on the suitability of housing, and which pays attention to the transitional nature of childhood by introducing support.⁶⁹ Viewed from the perspective of children's citizenship the approach under the CA 1989 is participative and recognizant of a child's evolving capacity. However, this is undermined as housing takes place against the backdrop of the dilemma introduced by property law, solution which has been worked out in this context undermines children's citizenship.

⁶⁸ This should not be seen as suggesting that the court would otherwise have embarked on a discussion of children's citizenship.

⁶⁹ A point noted by Baroness Hale in *R (on the application of G)*, above n.56, para.32.

The housing policies of remodeling and urban renewal in Spain and the United States: a comparative approach

Gómez Jiménez, María Luisa¹

Visiting Researcher IGLP Harvard Law School,
Asoc. Professor Malaga Social Studies School
University of Málaga, Malaga,
Spain

Email: fpmlgomez@gmail.com

Abstract:

One of the key issues in the Spanish Housing Policy now is the study of Housing remodelling and urban renewal processes, as the latest national Housing Project has recently emphasized. The significance of these two aspects is rooted in their response to a crisis context as it is the one taking place in the building industry these days. This paper approaches the comparative analysis of similarities and disparities between the United States and Spain legal frameworks regarding remodelling and renovation in Housing Policy with stress in its connection with the land use law policy.

Keywords:

Housing Policy, Remodelling, Urban Renewal, Housing Projects, Urban Planning

1 The case for a right to Housing in Spain and in the United States and its connection with social needs and social housing

The Universal Declaration on Human Rights recognizes in section 25:

“Everyone has the right to a standard of living adequate for the health and well-being of himself and of his family, including food, clothing, housing and medical care and necessary social services, and the right to security in the event of unemployment, sickness, disability, widowhood, old age or other lack of livelihood in circumstances beyond his control”¹

Most all countries have made a declaration in similar terms considering housing a right. Nevertheless, the kind of recognition made and the effects of it in a practical

¹ Universal Declaration of Human Rights U. N. Doc. A/ 810, at 25 (1948), available at <http://www.un.org/overview/rights.html>.

sense are still very far from the spirit of the original Legislators². In December 2000 in Nice, EU leaders, the European Commission, and the European Parliament proclaimed the Charter of Fundamental Rights of the European Union³. The charter draws together for the first time all the personal, civil, political, economic and social rights into a single text. The way the charter was drafted was itself an achievement by involving all the EU institutions, national Parliaments as well as and a broad spectrum of civil society, although the text does not dedicate one specific section to housing refers to the freedom of movements and residence, understanding that this right implies housing.⁴ Nonetheless, there is a specific right to good government: public authorities must work to address citizens' needs.

The right to housing, which exists within the text of the Spanish Constitution, is particularly important because of its connections with the rest of the rights regarding urban policy and land use. Although this right is recognized, it does not all fall within the specific section of "fundamental rights" in the Spanish Constitution. Instead, the right to housing is considered to be included in the main principle of the social and economic system⁵. Given that housing is not a fundamental right, but a social and economic one, public authorities must act to improve the affordability of and access to housing as a consumer good.

Nevertheless, the Spanish Constitution does not define the mechanisms for reaching this goal. Instead, it simply recognizes the role of the government and grants it the powers to achieve this goal⁶. As pointed out, it is important to bear in mind that housing is not a "fundamental" right. That is, no one has the right to go to the public authorities to apply for a house to live in based solely on the Constitution (García de Enterría, 2006). Another important concept that can be drawn from the constitutional text is that government must address the need for suitable housing for all Spaniards⁷. Although the Constitution does not specify the extent of intervention by the government, the text implies active policy or public intervention in the housing sector.

In Spain, the right to housing is inextricably intertwined with the social and economic state of the country and the need for decent shelter. References to urbanism and housing policy in Article 47 of the Spanish Constitution are connected in the text. This implies that urbanism is fundamental to achieving decent and suitable housing. By contrast, in the United States, housing has never been recognized as a constitutional right. However, this does not mean that housing policy is unimportant, as Professor Ted Koebel pointed out:

² The Spanish Constitution declared the right to Housing in Article 47:

All Spaniards have the right to enjoy decent and adequate housing. The public authorities shall promote the necessary conditions and establish appropriate standards in order to make this right effective, regulating land use in accordance with the general interest in order to prevent speculation. The community shall have a share in the benefits accruing from the town-planning policies of public bodies.
Constitución [C.E.] [Constitution] art.47 (Spain),

³ Charter of Fundamental Rights of the European Union 2 (2000), *available at* http://www.europarl.europa.eu/charter/pdf/text_es.pdf

⁴ Charter of Fundamental Rights of the European Union. 2, (2000), *available at* http://www.europarl.europa.eu/charter/pdf/text_es.pdf

⁵ SPANISH CONSTITUTION, section 53.

⁶ SPANISH CONSTITUTION, section 53

⁷ SPANISH CONSTITUTION, section 47.

“Housing policy initiatives in the United States have largely sought to augment private housing provision. During the Great Depression, when many people lost their homes due to the economic collapse and the short-term financing of homes, federal housing policy in the National Recovery Act of 1933 initially focused on direct federal production of housing through the Public Works Administration. This effort was rebuffed by the courts, which ruled that production of housing for civilians (other than emergency shelter) was not a power delegated by the states to the federal government (*United States v. Certain Lands in Louisville*). Federal policy then shifted to the federal financing, private production, and local government ownership and operation of “public housing” (Gómez Jiménez and Koebel, 2007).

a) A right to Housing and a right to the city: perspectives on social housing.

Spanish Housing policy begins in the Constitution in section 47 and points out that:

“All Spaniards have the right to enjoy decent and adequate housing. The public authorities shall promote the necessary conditions and establish appropriate standards in order to make this right effective, regulating land use in accordance with the general interest in order to prevent speculation. The community shall have a share in the benefits accruing from the town-planning policies of public bodies”.⁸

In 1949, Congress declared that the United States should realize:

“As soon as feasible...the goal of a decent home and a suitable living environment for every American family, this contribution to the development and redevelopment of communities and to the advancement of the growth, wealth, and security of the nation”⁹.

In the United States, the legislative branch determines the level of intervention by public authorities when it comes to housing. Housing policy is made by the United States Department on Housing and Urban Development (HUD), the states, the local housing authorities, and the Federal Tax policy (Swartz, 2006)¹⁰.

In Europe, public intervention in housing is called “social housing” and is broadly targeted to specific groups (Bachin, 1996)¹¹. It is important to note that the European Union does not hold any power in creating housing policy. These powers belong to each country’s national government. This means that there is no European housing policy, except for specific measures regarding building codes, energy efficiency, and environmental protection. As a result, the concept of public housing is defined on a country by country basis.

In the United States, today, although the State government subsidizes some housing, this is not its main priority. Local housing authorities are responsible for managing the stock of existing public housing by resorting to creative measures such as private and public financing and partnership with non profit organizations. Thus, Chester Hartman declares: “the question could be if we can afford the recognition of a

⁸ SPANISH CONSTITUTION (CE) art. 47

⁹ Housing Act of 1949, Pub. L. No. 171, 63 Stat. 413, 413 (1949).

¹⁰ A detailed analysis regarding this description can be read there.

¹¹ That’s the case for the housing for the disabled or frail elderly,

right to housing? (Hartman, 2006). Because, the cost of providing everyone with a decent affordable housing is greatly affected by the ways in which such program would be carried out".¹² Furthermore, the United States lacks a federal or state housing project, as it exists in Spain, although both the federal and state authorities regularly provide funds to local housing authorities.

Nonetheless, as Chester Hartman pointed out, both in Spain and the United States initial funds and support for the right to housing came from religious groups.¹³ Therefore, to deal with the cost of public intervention in housing, we need to come up with a strategy to achieve this public goal in both United States and Spanish legal systems.

In both the United States and Spain, part of this approach relies on defining "public housing." However, each country has different understandings of the level and type of intervention required in this public strategy with a different impact on the way housing needs are solved.

2 The role of public Housing in Spain.

Contemporary housing policy in both America and Spain is rooted in the countries' responses to post-war conditions¹⁴. In Spain, this formative period took place after the Spanish Civil War, while in the United States the formative period occurred after World War II¹⁵. The first public housing activities developed after the Spanish Civil War and were directed towards national reconstruction. At that moment, public authorities took the lead on encouraging housing initiatives, given that the private sector lacked the resources to meet the emergent need for housing (Gómez Jimenez, 2006, pp 39-45).

The measures enacted in the 1976 Social Housing Act encouraged the private sector to build public housing projects.¹⁶ However, the results were inadequate, despite urban reform that was taking place at the time to combat land use speculation.

The most successful attempt to comply with this initiative came in 1978, before the approval of Spanish Constitution, through the consolidation of the VPO, "Viviendas de Protección Oficial," that is, Official Subsidized Housing System¹⁷. The VPO's first act was passed in 1963, and its provisional regulatory by-laws were developed in 1969 and applied years later without changes.¹⁸

In 1976 and 1978, there was an attempt to integrate, in a unique text, the subsidized housing programs with all the public housing initiatives that had been issued up to that point (Gómez Jiménez, 2007). Based on this, public housing can be

¹² *Id* at same page.

¹³ In the United States, the same religious groups that provided leadership in the Abolitionist and Civil Rights movements. *Id.* at 178-79.

¹⁴ *Id* at 2

¹⁵ Even in the (20th) Century Spain's first attempt to provide housing came from the 1922, February 1st, (1921, Act December 10th) Act, so-called "Ley de Casas Baratas", enacting a legislation that had enforced the promotion of social or subsidized housing. An interesting link to this text can be seen at: http://www.ciudadjardinmalaga.es/ley_de_casas_baratas_css.htm

¹⁶ *Id.* at 332

¹⁷ The VPO housing system was enacted by the Royal Decree 31/78, October 31st

¹⁸ *Id.* at 347

understood as any kind of housing with public protection. That is, social housing is considered to be a kind of public housing, and subsidized housing is another type of public housing.¹⁹ The confusion that exists in the terminological aspect is so high that the housing policy of the Autonomous Communities demands its own definition of public housing that has the legal form of subsidized housing and that is called: “housing with some kind of public protection”.²⁰ The exact definition of this public protection means depends on the different set of laws of each Autonomous Communities. It is important to note that these types of public housing are targeted to segments of the population with certain levels of income and particularly to those with low-incomes.

Thus, all these expressions (“public housing” instead of “social housing” or “subsidized housing”) function synonymously, whether housing is being subsidized by the National Administration or by the Autonomous Communities. The significance is that it was built with public aid.

As for subsidized housing, the mechanism to achieve it in both countries is almost the same. Public funds are used to reduce costs and increase the affordability of housing. In addition, subsidized housing in the Spanish context could refer to the promotion of public housing by the government (Administración Pública).²¹ Therefore, an incentive, enabling, or “carrot and sticks” policy²² is generally used to respond to the demand of adequate housing in Spain.

The terminology that the legislative bodies use is confusing. For example, in the United States, public housing describes units sponsored by the federal government and managed by local authorities. In Spain however, any public authority or body can sponsor and manage public housing. The *Ministerio de la Vivienda* (National Public Administration on Housing) and the Autonomous Communities are free to implement their own public housing policies while the local housing authorities (under their local governments) develop a specific housing policy to subsidize units. This overlap occurs because each public body has the power to subsidize units as part of an overall enabling strategy. The power to subsidize is not exclusive but shared.

¹⁹ It is important to distinguish the concepts because there are different names for the same phenomenon. As pointed out previously, Social Housing is the European name for Public Housing or Subsidized Housing. The reason is that with the expression “Social Housing” we emphasize the importance of group populations. By contrast, when we use “Public Housing,” we refer to some kind of social housing. This type of housing is addressed towards low income population and is characterized by the role of the Government in providing funds and the protective system enacted to define the role this type of housing plays in achieving the goal of affordable housing.

²⁰ It means that it holds any degree of financial support to subsidize the unit with the target of increase the affordability of it. The term was not established like that in any law but it’s easily understood when you examine the different content of the acts in the Public Housing Law System.

²¹ By “Administración Pública”, we mean: Public Bodies (at national, regional and local level) able to develop the policy set by the Government as the Spanish Constitution has declared in article 97: “The Government shall conduct domestic and foreign policy, civil and military *administration* and the defense of the State. It exercises executive authority and the power of statutory regulations in accordance with the Constitution and the laws.

²² The carrots and sticks policy in the Spanish legislation comes with the territory if we examine the text of the different housing acts. Thus a carrot is for instance the economic support to promote homeownership through tax deduction. At the same time, the establishment of rigid controls to fight against speculation with subsidized housing works like a stick.

The difference between public housing in Spain and the United States exists in this context when we consider the different distribution of powers among public bodies. Although housing authorities or agencies are not new in Spain, the idea that they exclusively manage the stock of public housing under a federal program run is certainly new and not the reality. In fact, public housing policy in Spain today is defined by the Autonomous Community. Thus, the legislative branch at the regional level regulates the public units and the stock of public housing that must be built under the regional housing project²³.

Hence, the federal government provides subsidies for low-income households in three basic ways: supporting the construction and operation costs of specific housing developments; helping renters pay for privately owned housing; and providing states and localities with funds to develop their housing programs.

As described by Schwartz, the first form of assistance, known as supply-side or project-based subsidies, includes public housing and other programs such as “section 8 new Construction,” in which the federal government helps subsidize the construction and sometimes the operation costs of privately owned low-income housing.

Hence, the laws specifically addressed housing in cities and, as a result, these buildings evinced a mixed urbanism²⁴.

In the Spanish Public Government, public housing implies not just intervention in housing policy but also a budget allocation of public units that can be managed by the regional or local governments and function as subsidized units. The Spanish Minister of Housing does not build housing but gives funds to other public authorities and to entrepreneurs from the private sector to help in the building process. Subsidizing housing then exists through a system of incentives given to both the developers and the owners.

Social Housing in Spain has various connotations. First, it can refer to housing developed under the 1976 Social Housing Act that preceded the Spanish Constitution. Second, it can refer to a non-legal and broader concept of housing targeted to groups with specific social needs. Given that public housing is targeted to this goal, we could say that in Spain public housing and social housing are almost the same idea. Social housing is managed by the private or the public sector in Spain and has several legal restrictions²⁵ in order to preserve its public function: providing housing for low-income individuals, including senior citizens.

²³ It is difficult to summarize here the number of regional housing projects enacted in each of the Autonomous Communities, just to give an example, in Andalusia the last Regional Housing Project 2008-2012 was enacted by the Decree 395/2008, June 24th, in Madrid the last Regional Housing Project Decree 12/2005, 27th January.

²⁴ By mixed urbanism we understand the scenario developed by different legislation when issuing the same topic but with different results. So that we can recognize, just by looking at a geographical area, the effects of a specific Spanish urban planning act.

²⁵ Restrictions set in public housing “regime” is referred to the eligibility to access to the units, to the limitations of selling or renting the unit, and the duration of time which these units are supposed to be affected by the public protection. Besides, compulsory notice to the public authorities must be issued when any change regarding the unit's status arises.

3. Connecting Housing policy and Land Planning

Another important element must be taken into account: the inter relationship between housing and urban planning. The differences here are very clear between the European and the United States legal system. To understand the differences, the role of cities and localities in the juridical context need to be considered (Frug, 1980).²⁶

In this juridical context, it is very important to determine the level of authority that cities and towns hold in Spanish and in the United States systems. Particularly, *there is a lack of coordination in both countries between urban development decisions, environmental policies, and housing policies*. Currently, agencies are working to take into account the urban design the city has, but there is no effective and formal interrelationship between the authorities or agencies responsible for developing a land policy. This implies the parties seeking solutions in an informal manner.²⁷ The inexistence of a “local planning strategy”²⁸, in the sense that it is known in Spain, helps to understand the process in the United States, although, the mere existence of planning tools in Spain has not helped to improve the coordination between the land policy and the housing policy.²⁹

Besides, the development of city planning can be constrained due to the cost of bureaucracy. The positive aspect of the existence of urban planning in Spain is that it helps to understand the future design of a city and thus give more confidence to future investments. When this planning tool doesn't exist, in contrast, the lack of an instrument of land planning gives a city tremendous flexibility. Projects³⁰ help to define a concrete idea of what is needed for an area while at the same time contributing to an increasing level of participation by the population (which is very important). Thus, one can get a more defined image of a city through the knowledge of plans approved for that city.

In the United States the definition of land policy is done through each state, at a regional level, and the cities follow these rules, managing their own resources and approving their own zoning. Meanwhile, the housing agencies at the local level are the ones in charge of managing the stock of public housing. Federally, the U.S. Department on Housing and Urban Development (HUD) at federal level has the power to fund its specific programs throughout the United States.

²⁶ Gerald Frug, *The City as a Legal Concept*, 93 HARV. L. REV. 1059, 1062 (1980).

²⁷ This was called in Spain “*Urbanismo Concertado*” but its particular implication with this issue (housing for seniors) remains unsolved

²⁸ Here the word plan means the land use law tool used by the Spanish authorities to endorse the development of the cities. There are some similarities with zoning but planning implies more constraints in order to define the development and allowances of each zone in the territory. There are basic differences between the “plan” and “project” according to the Spanish legal framework. While plan refers to land use law and zoning definitions, projects refers to housing designs as an economic tool used to measure the number of units to be built in a period of time. The existence of housing projects contrasts with the inexistence of national land use plans, for historical reasons which explanation goes beyond these lines.

²⁹ In Spain we have had several urban plans, at a Regional and Local level targeted to solve urban affairs issues. These land planning tools didn't succeeded in a coordinated response in terms of housing and land use (MARÍA LUISA GOMEZ JIMÉNEZ, ADMINISTRATIVE INTERVENTION IN HOUSING POLICY IN SPAIN (1938-2005) [777] (MONTECORVO, 2006).

³⁰ Here again as in Housing Policy the word is Project and not Plan. So whenever there is no a real urban planning the urban projects define the structure and the uses of the land.

This dispersal of powers makes the integration of land planning in housing policy more complicated and allows it to remain isolated but with a powerful effect on the development of new facilities for specific groups with significant needs.

In the Spanish case, the Ministry of Housing in Spain, as an independent body of the urban policy administration, was enacted in 2004³¹. It was created as a result of the increasing importance of housing after a surge in prices³². Nevertheless, the National Government must adhere to the Spanish Constitution. Therefore, the National government can only make policy by controlling the public expenditure on housing. As stated above, the national government and the general public administration hold the exclusive power to coordinate economic activity.

The difference with the United States case here is clear. In the United States, the federal model, although it must be followed throughout the country, allows local housing authorities to issue their own policy without agreement with the state government. As a result, there is a conjunction of policies directed toward the same purpose. By contrast, in Spain, because the national government does not have authority over housing policy, it must create housing policy through general economic planning while taking into account the powers of the Autonomous Communities in this area.

In addition, as mentioned previously, the Autonomous Communities each have their own laws concerning land planning, and urban planning. There is even a complicated system integrated by acts enacted at a regional level to deal with different matters connected to housing and land planning (environmental, landscape, urban development, and social needs- seniors housing, among others). Also, local communities have been able to create policy directed towards specific sector of the population connected to housing and developed and intensive program regarding housing needs.

4. Defining Gentrification³³ in the Spanish legal Scenario

In the legal framework described above, any process which involves the regeneration of an area connected to urban renewal implies a new building process (or a remodeling one). Besides it requires more than one level of government to be informed because of their different layers of jurisdictions.

This means that the displacement of population must be promoted through economic incentives to allow the area recover the value it might have had in the past. The process itself is especially difficult in Spain since it requires a public investment

³¹ Royal Decree 533/2004, April 17th, (can be seen too at www.mviv.es)

³² Royal Decree 1718/2004, July 23th.

³³ According to the Meriam Webster dictionary Gentrification can be define as: "The process of renewal accompanying the influx of middle-class people into deteriorating areas that often displaces early poorer residents . As described in a general sense in the Wikipedia: " Gentrification or **urban gentrification**, is the change in an [urban area](#) associated with the [movement](#) of more affluent individuals into a lower-class area.^[2] The area experiences [demographic](#) shifts, including an increase in the median income, a reduction in household size, and often a decline in the proportion of racial minorities (if such minorities are present).^[3] More households with higher incomes result in increased real estate values with higher associated rent, home prices, and property tax taxes. Industrial land use can decline with redevelopment bringing more commercial and residential use.

connected to urban planning. The most feasible way to do this is by means of the National and regional housing projects.

This implies again a public intervention in housing policy and implies too, a needed integration between housing policy and land planning. Nevertheless, as it was stated in Gomez Jiménez & Koebel, Ted (2007) “it is important to note that even if the main public initiatives in Spain are focused on subsidizing public units in order to address the needs of these groups, the percentage of housing that is public within Spain is not very high”³⁴. This percentage has decreased in the last few years to the point where the Plan began to promote these units in order to increase the affordability of the housing market. This action is understood when the effects of other factors such as the price of the land, urban laws, and the 1998 Housing Plan and the 1998 land use law, that evaluates housing assets are considered. Therefore, it is not surprising to discover that the importance of public intervention notwithstanding, the number of public units remains low.³⁵

In the United States, the number of public housing units built is also low, however, for different reasons. Local housing authorities responsible for managing the stock of housing realize that the budget allocations for housing are decreasing as a result of federal actions. Leaving aside the operating subsidies and the rehabilitation policy that became important starting in the 1980s in the United States, the truth is that the map of public housing in the United States would not be very different from the Spanish one.

In the United States case it was clear as pointed out by the Freeman that the inclusionary zoning would help in the process of gentrified areas, not the same in Spain, when the bureaucratic structure needed in the urban planning process takes its time.

Besides, the urban process linked to the intervention in housing policy has been generally the opposite. In Spain, the “Nimby”³⁶ phenomenon has affected public housing determining the value of land and creating ghettos where public housing areas are defined (Gomez Jimenez, 2006). This is not new either in the United States where the revival of neighbourhoods with a significant number of public housing has been claimed by some authors (Grogan, 2000)³⁷.

Thus, the Spanish National Housing project is supposed to function like the real tool to connect the renewal process in an urban area with a suitable gentrification area. Nevertheless it's important to remember that National Housing projects work only when the national government makes agreements with the Autonomous Communities to realize the plan. Although the national government provides financial resources under its exclusive power to coordinate economic activity, the exclusive power to issue housing policy is held by the Autonomous Communities.

³⁴ For example, the percentage of public housing in Spain represents just 8 % of the total stock.

³⁵ Between April 2005 and March 2006, 820,107 new units have been authorized in Spain. 92% of these were units listed at market rate. Only 8 % of the total amount provided housing with any kind of public protection or subsidy. By comparison, in the United States “[t]here are approximately 1.3 million households living in public housing units, managed by some 3,300 Housing Authorities,” according to HUD.

³⁶ Not in my Back yard.

³⁷ Paul Grogan (2000); shows this in the mixture of public and private partnership in “Comeback cities”, Westview. As for the public Housing story, can be read;

It means that Autonomous Communities are the ones to define not just the housing policy or the urban planning policy but to define the regional housing tools to improve the remodelling process in a determined area, promoting a gentrification process. Nevertheless it's good to know that this phenomenon is not operating as it did in the United States case and its impacts in the Spanish legislation is very low.

5. Conclusions

In this comparative approach several issues have been considered. Difficult as it is to summarize the main ideas which come out of this research I'll just add some of them as to emphasize the most relevant for our study.

1. In Spain, the right to housing is inextricably intertwined with the social and economic state of the country and the need for decent shelter. This is due to the fact that the recognition of a right to housing is not a fundamental right but a principle of the socio-economic mainstream defined at a constitutional level. In the States, housing has never been recognised as a constitutional right but the legislation enacting the right to housing makes a huge contribution to the matter through the Housing Authorities at a Local level, the HUD at a Federal Level and the States in their sphere of jurisdiction. By contrast, although a possible effect in certain areas connected with economics and environmental protection, the EU doesn't hold any jurisdiction directed connected to housing, so the housing policy must be developed on a country by country basis.
2. Stated as it is that the housing policy is an economic policy the Spanish approach at a National level is connected with the idea of Economic Planning in the housing projects. Nevertheless due to the different development of the enabling strategy at regional level the terminology that the legislative bodies use to define Subsidized Housing is confusing. From the well known label of "Viviendas de Protección Oficial" to the new coined term of public protected or subsidized unit a wide range of possibilities coexist. Besides the definition of Public Housing differs clearly from the same idea in the United States. In Europe the term more commonly used is "Social Housing" to refer to this kind of units targeted to some specific needs or groups of people. Thus, there has been a different approach to the same phenomenon, which in the case of Spain emphasizes more the weight of homeownership instead of rental.
3. Therefore, regarding the interrelationship between housing and urban planning. The differences here are very clear between the European and the United States legal system. Here we must differentiate between the idea of planning and the idea of projects. In Spain Housing Policy has been rooted as a national level through a legislative definition and National Housing Projects while urban planning has required the definition of regional or local planning tools. In the States the significance of zoning works more close to the idea of project rather than planning. This implies more flexibility and a

different understanding of the model of city which is behind the process. Nevertheless neither one model nor the other (neither in the United States nor in Spain) has succeeded in prevent the lack of coordination between the urban planning policy and the housing policy as a overall strategy for a variety of reasons

4. In this scenario is not difficult to conclude that any urban renewal process which might imply a gentrification process must be considered not just from an economic perspective but from an urban planning perspective. In this sense in the United States case it was clear that the inclusionary zoning would help in the process of gentrified areas, not the same in Spain, when the bureaucratic structure needed in urban planning enlarge the decision making process and do not help to prevent the effects of the phenomenon
5. Besides, due to the fact that the Regional Governments in Spain are the ones entitled to develop a comprehensive housing and urban planning policy. Urban renewal falls in their jurisdiction with the effects of diversifying the number of responses given upon the same phenomenon depending on the Autonomous Community you are looking at.

Lastly the huge literature on Gentrification and its connection with urban renewal in the States can help to prevent some of the effects of it in the Spanish case if the legal tools available in land planning and housing law are able to adapt to it.

6. Acknowledgements

This paper is part of a book I'm working on now. I would like to acknowledge the supervision and support in the process of researching on this topic by Professor Duncan Kennedy Carter Professor of Jurisprudence at Harvard Law School, without his supervision and patience this simply couldn't have been possible. At Harvard Law School I would like to thank too for the opportunities given by the Institute Global Law and Policy, and the comments and feed-back and insights got at the Joint Center for Housing Studies at the KSG while there. The wise orientation in any step of the research from the colleagues there were simply great. Not to mention the support of the Real Colegio Complutense at Harvard University, the "alma mater" of Spanish research at Harvard.

In Virginia Tech I acknowledge the patience, encouragement, long conversations and always valuable help from Professor Ted Koebel at the Virginia Center for Housing Research.

I just simply couldn't summarize in these lines all the colleagues' and experts' names I've had the pleasure to talk to.

As for the polishing of the English my acknowledge to the Sheila Lopez, now at New York Town Hall, Justinian Dorste at Harvard Law School, and Eric Engels on the WPS at HELA.

And last but not least to FPJ for his encouragement in the continuous process research and living is.

Thanks to all of them error are just mine.

7. References

Bachin, Paul, (1996) Social Housing in Europe.

Charter of Fundamental Rights of the European Union, 2, (2000), available at http://www.europarl.europa.eu/charter/pdf/text_es.pdf

Frigg, Gerald (1980), *The City as a Legal Concept*, 93 HARV. L. REV. 1059, 1062.

García de Enterría, (2006). *La constitución como norma y el tribunal Constitucional*, Civitas, Madrid.

Gómez Jiménez, (2006), *Administrative intervention in Housing Policy in Spain (1938-2005)*, Montecorvo, Madrid.

Gómez Jiménez, (2008) “La intervención administrativa en la vivienda: de la Constitución Española de 1978 a la Ley Estatal de Suelo de 2007”, in *Derecho Urbanístico de Extremadura. Tomo III. Gestión Urbanística*. Aranzadi, Junta de Extremadura. Consejería de Fomento, pp 1101- 1105,

Gómez Jiménez, María Luisa & Koebel, Ted) (2007), “Public Initiatives and Social Needs: an American and Spanish Perspective on Housing”, *International Journal of Interdisciplinary Social Sciences*, Volume 1, pp 1-20, Common Ground Publishing, Melbourne, Australia.

Grogan Paul (2000), *Comeback cities*, Westview.

Hartman Chester (2006) “The Case for a Right to Housing”, In Rachel G. Bratt, Michael E. Stone, Chester Hartman eds, *A Right to Housing*, Foundation for a New Social Agenda 177, 184, Temple Univ. Press

Koebel, T. (2007), *International Land Planning Congress*. Malaga Land Planning Conference Proceedings,

Schwartz (2006), *Housing Policy in the United States*,

Balancing housing need and prevention of homelessness in England

Patricia Ng¹

Email: patricias.w.ng@btinternet.com

Abstract:

This paper explores the tension between Part VII of the *Housing Act 1996* (the local authority's reactive emergency housing duty to assist statutory homeless people) and the *Homelessness Act 2002* (strategic homelessness prevention duty). A question, which this paper addresses is whether local authorities are placing a greater emphasis on preventing homelessness in practice than the legal requirement imposed on authorities to ensure that people in housing need are accommodated. In carrying out its homelessness prevention work, many authorities have adopted mediation schemes. It is questionable whether mediation is an appropriate process to be used in order to prevent immediate homelessness. The information contained in this essay is partly derived from a wider study carried out in relation to access to housing justice for homeless people, in particular homeless applicants. The wider study drew upon data from fifty case studies of homeless people, including those who had tried to access emergency housing assistance from authorities in London, a total of twenty-one interviews with housing law practitioners, community mediators as well as people who had experienced or were still experiencing homelessness, as well as housing and related problems at the time of the interviews. This paper concludes that a two-fold approach is essential in order to successfully prevent homelessness from arising: a balanced approach to meeting housing need by the supply of housing when necessary, and the provision of mediation, which requires adequate funding, and adherence to the principles and professional standards and codes of mediation. Greater guidance from central government in such areas to local authorities is vitally important.

Keywords:

Homelessness mediation, housing need, England.

1 Introduction

The purpose of this paper is to explore the use of mediation in a crisis situation.² This process is used by local authorities in their dealings with potentially homeless people,

¹ Formerly School of Law, SOAS, University of London

² This paper is based to a significant extent on Chapter Eight of my doctoral dissertation entitled "Mediation and Homeless Applicants." The thesis is entitled *Down and Out and Denied in London: Appropriate and Inappropriate Dispute Processes for Homeless Applicants*. The major themes explored in the thesis were as follows: the notion of "access to housing justice," dispute processing, access to justice, civil justice, homelessness within the homelessness legislative framework. A total of fifty case studies were reported and analysed, and interviews were carried out with housing law practitioners, community mediators who took part in a pilot homelessness mediation scheme for a London authority. People, who had experienced housing and housing related problems, including homelessness, were also interviewed. It is hoped that this paper will give a sense of the important and pressing issues relating to homeless applicants and the mediation models adopted by some local authorities in London explored within the chapter.

primarily in order to prevent homelessness. The question that this essay addresses is whether or not the mediatory assistance given by local authorities to family members in facilitating effective communication (or even to begin to communicate again), in order to avoid homelessness, has become more important in practice than the legal requirement imposed on local authorities to ensure that people in housing need are accommodated. The government in the UK has advocated the use of mediation by local authorities in order to prevent homelessness since the implementation of the *Homelessness Act 2002*. This causes tension between the reactive duties arising from Part VII of the *Housing Act 1996* and the preventative duties imposed by the 2002 Act. The situation is complicated by the impact of the rights of certain types of social housing tenants to buy the home that they are occupying – the introduction of such rights has led to a depletion in social housing stock over the years. This loss of housing stock has made it difficult for housing authorities to fulfil their reactive housing duty to accommodate the priority need homeless. The *Housing and Regeneration Act 2008*, among other things, now enables local authorities to invest in new social housing again, which might eventually lead to a welcome increase in the supply of social rented housing.

Local authorities in England have had a duty to assess, and in some cases provide emergency housing assistance to persons who are homeless, in priority need and eligible for housing assistance since 1977.³ An authority is obligated to accept a homeless application and if necessary, provide interim assistance, while it continues to assess an applicant in relation to the other stages of the inquiry process.⁴ In assessing an application, officers must have regard to the Code of Guidance, which contains much more information to assist officers in their assessment duty (CLG 2006a).

³ *Housing (Homeless Persons) Act 1977*. The 1977 Act has been consolidated into the *Housing Act 1985* (Part III), and further consolidated into the *Housing Act 1996* (Part VII) and was amended at the same time. Part VII of the *Housing Act 1996* was amended by the *Homelessness Act 2002*. Hence, the emergency housing duties imposed on local authorities can currently be found in Part VII of the *Housing Act 1996*, as amended by the 2002 Act.

⁴ The interim or temporary housing duty (section 188 *Housing Act 1996*) is triggered when an authority has *reason to believe* (a lower threshold) that a person is homeless, in priority need and eligible for assistance. The full housing duty is owed to an applicant after a full enquiry has been carried out, and the applicant is found to be homeless, in priority need, eligible for assistance, not intentionally homeless, and has a local connection with the authority where the application was made (section 193). Once an authority has carried out full enquiries into the homeless application, the authority has a duty to issue a written decision (section 184). The priority need categories can be found in section 189(1). 'Eligibility' assesses the applicant's immigration status in terms of whether he or she is in a position to be able to access public funds. The concept of 'intentionality' is complex (section 191), and the Communities and Local Government department provided further guidance to local authorities in 2009 in the form of a supplement to the Code of Guidance – the advisory document which local authorities must have regard to when assessing homeless applications. The question in relation to intentional homelessness that authorities need to ask applicants is whether he or she took action or failed to take action, and as a result became homeless. The point is that if the applicant could have taken action prior to homelessness, but failed to do so, and as a result became homeless then potentially the claimant could have caused his or her homelessness. The same reasoning applies in the situation where the applicant took action when he or she should not have done so, and as a result became homeless, then the applicant potentially made him or herself intentionally homeless. See CLG (2009) *Homelessness Code of Guidance for Local Authorities. Supplementary Guidance on Intentional Homelessness*, at <http://www.communities.gov.uk/publications/housing/intentionalhomelessnessguide>, viewed: 29/7/2010. Finally, the applicant would need to have a connection with the authority at which he or she presents as homeless in order for that particular authority to be under a duty to provide the full housing duty (section 199, see also section 184(2)). Applicants should be able to request assistance from any authority in relation to the interim housing duty (section 188).

2 The Meaning of ‘Mediatory Assistance’

The Community and Local Government (CLG), the government department responsible for homelessness since May 2006, does not provide a definition of family mediation although it does provide a ‘dictionary definition’ without attributing the source. “‘Mediation’ involves a process of intervention (by an intermediary agency) between parties in a dispute to produce agreement or reconciliation” (CLG 2006b:para 5.2). The CLG does however acknowledge the Shelter guide on mediation, which defines mediation as “a process for resolving disagreements in which an impartial third party (the mediator) helps people in dispute to find a mutually acceptable resolution” (Shelter 2004). The CLG promotes the view that there are differing opinions in relation to what is involved in ‘family mediation’ as a homelessness prevention technique, and suggests the key elements of mediation as being the following (specifically mentioning the Lemos and Crane [2001] report on mediation services in Scotland, as well as the above mentioned Shelter publication):

- Those in dispute being willing to take part
- Openness and honesty on the part of all parties
- Commitments to working cooperatively with the other party to find a solution
- Clients feeling that they are in a safe atmosphere and that confidentiality will be respected

The CLG concludes that ideally, family mediation should include both the young person and their parent or other host householder. The CLG acknowledged at the same time that in many local authorities, where mediators are willing to work with only one of the estranged parties, such sessions ought strictly to be described as involving counselling (CLG 2006b:para 5.20). Further, it appears that the CLG accepts that in some cases, mediators see their role as extending beyond simply attempting to reconcile the estranged parties (CLG 2006b:para 5.22). However, the government department considers it crucial for mediation officers to be trained to enable them to understand better the homeless application process, just as homelessness officers ought to understand the mediation process as well as the mediator’s role (CLG 2006b:para 5.24).

3 Tension between ‘Reactive’ and ‘Preventative’ Duties

For about twenty five years, local authorities in England have had a reactive duty to assess any homeless or potentially homeless persons, and assist the ‘priority need’ homeless, along with their family, by the provision of emergency accommodation (*Housing Act 1996* Part VII (as amended by the *Homelessness Act 2002*), hereafter the 1996 Act). However the *Homelessness Act 2002* imposed a strategic duty on authorities to prevent homelessness at a local level. Thus, there is tension between reactive duties contained within the 1996 Act and the strategic duties under the 2002 Act.

Homeless people have never had an automatic entitlement to emergency housing assistance because authorities have discretion in the decision-making process

throughout the different stages of the enquiry. Single people would need to demonstrate ‘vulnerability’ in legal terms. Local authorities are expected to accommodate family members along with the applicant who is in ‘priority need.’ Even so, the legislation gives local authorities discretion to decide, within their local area, the particular definition of ‘vulnerability’ under the 1996 Act. As a result, local authorities are failing to assist those who are not strictly provided for by the legislation. In particular, in London, where there is a severe shortage of social housing, many single homeless people with multiple problems, which might include physical health problems, as well as undiagnosed psychiatric problems, fall through the net.

The homelessness legislation allows for a narrow interpretation of the homelessness law by local authorities, assuming they take the law into account at all (see for example, Loveland 1995, Cowan, Halliday and colleagues 2003). This is a problem which many people face when approaching a local authority in London for emergency housing assistance. Social changes should have an impact on official perceptions of what constitutes a vulnerable person. However, especially in the London boroughs, where there are vast housing shortages, local authority duty homelessness officers have a dual role of assessor and ‘gatekeeper’ of the local authority’s resources.

The 2002 Act, which was implemented on 31 July 2002, imposes a new duty on local authorities to be pro-active and strategic in managing the local homelessness situation.⁵ Since 2002, the prevention of homelessness has been a key matter for local authorities to address, and social services departments have a clear duty to take into account the homelessness strategy in their work. Further, social services departments are expected to assist the housing authority in both the homelessness review and strategy process. Local authorities have a duty to construct homelessness strategies on a five-yearly basis in order to deal with homelessness at a local level. Authorities completed their first homelessness strategies on 31 July 2003, and the second reviews and strategies were completed in July 2008. The strategies are informed by a comprehensive review of the current homelessness situation, and include an estimate of the likely future levels of homelessness. At the same time, the local level of available resources must be reviewed.

The strategic approach to dealing with the local homeless population that authorities are expected to take has been underpinned by a change in policy direction of central government – that of preventing homelessness. The Government’s new policy direction was confirmed in documents, especially *More Than a Roof* in 2002. In January 2005, the then Office of the Deputy Prime Minister (ODPM) – the government department responsible for overseeing homelessness – announced its *Five Year Plan for Housing: Sustainable Communities: Homes for All*. In March 2005, the ODPM published *Sustainable Communities: Settled Homes; Changing Lives – a Strategy for Tackling Homelessness*. This explained the government’s strategy for combating homelessness, raising the question of legislative change in order to encourage homelessness prevention. In that document, the government outlined its plans to start consultation on proposals to assess the need for possible changes to the existing homelessness legislation. The consultation would also consider whether legislative changes could be made to encourage greater use of the private rented sector to provide settled housing

⁵ *Homelessness Act 2002*, sections 1–3.

options, should such options be suitable and meet the needs of households in temporary accommodation (ODPM 2005:para 3.15). Additional aims of the government's strategy included the promotion of effective mediation and counselling services with the aim of reconciling families and prevent homelessness (ODPM 2005:13). Already then, we can see that the policy direction of the government has created tension between the local authority reactive duties to accommodate individual homeless households under the 1996 Act, and the more strategic homelessness prevention duty arising from the 2002 Act to assist the local homeless population.

4 Homelessness Prevention

The policy direction for the homelessness prevention duties can be found in various documents, including the *Homelessness Prevention – A Guide to Good Practice* published in 2006, which provides authorities with information on homelessness prevention tools, along with examples of circumstances when authorities would be expected to focus on preventing homelessness. The Good Practice Guide is merely a document giving authorities recommendations of good practice in the delivery of a service, with what the CLG calls “a new ethos” of homelessness prevention for people in housing need, including people who need emergency housing assistance. In its Good Practice Guide, the government openly admits that the housing options approach to preventing homelessness “recognises the limited scale of social housing resources” (CLG 2006b:para 2.10).

In contrast to the Good Practice Guide, authorities need to demonstrate that they have had due regard to the *Code of Guidance* when considering a homeless application. Yet, even the Code has been diluted to incorporate the homelessness prevention ethos. Thus, Chapter Two of the Code focuses on homelessness prevention work. The definition of such work means “providing people with the ways and means to meet their housing, and any housing-related support needs in order to avoid experiencing homelessness” (CLG 2006a:para 2.2). In addition, the Code gives guidance on the three stages at which authorities may intervene in an attempt to prevent homelessness. The three stages mentioned are early identification, pre-crisis intervention, and preventing recurring homelessness (CLG 2006a:para 2.6). Further, authorities are advised that it is open to them to suggest alternative solutions in cases of potential homelessness where these would be appropriate and acceptable to the applicant (CLG 2006a:para 2.3).

The 2006 Code also advises authorities of which situations homelessness prevention tools could be considered to be appropriate:

- Where people have been asked to leave accommodation by friends or family (CLG 2006a:paras 8.9–8.12, 8.11).
- 16 and 17 years old are expected to live in the family home unless it would be unsafe or unsuitable for them to do so (CLG 2006a:paras 12.7, 12.8–9).

- The government suggests that counselling, negotiation, home visits, mediation and practical solutions could help prevent homelessness in relation to three main causes of homelessness (CLG 2006a:Annex 7 paras 4–6, 10).⁶

At the same time, authorities are minded not to avoid their obligations under Part VII of the *Housing Act 1996* (the homelessness duty), which includes the duty to make inquiries under section 184, if the authority has reason to believe that an applicant may be homeless or threatened with homelessness. Authorities are reminded of their interim duty to accommodate, if necessary, until completion of the homelessness enquiries and a decision is issued. Also, that the threshold for interim duty to accommodate is low. The interim duty takes immediate effect, even where the authority considers the applicant may not have local connection with its district, but may have a connection with the district of another local housing authority.⁷ The Code advises authorities that they must not avoid their obligations under Part VII – the homelessness duties (CLG 2006a:para 6.4). It appears that the CLG were concerned about authorities adopting a homelessness prevention approach in inappropriate cases (CLG 2009:Part 3). The reactive safety-net duty should run parallel to the authorities' homelessness prevention work (CLG 2006a:para 6.4).

In reality, homeless households now have to overcome a further barrier in attempting to gain emergency housing assistance under the 1996 Act. This is because the homelessness prevention policy direction is reinforced by “a crucial component of the new ethos of homelessness work,” namely, the housing options approach (CLG 2006b:para 2.7). Such an approach means that applicants who are potentially in priority need for accommodation would be given a ‘housing options’ interview first, when the officer will determine whether homelessness can be prevented (CLG 2006b:para 2.8).

In terms of the homelessness prevention work authorities could consider adopting, along with the promotion of rent deposit schemes, housing advice, and so on, the CLG considered mediation to be an appropriate process to prevent homelessness. However, the CLG did not provide authorities with detailed guidance in terms of the manner in which mediation schemes ought to be established and developed. The then government department responsible for homelessness, the Office of the Deputy Prime Minister (ODPM), did however, grant funding to Relate from 2005 until 2008 – a voluntary sector organisation that provides relationship counselling – in order to develop a homelessness prevention service, and frameworks for the setting up and establishment of mediation schemes to prevent homelessness. However, it is not clear whether many authorities adopted Relate's frameworks in setting up their mediation services, as this information is not readily available.

As a homelessness prevention tool, ‘mediation’ appears to be a popular process adopted by many authorities (CLG 2007). Yet, in some circumstances the process adopted by some local authorities would not be recognised as mediation. A key question that

⁶ The three most common causes of homelessness are: parents, relatives or friends not being able to provide accommodation, relationship breakdown (which includes domestic violence), and the end of an assured shorthold tenancy (CLG 2006a:Annex 6).

⁷ Section 188, *Housing Act 1996*. See also CLG 2006a:paragraphs 7.3, 6.5 and 7.4.

requires investigation, which this paper addresses, is whether mediation, as a process to facilitate communication, for example, could play a meaningful part in assisting homeless people as a genuine homelessness prevention tool. Before such a question can be addressed, one comment needs to be made at this stage. It appears that there is great potential for authorities to contain the number of homeless applications in the name of homelessness prevention. It can be argued that a general lack of detailed guidance from central government in terms of the manner in which mediation schemes ought to be set up, along with the nature of monitoring required by central government on authorities, are further concerns which reinforces the issue in relation to homeless application containment.⁸

The government has been honest enough to admit that the issue of financial constraints has been the reason why it advocates the homelessness prevention approach. However, practitioners who were interviewed in 2005 for this study complained that, in practice, the safety-net of the local authority homelessness duties is being pulled away from even the most vulnerable homeless in society, unless these potentially homeless or homeless people have representation to assist them in making a homeless application.

The housing law practitioners who were interviewed for this study were specifically asked for their views on the 'homelessness prevention' work that authorities carried out, and whether the practitioners believed that the homelessness prevention work was effective in assisting those that needed help. CW1's response was that "it depends [since] there's a whole spectrum of work [that the local authority carries out]." In terms of the rent deposit scheme that authorities set up as part of the homelessness prevention work "in my eyes [the rent deposit scheme] is a gatekeeping scheme. The way it works as gatekeeping is the local authority will say, for example, 'We've got this lovely property [for] this client, [you can move in with a] rent deposit [which we will give you] or of course you can [make a homeless application and] go into that grotty B&B and wait for a council house that you won't like. What would you prefer to do?'"

CW2 did not believe the homelessness prevention work that the local authorities carried out was very effective, declaring that such work was mere " 'window dressing,' delaying the inevitable [for clients who ought to have had a homeless application taken

⁸ The Audit Commission working with central government monitors the performance of authorities. The performance monitoring system in place up until the financial year 2007-08 was the Best Value Performance Indicator. The Audit Commission through the Comprehensive Performance Assessment (CPA) process formally assessed local authority performance. The CPA process examined how well councils delivered their services as well as how well the council was run. The CPA was based on performance information from a number of sources, which included Best Value Performance Indicators set by the government. The last time authorities were measured against the BVPI was the 2007/08 financial year. Up until April 2008, success was measured in terms of reductions in official acceptance levels, for example, Best Value Performance Indicator 213. BV213 was first introduced in the financial year 2005/06, and measured the impact of housing advice in averting actual or imminent homelessness. Good performance would be a low figure. The National Indicator Set or NIS replaced BVPI from 1 April 2008 – a single set of 198 national indicators. From the financial year 2008-09, authorities were measured against, for example, National Indicator 156 in relation to the number of households living in temporary accommodation. Councils already gather PIE statistics – a record of how many homeless applications local authorities accept. Annex 4 (page 15) of the *National Indicators for Local Authorities and Local Authority Partnerships: Handbook Definitions* – published by the DCLG, and revised in May 2008 – gives guidance to authorities on what good performance in relation to N1 156 entails. "Each local authority has submitted projections showing how they plan to reach their own targets, which we monitor against actual performance each quarter. Good performance is typified by a lower figure." <http://www.communities.gov.uk/publications/localgovernment/finalnationalindicators> , viewed: 29/7/2010.

when they first approached the local authority for emergency housing assistance], diverting people from access to social housing.” CW2 echoed the thoughts of CW1,

CW2: I think it's legitimate to offer [rent deposits and rent in advance] as an option, but I think that certain local authorities have been quite cynical in the way that they've used it, and they've actually diverted people away from access to social housing.

Me: So, you think that the key problem with 'homelessness prevention' work is diverting people.

CW2: It's not making it clear that it's one among other options.

However, CW5 commented that the local authorities are “very good in assisting with preventing unlawful eviction.”

Interviews with these practitioners took place from June until July 2005, almost three years after the implementation of the *Homelessness Act 2002*. Moreover, the comments made by the housing law practitioners about the local authorities' homelessness prevention work were an indication of the subsequent guidance the central government would give in this exact area in 2006 (CLG 2006a).

5 The Promise of Mediatory Assistance

The following table presents the range of mediation approaches in London in 2008. The information that was gathered in relation to the types of mediation services that existed at the time has been organised and grouped into 'models.'

Model	Description
1	Landlord and Tenant mediation scheme. (In general, Tenancy Relations Officers – who are also law enforcers in relation to certain aspects of Landlord and Tenant law affecting private rented sector tenants – would negotiate with private landlords on behalf of private rented sector occupiers.)
2	YP expected to attend independent mediation scheme first.
3	Information from 2003 states that mediation is offered where appropriate. The One Stop Shop for YP is operated by the HD, SS and a voluntary sector organisation for the young single homeless people, is linked to Connexions and plays a key role in arranging mediation, accommodation and support for 16–19 year-olds but the service will be extended to cover 20–25 year-olds.
4	Council staff will make a referral to mediation scheme.
5	Independent mediation scheme that reports back to the council, the outcome of mediation will be assessed as part of the HPA.
6	Mediation is offered to homeless 16 and 17 year-olds, where appropriate, to help them remain at home.
7	Housing Advice Team can offer mediation between disagreeing parties to try to address problems.
8	The Housing Options and Advice Teams are assisted by mediation workers, including a specialist worker for YP threatened with eviction from their family home.
9	Variations on following: A. Mediation service required to report to the LA on the outcome of the

	mediation. The LA has an expectation that the HPA will attend mediation because the outcome of the referral might influence the decision of the HPA. B. The housing advisor at the LA might suggest a referral to CALM (community mediation service) or the HPA might approach CALM him or herself. HPAs are encouraged to use CALM as a realistic option to solve their housing needs. The mediation process will form part of the council's HPA assessment. The mediator will feed back to the housing officer at the LA the outcome of the mediation. C. The service is available to anyone living in the LA area, where a household dispute may result in one party becoming homeless – this includes parents, relatives or friends who wish a person to leave. The outcome of the mediation will be fed back to the referring officer.
10	If the HPA is living with his or her parents or friends, the LA will mediate on the HPA's behalf to try to keep him or her at home.

Key

YP	Young People
HD	Housing Department
HPA	Homeless Application
SS	Social Services
LA	Local Authority

The models of homelessness mediation schemes appearing in the table above raises issues that would benefit from discussion, particularly because these matters span across many of the homelessness mediation models adopted by local authorities.⁹ A primary concern is whether mediation is an appropriate process for councils to use as a homelessness mediation tool. The answer really depends upon the manner in which authorities use mediation. There is support in the alternative dispute resolution (ADR) literature, and among mediators that I interviewed, for mediation to be used as an effective communication tool, to facilitate communication or assist in the repair of relationships. After all, the mediation process involves the exchange of information between parties (see for example, Roberts and Palmer 2005, Roberts 2008, Goldberg, Sander and colleagues 2007). Mediation used in such a manner might or might not result in a return home for the person being evicted. If mediation were to be used to repair relationships or to communicate, the timing is crucial as to when mediation becomes available.

Further, it is questionable whether local authority officers give adequate enough consideration as to whether or not mediation is appropriate before making the referral for mediation – an issue that the Advice Services Alliance raised.¹⁰ This was also the reason why Med3 and Med4, the mediators I interviewed, worked hard at getting to know local authority staff that might be making the referrals. They also ensured that these members of staff understood the nature of the mediation scheme. Unfortunately, referrals were not always forthcoming, or inappropriate referrals were made.

⁹ An internet survey of mediation models was carried out in August 2008.

¹⁰ ASA is a national membership body, as well as umbrella group for independent advice services in the UK. ASA's members are national networks of not-for-profit organisations providing advice and assistance on law, access to services and related issues. See http://www.asauk.org.uk/go/MiscPage_52.html, viewed: 29/7/2010.

As already mentioned, the timing of access to mediation is crucial. Med4 informed me that, “there’s so much that happens within a family in a very short space of time, and you can have some people that can tolerate things for years, and some people that are like, ‘I’m not having it, and I’m out of the door.’” Certainly, in relation to someone’s housing situation, a referral to mediation at the point that person has become homeless is too late. Med4 comments:

For me personally, it didn’t matter at what point they [the young people] presented to the HPU. Often it was too late. They’d already had the mass of the damage done to their relationship with their parents. Or their parents have had the mass of damage done long before they got to the point when they asked [the young person] to leave the house. Or they decided that their parents weren’t listening and they wanted to leave.

In addition, Roberts discusses the disadvantages of someone undergoing mediation at too early a stage, which really is as much of a concern as someone who agrees to mediation when it is too late (2008:179).

In terms of which types of problem mediation should not be used to resolve, the CLG recognises that where the relationship breakdown is due to violence, a referral to mediation ought not to be made.

The issue as to whether an evictee does really attend mediation on a voluntarily basis is a further area of concern. The authors of the CLG Good Practice Evaluation report acknowledged this as a problem, but did not address the difficulty within their report (CLG 2007:81 at para 5.10). Ada, a client, was not given an option whether or not to attend mediation. A local authority officer told her she had to attend mediation because she had been asked to leave her parent’s home. MedLA informed me that should a young person approach her local authority with a request for accommodation, but did not want mediation the young person would be placed in a hostel. However, approximately 50 per cent of the young people would later ask for mediation in order to be able to return home. MedLA told me that usually, after a young person has lived in a hostel for a few weeks, he or she tended to miss home.

It is necessary at this point to mention the problem of some of the mediation models intertwining the mediation process with the statutory homelessness process, thereby incorporating the mediation process as part of the homelessness assessment. A number of local authorities appeared to have intertwined the processes – see Model 9 (A and B). The CLG probably caused confusion in its 2006 Good Practice Guide, by discussing the need for the mediator’s judgement about the scope of reconciliation between the applicant and excluder, when the mediation process is integrated with the homeless application process (CLG 2006b:67–68). However, the case of *Akilah Robinson v Hammersmith and Fulham LBC* [2006] EWCA Civ 1122, explicitly states that a local authority is not entitled to postpone making a decision on a homeless application in order to avoid a duty. Hence, the mediation process ought to proceed independently of the homeless application process, although both processes can run parallel to each other. Moreover, the homelessness officer should not wait for the outcome of the mediation

process before making a decision on the homeless application, nor can an authority expect the young person to attend mediation prior to making a homelessness application, if housing was an urgent need at the point the young person required assistance from the authority.

Where mediators are asked to make judgements about whether a client is homeless or not, these mediators are clearly asked to take on an evaluative role, which, in itself, continues to cause disagreements among ADR commentators (discussed below).¹¹

The referral of a homeless or potentially homeless applicant to an independent mediation service is good practice. However, a related issue that needs to be discussed is that of independent mediation organisations being funded by the local authority, and the expectations placed by the authorities on the mediation organisations. As we saw in the table at the beginning of this section, some local authorities expect the independent mediation schemes to report back to the council the outcome of the mediation in a manner that clearly identifies the client. This issue is linked to another concern, which is the fourth issue, that of the integrity or characteristics of mediation, which includes confidentiality. In Ada's case, a local authority officer provided the mediation. However, another officer divulged confidential information of an extremely sensitive nature about Ada to her father after the mediation. The disclosure of such information to her father caused Ada to fear for her own safety, and probably caused irreparable damage to the father-daughter relationship. Clearly, the independent mediation service reporting to the local authority the individual circumstances of a client is a problem.

There are more serious considerations, however, which Marian Roberts addresses in relation to the key issue of confidentiality of the relationship between the mediator and the parties (2008:10 and Chapter 9, see also Roberts 2007:94, McCrory 1988). Roberts suggests that it is a matter of the discretion of the parties to decide what information they impart to their solicitors or anybody else (2008:187). Yet, it is worth bearing in mind that "the court will be very reluctant to allow confidential exchanges between the parties to be used as evidence in any subsequent proceedings" (2008:187). Hence, following court practice, should a homeless application decision be challenged, which could lead to litigation, local authorities should not expect independent mediation organisations to divulge information of a more personal nature in relation to the mediation process. Alternatively, independent mediation organisations should not allow themselves to fall into that situation, and thereby contravene an ethical and professional code by breaching confidentiality with the client.

Other than the potential unlawful use of the information, should mediation organisations be expected to divulge information of a private nature that ought to remain confidential between the mediator and client? It is worth bearing in mind Roberts' comments about confidentiality, which should be:

¹¹ Shelter 2004:15. See also Riskin 1996, Stulberg 1997, Love 1997, Kovach and Love 1998, Guthrie 2001, Love and Cooley 2005, Roberts 2005, also Riskin 2003.

integral to the relationship between the mediator and the parties... one of the four fundamental and universal characteristics of mediation. It is the cornerstone of the relationship of trust that must exist between the mediator and the parties, and of the free and frank disclosure that is necessary if obstacles to settlement are to be overcome (1997:133).

This raises a fifth issue, in relation to the integrity of mediation and the model where council officers mediate, such as in Model 10. Confidentiality, as discussed above is one of the four core principles of mediation – both ethical and professional. The other three core principles of mediation are the impartiality of the mediator, and the voluntary nature of the process. The mediator does not have power to compel participation in the process, nor impose an agreement on the parties. This brings us to the procedural flexibility available to the mediator (Roberts 2007:94, see also Gulliver 1979:212).¹² Should mediation be carried out by an officer while visiting the home or former home of the homeless applicant, and the officer succeeds in delaying homelessness, the first question that needs to be asked is whether mediation has really been carried out? Not so, in accordance with the core principles of mediation outlined above. The officer might call the process mediation, but the householder meeting with the officer on that one occasion would have felt pressure to agree to the applicant staying longer. In addition, the local authority officer clearly has a vested interest in the outcome of the ‘mediation.’

We now need to examine the problem of the situation where only one party attends mediation. Of concern is that many local authority procedures allow for ‘mediation’ to take place with only one party, and of greater concern is that mediators are willing to work with only one of the estranged parties (CLG 2006b). Where a model of mediation includes a preliminary session with the applicant only – initial information is gathered at this session, prior to the start of mediation – this is understandable. However, time spent with the applicant in the preliminary session to “developing problem-solving strategies which may preclude the need for the young person to leave home,” or “removes the possibility that one of the parties can effectively ‘sabotage’ the process by refusing to participate” is worrying (CLG 2006b:68–69). The first situation is clearly not mediation, however, as the circumstances suggests, ‘problem-solving’ to prevent the young person becoming homeless. The second scenario outlined, either demonstrates the outcome of a situation where mediation is not the most appropriate mechanism for resolving that particular problem, or that more work needs to be done to enable both parties to understand the process of mediation, what mediation could achieve, as well as the voluntary nature of attending mediation. After all, the parties will be agreeing that they vest power in the mediator to assist them to resolve their difficulty, and to arrive at a joint agreement.

¹² See also, Roberts and Palmer 2005:154 for a discussion of the ‘neutrality’ expected of a mediator, and Advice Services Alliance at www.adrnow.org.uk, viewed: 29/7/2010.

6 The Need to Balance Housing Need and Prevention of Homelessness

Thus far, we have seen how since the *Homelessness Act 2002*, homelessness prevention has become a priority for local authorities. It would appear that the prevention work has become more important in practice than the legal requirement imposed on authorities to ensure that people in emergency housing need are accommodated. The level of homeless application acceptances has decreased over the past three financial years.¹³ Furthermore, as discussed above, authorities are monitored on the number of households living in temporary accommodation (in relation to homeless application acceptances). The right to buy scheme, has been a contributory factor to the shortage of social rented housing being available, and would appear to be part of the reason as to why authorities might be containing the number of homeless applications, particularly in London, where there is a severe shortage of social rented housing.¹⁴ Measures have been taken by the recent Labour administration to enable authorities to build more social housing, including social rented housing, under the *Housing and Regeneration Act 2008*. The 2008 Act allows local authorities to opt out of the Housing Revenue Account subsidy arrangements where an agreement is reached with the CLG. It remains to be seen in this economic climate what action authorities would take. There are lessons to be learned from the situation in London where the increasing shortage of housing over the years has led to local authority officers acting as both assessors of homeless applications and financial gatekeepers of local authority resources. However, where a local authority owes a duty to a homeless applicant, an assessment, and if necessary, accommodation, ought to be provided, otherwise an authority would be in breach of its statutory duty.

In the true spirit of homelessness prevention, provided the mediation process is not abused, mediation could have a positive impact, as a homelessness prevention tool, on people's lives whose personal circumstances contain the danger of homelessness. This author suggests that the CLG should give guidance to authorities on ensuring that mediation is maintained as a separate process from the homeless application assessment, and not allow outcomes of mediation to be considered as part of the application process. Both processes need to remain distinct from each other. Indeed, the case of *Robinson* has determined that this should be so. In addition, it would be very helpful for the CLG to give guidance to local authorities to the effect that mediation as a homelessness prevention tool ought to be used to assist with the longer-term housing needs of

¹³ Homeless acceptance figure from 2007–2008 was 63,170; 2008–2009: 53,430 and 2009–2010: 40,020. See Communities and Local Government, <http://www.communities.gov.uk/housing/homelessness/homelessnessstrends/>, viewed: 29/7/2010.

¹⁴ The “right to buy” (RTB) scheme enables local authority (secure) tenants – within the public sector – to acquire the freehold or leasehold of their homes, depending upon whether the landlord owns the freehold or leasehold of the property. A secure tenant can only exercise his or her RTB the property provided he or she has lived in this property from between two and five years. The property is sold subject to the value of the dwelling at the time of the purchase. Tenants are given a discount of the sale value depending upon length of stay in the property, prior to exercising his or her RTB – the longer the stay in the property prior to purchase the greater the discount up to a maximum amount: *Housing Act 1985*. The maximum discounts available to purchase under this scheme – in relation to a number of local authorities (twenty-seven in London) – were not reduced until January 2003. Changes to the RTB scheme came into effect on 22 September 2008 after the *Housing and Regeneration Act 2008* received Royal Assent on 22 July 2008.

homeless people in terms of strengthening their support network.¹⁵ The support network of individuals in society would at the same time lead to the development of a more cohesive society, and ultimately save public money.

The Good Practice Guide refers to the dictionary meaning of mediation, as involving “a process of intervention (by an intermediate agency) between parties in a dispute to produce agreement or reconciliation” (CLG 2006b:64). Additionally, the CLG evaluation report highlights the fact that there is limited literature, which focuses specifically on the evaluation as to the effectiveness of homelessness mediation, which this author agrees (CLG 2007:80). Yet, neither of the two reports considers the rich source of information, which can usefully be referred to within the ADR literature, particularly in relation to mediation, and more specifically, family mediation. If reference had been made to the ADR literature on mediation, then a greater understanding would have been gained in relation to the mediation process.

Information gathered for this study indicates that timing, in terms of when mediation is offered, is important. The data considered appears to indicate that parties might not attend mediation sessions, particularly if the evictor understands the purpose of the mediation is to allow the evictee to return home on a short-term basis as an outcome of the mediation. There might well be a greater uptake of mediation from both the evictor as well as the evictee, if the focus of the mediation is only on facilitating communication, or in repairing the relationship. In addition, a short-term session of counselling or even another form of support, which would enable both parties to understand the benefits of, and prepare the parties for mediation, could encourage voluntary participation. This means that although mediation could still be used as a homelessness prevention tool, the aim of mediation would be to assist the evictee with his or her medium to longer-term housing needs. Once the parties are engaged in the genuinely facilitative mediation process itself, the parties would be able to benefit from the assistance of a mediator to exchange information, to gain an understanding of each other’s views, to shift the focus from the past to the future, and to find a creative solution to the problems experienced by the parties so far.

Time spent with the mediator, jointly as well as separately (in caucuses), is essential. At the joint meetings, the mediator can explain and clarify the reasons for participating in the mediation sessions, as well as assist the parties to gain a clear understanding of the issues. While the separate sessions with the mediator would enable the parties individually to tell the mediator on his or her own of any concerns about attending mediation, to give background information, and to inform the mediator what he or she wants to achieve from the mediation (Roberts 2008:151-158). By telling their story to the mediator – an independent person – each party is able to gain cathartic relief. The

¹⁵ Social networks, in general, tend to be friends and family (see also Lemos and Durkacz 2002 for further information). In this context, it would mean those who are able to provide emotional and practical support, including advice. In a wider social setting, people who could be included in the social network would be colleagues through employment. Or if the person is unemployed and uses a day centre, it would include other people who use day centres – places where homeless people can go during the day. Medical practitioners, such as a GP or dentist might hold sessions at a day centre, as might housing advisors, in order to give advice or to assist in finding an emergency hostel vacancy.

mediator, by summarising the issues, objectives and feelings of each party in the joint meeting, could give each party some new information, and would certainly enable each party to gain a better understanding of the other party's position: "it makes sense... to address the summary to the *same* person whose account it is. In this way, the validity of each person's viewpoint is affirmed, whatever the difference" (Roberts 2008:154).

Secondly, there is much debate within the ADR literature about whether mediators ought to evaluate. Riskin, in his 1996 article produced a grid, which was meant to assist parties to choose a mediator with appropriate qualities that could assist in reaching an agreement by facilitating or evaluating. The grid garnered much criticism, and arguments about why mediators should not evaluate.¹⁶ Lela Love offers a definition of an 'evaluative' mediator, who "gives advice, makes assessments, states opinions – including opinions on the likely court outcome, proposes a fair or workable resolution to an issue or the dispute, or presses the parties to accept a particular resolution" (1997:938). It is clear that where authorities have asked community mediators to make a judgement about whether a client is homeless or not, these mediators have been asked to provide an evaluation. This author considers that in the context of mediation being used as a homelessness prevention tool, facilitative, rather than evaluative mediation would be much more appropriate.

Thirdly, the manner in which mediation organisations are expected to work with local authorities is an issue, in terms of information that mediation organisations are expected to report to local authorities. For example, reporting the outcome of mediation in relation to individuals to the local authority, and reporting to the local authority to inform the council's own decision as to whether the young person was statutorily homeless (CLG 2007:85 at para 5.2.7). This could be an issue in relation to terms agreed within service level agreements. However, the case of *Robinson* should prevent such practices from continuing (CLG 2007:88). Moreover, this author considers that monitoring the performance of mediation organisations in order to assess the effectiveness of service is generally acceptable only when information is anonymised.

Fourthly, provided mediation is carried out properly, and within its integrity and values, mediation has great potential to prevent homelessness, but should not be used to prevent immediate homelessness. Mediation can contribute to prevent homelessness for people who have medium to longer-term housing needs. A greater recognition is needed of the more 'indirect' benefits of mediation, for example, the reduction or elimination of conflict between the evictor and evictee. Lemos and Crane were clear that the purpose of homelessness mediation would be to strengthen social networks, rather than to enable people to return home (2001:22), although the authors did not discuss the potential benefits of mediation as a communication tool. Moreover, mediation needs to be offered well in advance of a homeless application being made, because if mediation fails, the young person would need to make a homeless application in any event. If mediation were to be offered to young people at secondary school, and young people thereby better understand the mediatory process, this could prevent disputes between parents

¹⁶ See for example, Stulberg 1997, Love 1997, Kovach and Love 1998, Guthrie 2001, Love and Cooley 2005, Roberts 2005, See also Riskin, 2003.

and young people from escalating into greater misunderstanding, and a breakdown in communication, leading eventually to homelessness of the young person.

7 Conclusion

The existence of both the 1996 and 2002 Act, and the tension between the local authority reactive and preventative duties, as well as the policy direction on homelessness prevention work imposed by central government, has meant that, in practice, authorities tend to prioritise the prevention work. While there continues to be a severe shortage of social rented housing, and central government monitors authorities against the number of people staying in temporary accommodation, the easier route would appear to be homelessness prevention work. Mediation could potentially offer effective assistance to homeless people, provided authorities also fulfil their reactive safety-net duty when needed. The availability of potential assistance in the form of emergency housing assistance to homeless people is crucial, since mediation, as a homelessness prevention tool, is not designed to provide a ‘quick-fix’ solution to ‘encourage’ homeless people to return to the accommodation they had to leave. There are many lessons that could be learned from the response of London authorities in tackling homelessness, particularly because London has such a severe shortage of social rented housing.

A balanced approach is required in order to effectively prevent homelessness from arising. First, housing need should be met when required, through an increased supply of social rented housing, as well as through the reactive emergency housing assistance. Secondly, mediation should only be offered if it is an appropriate means of intervention to a particular homeless person. In making available mediation to assist homeless people, adequate funding is required, as is an adherence to the principles and professional standards and codes of mediation. Greater guidance from central government in such areas to local authorities is vitally important.

Acknowledgements

The author would like to express her gratitude to the interviewees: the housing law practitioners, community mediators and clients, and especially to Professor Michael Palmer for his supervision, and to Catherine Gardner for her helpful comments on this paper.

References

Communities and Local Government (2006a) *Homelessness Code of Guidance for Local Authorities*, CLG, London

Communities and Local Government (2006b) *Homelessness Prevention: A Guide to Good Practice*, CLG, London

Communities and Local Government (2007) *Evaluating Homelessness Prevention*, CLG, London

- Communities and Local Government (December 2009) *Recording Homelessness Prevention and Relief at E10 of the PIE Quarterly Return*, CLG, London
- Cowan, D, Halliday, S and colleagues (2003) *The Appeal of Internal Review, Law and Administrative Justice and the (Non-)emergence of Disputes*, Hart Publishing, Oxford and Portland, Oregon
- Goldberg, S, Sander, F, Rogers, N and Cole, S (2007) (5th ed) *Dispute Resolution. Negotiation, Mediation, and Other Processes*, Wolters Kluwer Law & Business, Gaithersburg, Maryland
- Gulliver, P (1979) *Disputes and Negotiations: A Cross-Cultural Perspective*, Academic Press, New York and London
- Gutherie, C (2001) "The Lawyer's Philosophical Map and the Disputant's Perceptual Map: Impediments to Facilitative mediation and Lawyering" 6 *Harvard Negotiation Law Review*, pp 145–188
- Kovach, K and Love, L (1998) "Mapping Mediation: The Risks of Riskin's Grid" 3 *Harvard Negotiation Law Review*, pp 71–110
- Lemos, G and Crane, P (2001) *Mediation and Homelessness, A Review of the Literature and the Views of Service Providers in Scotland*, Scottish Homes, Scotland
- Lemos, G and Durkacz, S (2002) *Dreams Deferred. The Families and Friends of Homeless and Vulnerable People*, Lemos and Crane, London
- Loveland, I (1995) *Housing Homeless Persons, Administrative Law and the Administrative Process*, Clarendon Press, Oxford
- Love, L (1997) "The Top Ten Reasons Why Mediators Should Not Evaluate" 24 *Florida State University Law Review*, pp 937–948
- Love, L and Cooley, J (2005) "The Intersection of Evaluation by Mediators and Informed Consent: Warning the Unwary" 21 *Ohio State Journal on Dispute Resolution*, pp 45–71
- McCrory, J (1988) 'Confidentiality in Mediation of Matrimonial Disputes' 51 *Modern Law Review*, pp 442–466
- Office of the Deputy Prime Minister (March 2005) *Sustainable Communities: Settled Homes, Changing Lives. A Strategy for Tackling Homelessness*, ODPM, London, <http://www.communities.gov.uk/documents/housing/pdf/137815.pdf>, viewed: 17/5/2010
- Riskin, L (1996) "Understanding Mediators' Orientations, Strategies, and Techniques: A Grid for the Perplexed" 1 *Harvard Negotiation Law Journal*, pp 7–51

Riskin, L (2003) "Decisionmaking in Mediation: the New Old Grid and the New New Grid System" 73 *Notre Dame Law Review*, pp 1–54

Roberts, K (2005) "Mediating the Evaluative-Facilitative Debate: Why Both Parties Are Wrong and a Proposal for Settlement" 39 *Loyola University Chicago Law Journal*, pp 187–213

Roberts, M (1997) *Mediation in Family Disputes: Principles of Practice*, Arena, Aldershot

Roberts, M (2007) *Developing the Craft of Mediation. Reflections on Theory and Practice*, Jessica Kingsley Publishers, London and Philadelphia

Roberts, M (2008) *Mediation in Family Disputes: Principles of Practice*, Ashgate Publishing Limited, Hampshire

Roberts, S and Palmer, M (2005) *Dispute Processes. ADR and the Primary Forms of Decision-Making*, Cambridge University Press, Cambridge

Shelter (2004) *Mediation for Young People, A Good Practice Guide*, Shelter, London

Stulberg, J (1997) "Facilitative Versus Evaluative Mediator Orientations: Piercing the 'Grid' Lock" 24 *Florida State University Law Review*, pp 985–1005

Black Economic Empowerment in South Africa: an update

Charles Nsibande¹

¹Bogoje Properties, Johannesburg, South Africa

Email: m_nsibande@yahoo.co.uk;

Elizabeth Pienaar²

²School of the Construction Management and Economics,

University of Witwatersrand, Johannesburg,

South Africa

Email: edeling@worldonline.co.za

Abstract:

The Republic of South African Government enacted the broad-based black economic empowerment (BEE) Act No. 53 of 2003, in January 2004, to enforce BEE in all sectors of the South African economy. Several papers, covering different BEE topics, were presented at the COBRA 2009 conference highlighting the importance and relevance of BEE legislation to the construction and property industry. This paper gives clarity on the relevance of BEE to each of the papers presented at the conference and describes how each paper fits with the broader objectives of BEE legislation.

The aim of the paper, as requested by the RICS Research Department, is to bring clarity and make it easier for people who will in future read research papers on BEE to understand its relevance and how each research topic on BEE fits with the objectives of BEE legislation.

The paper brings the broader perspective on BEE objectives and how the legislation came about. Background on BEE legislation and its development is given. An update on the BEE implementation is given. The paper gives reviews of each paper (that had BEE on its content) presented at the COBRA 2009 conference and gives synopsis of the main topic on the paper, highlighting BEE legislation's objectives in the content of the paper.

In conclusion, the prospects of the BEE Act are evaluated based on the contents of the papers presented at the COBRA 2009 conference.

Keywords: Black Economic Empowerment (BEE), COBRA 2009 BEE papers.

1. Introduction

Since the enacting of the black economic empowerment act (BEE Act) of 2003, there has been a lot of activity in all sectors of the South African economy, including the construction industry, in implementing the requirements of the BEE Act. In the same light, it is important to note that the earlier concepts of black economic empowerment (BEE) has evolved since they were first introduced in 1994 to adapt to the changing needs in the implementation process. Nowadays BEE is driven by principles of broad based black economic empowerment (BB-BEE) which emphasises the importance of larger numbers of “black” representation in BEE transactions and in the implementation of the Act. Game theory is relevant in understanding the resistance by some companies and acceptance by others in the implementation process. The purpose of the paper is to review the impact, relevance of game theory and to give an analysis of the papers that were presented at COBRA 2009 conference. The paper gives clarity on key elements and objectives of BEE while drawing possible scenarios on whether the BEE objectives are likely to be achieved or not.

2. Objectives of the BEE Act

According to Woolley (2005:12) the government defines BEE as “an integrated and coherent socio-economic process that directly contributes to the economic transformation of South Africa and brings about significant increases in the number of black people that manage, own and control the country’s economy, as well as significant decreases in the income inequalities”. Woolley emphasis that BEE is strategic and should match the company goals to increase value creation and stresses that the companies that will benefit from BEE are those that have strong BEE strategies as all companies in the sectors will eventually be BEE empowered hence BEE, on its own, will not offer competitive advantage. Balshaw and Goldberg (2005) also note the importance of BEE in the “new” South Africa. Balshaw *et al* states that BEE affects all sectors of the economy and every company irrespective of its size, nature or ownership structure.

2.1 *Reasons for BEE legislation*

The main purpose of the Republic of South African Government passing BEE legislation was to enforce BEE in all sectors of the South African economy in trying to redress the imbalance in the South African economy which is racially biased. The racial imbalance in the control of the economy is a result of the discrimination and segregation laws passed by the pre-1994 governments which disempowered certain races of the population. Most of the segregation laws were passed in the 1950’s and 1960’s and include the Group Areas Act and the Bantu Education Act, as two examples referred to

in this paper. The Group Areas Act of 1950 introduced restrictions for “non-whites” to move around the Republic freely and also introduced restrictions on “non-whites” to land ownership in certain areas of the Republic, notably urban areas. This was a great source of economic disempowerment and resulted in adverse poverty amongst the “non-whites”. The Bantu Education Act of 1953 introduced separate curriculum for different races in the Republic which resulted in certain races in the population having inferior education and again disempowered as they could only qualify for unskilled employment. The BEE legislation seeks to address some of these imbalances through an organised and transparent process.

2.2 *Constitutional framework for BEE legislation*

Many people in South Africa have questioned the legitimacy of BEE legislation and have challenged affirmative action in the South African courts. A former constitutional judge, Abie Sacks, in Marcus, Mangcu, Shubane and Hadland (2007) best describes the legitimacy of the BEE legislation. Sacks refers to section 9(1) of the constitution, the equality protection clause, section 9(2), the affirmative action clause and section 9(3), the anti discrimination clause. In his arguments, Sacks emphasizes that each of the three clauses must be viewed together and notes that it is often difficult to best integrate the three provisions of the clause into a constitutionally and socially meaningful whole. Sacks’ arguments are based on the contradictory principles of affirmative action and anti-discrimination. For example, Sacks states that “a holistic view of equality must be in the South African content”, where there is a history of inequality due to discrimination legislation, in order for one to fully understand the provisions of the three constitution clauses.

BEE legislation has evolved since the constitution was enacted and several other pieces of legislation have been enacted to support BEE. These laws work in conjunction or support the objectives of the BEE Act. Balshaw *et al* noted some of these laws and below is a list of important legislation that work with the BEE Act.

Table 2.1: BEE legislation framework

1994	The ANC-led government assumes power and the interim constitution is promulgated The Reconstruction and Development Programme Promulgation of the National Economic, Development and Labour Council Act 35 of 1994
1995	The Labour Relations Act 66 of 1995 The government develops a strategy for the Development and Procurement of small business The Affirmative action policy development forum is established
1996	Final constitution is adopted

	Green paper on employment and development equity released National small business Act 102 of 1996 introduced
1997	Green paper on public sector procurement reform Employment equity bill
1999	Employment Equity Act 55 of 1998 Skills development Act 97 of 1998 and Skills development levies Act of 1999 Competition Act 89 of 1998 National Empowerment Fund Act 105 of 1998
2000	Public Finance Management Act 1 of 1999 Promotion of Administrative Justice Act 3 of 2000 Preferential Procurement Policy Framework Act 5 of 2000
2003	Broad Based Black Empowerment Act 53 of 2003 Promotion of Equality and Prevention of unfair discrimination Act 4 of 2000
2004	Draft codes of good Practice released by the department of Trade and Industry Municipal Finance Management Act 56 of 2003
2005	Department of Trade and Industry releases the 1st phase (code 000, 100 and 200) of the BB-BEE Codes of Good Practice
2007	The Code of Good Practice is gazetted in February 2007

2.3 *Evolution of BEE*

The evolution of BEE transactions has been recorded since the early 1990's when all BEE transactions involved shares' transfer to independent BEE companies through special purpose vehicles which were highly geared structures. These BEE transactions experienced problems in the early 2001's as there were dependent on the shareprice going up to service the interest on the debt. 2001 saw the lowest number of BEE transactions with only about fifty five BEE transactions (Woolley, 2005). The government immediately responded with the introduction of the BEE Act of 2003 to address the shortcomings of the earlier BEE transactions and its implications.

Vuyo Jack in Marcus *et al* (2007) accounts for the evolution of BEE transactions in three waves. The early wave of BEE transactions was typically shareholding or ownership transfer and had little or no other dimensions of empowerment. The 1993 transaction, involving the sale of Metropolitan controlling shareholding to New Africa Investment Limited (Nail), is a typical example of the first wave of BEE transactions. This transaction was done through the establishment of a special purpose

vehicle that was highly geared to facilitate the purchase of the shares from Sanlam. Nail proceeded and bought further shares in other listed companies including MTN, African Merchant Bank, Theta, Sowetan, Radio Jacaranda and Randmark. Nail grew to be the largest black owned publicly traded company with a market capitalization of nearly R6 billion at its peak in 1998. Jack states that the second wave of BEE began in 1997 through the establishment of the BEE Commission (BEECom) under the leadership of Cyril Ramaphosa, a former trade unionist and politician who later became a successful businessman. The BEECom set the background for government policy on BEE with the release of a report in 2001, which introduced the change of focus from only ownership and introduced other elements of business strategy like employment, skills development and preferential procurement. This was the beginning of the codes on good conduct for BEE transactions as there are currently applied in the current BEE scorecards. The third wave came with the introduction of the preferential procurement legislation in 2000. The preferential procurement legislation empowers government to award tenders based on allocation of points for transformation (and BEE) issues for contracts worth less than R500 000. This third wave of BEE seeks to encourage black business to start their own businesses from scratch and grow them through procurements and enterprise development opportunities from already established companies. According to Jack, the third wave has broadened the pool of BEE companies and has increased job opportunities in the country. This type of BEE is often referred to as the second economy and there is government policy to further develop the second economy as a strategic means of reducing unemployment.

3. Synopsis of the papers presented at COBRA 2009 conference

Nsibandé and Pienaar (2009) presented a paper on the impact of BEE legislation on shareprice. The paper looked at the “seven areas of focus where a company can score points and achieve BEE status or BEE accreditation, namely; ownership, employment equity, managerial control, skills development, preferential procurement, enterprise development and socio-economic investment. The main purpose of their research was to determine if the different types of strategies employed in the BEE transactions yielded different results on shareholder value by monitoring the impact each strategy had on the shareprice. The main hypothesis was to either support or disprove the hypothesis that the different strategies or methods employed to achieve BEE status will impact differently on shareprices of companies. From the main hypothesis, several sub-hypotheses were drawn: firstly, the transfer of ownership as a means of achieving BEE status has no effect on share-price, that is, when companies undergo an acquisition or a merger, there should be no impact on the share-price of the new company; secondly, employment equity has a positive impact on share-price; thirdly, managerial control as a means of achieving BEE status has a positive effect on share-price, that is, companies employing managerial control as a means to achieving BEE status should impact positively on their share-price;

fourthly, the announcement of skilled employees taking key positions in companies as a means of achieving BEE status has a positive effect on share-price, that is, companies employing skills development as a method of achieving BEE status are likely to have a positive impact on their share-price; fifthly, that preferential procurement as a means of achieving BEE status can either have a negative or positive impact on share-price depending on the nature of the vendor providing the service to the company, that is, when a vendor is in the same industry as the company, positive effects can be expected on the share-price; sixthly, that companies that choose enterprise development as a means of achieving BEE status will be negatively impacted on their share-price; and lastly, companies that choose socio-economic development as a means of achieving BEE status will be impacted negatively on their share-price". Results of the research indicated that overall BEE transactions did not impact on shareprice.

Khatleli and Root (2009) presented a paper on whether the private sector was obstinate agents or maturing stewards in the implementation of BEE. The paper reviewed the private sector's commitment to BEE policy in South African PPPs. The paper stated that "the South Africa's Black Economic Empowerment policy, due to practical exigencies, depended on the previously privileged white-owned companies for its implementation. This arrangement required the government to adopt a management trusted; a pertinent tenet in a country with a history of an institutional racial economic policy". The paper concluded with a PPP case study that demonstrated that "PPP consortia can default transcendental motives which are not purely utility maximizing, but are mutually beneficial". The paper concluded that "agents in PPPs, as a result of pre- and past-contractual interactions, are able to evolve and adapt a posture more akin to a trusted steward on BEE implementation and provides insights for policy makers and practitioners on the utility of PPPs in delivering socio-economic benefits.

Viruly (2009) presented a paper on the transformation of the South African built environment with an institutional approach. The paper developed "the proposition that the characteristic of the South African built environment can largely be explained by the economic and social institutional framework that continues to characterise the market. The proposition is developed within the context of the institutional economics literature and demonstrates the relevance of the theoretical framework in explaining the present and future trajectory of the South African built environment. While recent public policy and private sector policies and interventions have attempted to alter the players, the South African economy, there is growing evidence, and concern, that a significant institutional "lock in" continues to perpetuate historical market outcomes. There is realisation that government intervention may not have been successful in changing the institutional framework and structure of the South African built environment. This suggested that when strong institutional "lock in" exists, public

sector interventions that attempt to alter market institutional arrangements should target different echelons of the institutional hierarchy”.

Cloete (2009) presented a paper on transformation in property education in South Africa. The paper described transformation of real estate education programmes since the election of a democratic government in 1994. The paper referred to issues related to capacity building: student enrolment, the structure of formal real estate courses offered, skills development, the role of the South African Qualifications Authority and adult educational programmes. Results of the research showed conclusively that there was an increase in participation of the previously under-represented groups and mirror trends in transformation of tertiary education in South Africa”.

Rowland and Hoffman (2009) presented a paper on evaluation of credit risk assessment criteria for commercial property finance applications for BEE borrowers. The paper stated that “significant barriers exist which are preventing the redistribution of wealth to “blacks”. Principal amongst them was the lack of capital in black hands a lack of appropriate skills and experience which combine to prevent blacks from gaining access to commercial banks finance. The research indicated that lack of capital and collateral were key challenges facing BEE companies in raising finance from commercial banks. The paper presented a solution that could work, that BEE companies could form partnerships or joint ventures with white and experienced property developers and investors in order to mitigate the risks associated with lack of capital, skills and collateral to satisfy banks lending criteria”.

4. Prospects of the BEE Act

Based on the papers presented at the COBRA 2009 conference, one could develop possible outcomes and prospects of the BEE Act and hypothesize if the objectives of the Act will be met. Game theory is relevant in understanding and analysing the behaviour of industry players and their reaction to the legislation.

There are four possible outcomes of the BEE legislation. The outcomes could be explained using game theory as the BEE implementation is still in progress and the BEE legislation is evolving from time to time. The first possible outcome is a “win-win” situation whereby government achieves the objectives of the legislation while the construction industry strives. The second possibility is a “win-lose” situation whereby the government achieves the BEE legislation objectives while the construction industry fails as a result of implementing BEE. The third possibility is a “lose-win” situation whereby the government fails to achieve the objectives of BEE legislation while the construction industry continues to strive. The last possibility is a “lose-lose” situation whereby both the government and the construction industry fails as a result of implementing the BEE legislation, that is, the objectives of the

BEE legislation would not be met while the construction industry also fails to perform as a result of implementing BEE.

The papers presented at the COBRA 2009 conference give insight to which scenario is developing in the implementation of BEE legislation. Nsibandé and Pienaar (2009) gave a positive outlook, with BEE implementation having no impact on shareprices of companies that implemented BEE legislation. This is certainly a “win-win” situation whereby companies can comply with BEE legislation, that is, assist the government achieve its objectives while the construction companies are not negatively affected by BEE. This situation indicates that companies that embrace the principles of BEE do not necessarily suffer as a result of implementing BEE. Khatleli and Root (2009) also supported a scenario of “win-win”. The research in the paper indicated that PPPs that were implementing BEE principles in delivering large projects were indeed “adapting a posture more akin to a trusted steward on BEE implementation”.

On the other hand Viruly (2009) gave a less positive outlook which was more of a “lose-win” situation. The paper indicated that government was failing to achieve its objectives through BEE legislation measures. The paper postulates that government is unlikely to achieve BEE legislation objectives since the construction industry has a strong institutional “lock-in” hierarchy. The paper gives no indication of adverse impact on the construction industry as a result of the BEE legislation.

Cloete (2009) is not conclusive as to whether the educational institutions were impacted positively or negatively as a result of achieving the transformation as laid out in the BEE legislation. The paper concluded strongly that transformation had indeed been achieved in construction and property education sector. From the paper the two possible outcomes could either be a “win-win” situation whereby the results of transformation have been achieved while the standards of education have been upheld. A “win-lose” situation could be when the transformation has been achieved at the expense of the standards in the quality of the education.

Rowland and Hoffman (2009) indicates a “lose-lose” situation with financial institution failing to embrace BEE objectives while the BEE construction companies also failing to realise their ambitions of improving performance in the property development industry. The situation, as reported in the paper, is stagnation and is what both government and the construction industry would worry of. This is the worst case scenario that would result in government changing BEE legislation and imposing possible fines and punishment to role players. Government has indicated that there is a possibility of making changes to BEE legislation to make it more effective in transforming the economy, including the construction industry.

Overall, the papers presented indicated a positive outlook with eight “win” cases and four “lose” cases. The result indicates that BEE legislation is likely to achieve its objectives. The “win” cases and “lose” cases were derived from the analysis given above.

5. Hypothesis

From the four papers reviewed and discussed above, a hypothesis has been developed:

H_O: The government will achieve BEE transformation objectives without impacting negatively on the construction and property industry.

H_A: The government will achieve BEE transformation objectives at the expense of negatively affecting the construction and property industry.

6. Data: The impact of BEE on shareprice

From Nsibande and Pienaar (2009), further research was undertaken in the last twelve months to investigate the impact of BEE on shareprice. Fifteen construction and property companies have made BEE transactions since the enacting of the BEE Act. They all made their BEE transactions public through the investors’ news agency of the Johannesburg Securities Exchange (SENS). The fifteen companies are all listed on the JSE Securities exchange and include the following companies and the dates of their BEE transactions:

Construction companies:

Company	Date of BEE Transaction -SENS announcement
Aveng	16 February 2005
Basil Read	15 December 2005
Group 5	10 August 2005
Murray and Roberts	31 August 2005
WBHO	12 September 2006

Property companies:

Company	Date of BEE Transaction - SENS announcement
Acupac	13 February 2006
Ambit	11 June 2007
ApexHi	09 May 2006
Emira	05 June 2006
Growthpoint	30 August 2005
Hyprop	29 March 2006
RedefineMadison	30 August 2007
Redefine Income Fund	11 October 2007
Resilient	08 June 2006
Vukile	31 January 2006

7. Methodology

Brown and Warner (1980, 1985) gave a good background on the use of event studies' methodology in measuring shareprice performance. The event date for the research project was the SENS announcement date when the details of the BEE transaction were divulged. For each company, the shareprice was tracked and analysed around the SENS announcement date, as the event date (day 0) for the research. The performance of the shareprice was monitored 20 days prior to the announcement and 10 days after the announcement, to determine if there was any abnormal movement on the shareprice around the announcement date. The cumulative average residuals (CAR) technique, as employed by Fama, Fischer, Jensen and Roll (1969), was used to determine the exact point of impact on the shareprice.

8. Results and discussion

The results of the BEE transactions, analysed using the methodology discussed above, are tabulated below.

Company	No impact	Impacted
Basil Read	X	
Ambit (Ownership)	X	
Group 5		X
Aveng		X
Vukile	X	
Redefine Income Fund		X
WBHO	X	
Growthpoint (Ownership)	X	
Emira		X
Murray and Roberts	X	
Acupac		X
ApexHi A	X	
ApexHi B	X	
Resilient		X
Ambit (Managerial control)	X	
Growthpoint (Managerial control)	X	
Group 5 (Enterprise development)		X
Hyprop	X	
Madison	X	
Redefine	X	
Total	13	7

Thirteen of the companies' shareprices were not impacted by the BEE transactions while seven were impacted. The results clearly indicate that a majority of the BEE transactions did not impact the shareprices of the companies that implemented BEE post the enacting of the BEE Act.

9. Conclusion

As it has been shown in Nsibandé and Pienaar (2009) updated work, BEE does not have a negative impact on the majority of construction and property companies hence it can be drawn that it will not negatively impact on the industry. However, more analysis of data, looking at different aspects of BEE, as in the other papers reviewed, needs to be undertaken to fully understand the overall impact of BEE on the construction and property industry of South Africa. There will be follow on work from this paper.

References

Balshaw, T and Goldberg, J (2005) *Cracking Broad Based BEE: Codes and Scorecard unpacked*, 1st edition

Human and Rousseau, Cape Town, South Africa.

Brown, S. J. and Warner, J. B. (1980): Measuring security price performance, *Journal of Financial Economics*, Volume 8, Page 205–258.

Brown, S. J. and Warner, J. B. (1985): Using daily stock returns: the case of event studies, *Journal of Financial Economics*, Volume 14, Page 3-31.

Cloete, C.E. (2009) *Transformation in property education in South Africa*, a paper presented at COBRA 2009 conference, University of Cape Town, Cape Town, South Africa.

Department of Trade and Industry, Republic of South Africa (2005): *The codes of Good Practice on Broad Based Black Economic Empowerment – Phase one*.

Government Printer, Pretoria, South Africa

Fama, E.F., Fischer, L., Jensen, M. and Roll, R (1969): The adjustments of stock prices of new information, *International Economic Review*, 10, February, Page 1 -21.

Khatleli, N and Root, D (2009) *Obstinate agents or maturing stewards? The private sector's commitment to BEE policy in South African PPPs*, a paper presented at COBRA 2009 conference, University of Cape Town, Cape Town, South Africa.

Marcus, G., Mangcu, X., Shubane K. and Hadland, A. Ed. (2007) *Visions of Black Economic Empowerment*

Human Sciences Research Council, Johannesburg, South Africa.

Nsibande, C. And Pienaar, E (2009) *The impact of BEE legislation on shareprice*, a paper presented at COBRA 2009 conference, University of Cape Town, Cape Town, South Africa.

Rowland, Q and Hoffman, D (2009) *Evaluation of credit risk assessment criteria for commercial property finance applications by black economic empowerment borrowers*, a paper presented at COBRA 2009 conference, University of Cape Town, Cape Town, South Africa.

Viruly, F (2009) *The transformation of the South African built environment – an institutional approach*, a paper presented at COBRA 2009 conference, University of Cape Town, Cape Town, South Africa.

Woolley, R. (2005) *Everyone's guide to BEE and how to implement it*, 1st Edition
Zebra Press, Cape Town, South Africa.

Who owns the copyright of architectural works and designs?

Phebe Mann

School of Construction Management and Engineering,
University of Reading, Whiteknight, Reading, RG6 6AH,
United Kingdom

Email: p.mann@reading.ac.uk

Abstract:

Copyright protects the rights and interests of authors on their original works of authorship such as literary, dramatic, musical, artistic, and certain other intellectual works including architectural works and designs. It is automatic once a tangible medium of expression in any form of an innovative material, which conforms the Copyright Designs and Patents Act 1988 (CDPA 1988), is created. This includes the building, the architectural plans and drawings. There is no official copyright registry, no requirements on any fees need to be paid and they can be published or unpublished materials. Copyright owners have the rights to control the reproduction, display, publication, and even derivation of the design. However, there are limitations on the rights of the copyright owners concerning copyright infringements. Infringement of copyright is an unauthorised violation of the exclusive rights of the copyright author.

Architects and engineers depend on copyright law to protect their works and design. Copyrights are protected on the arrangements of spaces and elements as well as the overall form of the architectural design. However, it does not cover the design of functional elements and standard features. Although copyright law provides automatic protection to all original architectural plans, the limitation is that copyright only protects the expression of ideas but not the ideas themselves. It can be argued that architectural drawings and design, including models are recognised categories of artistic works which are protected under the copyright law. This research investigates to what extent copyrights protect the rights and interests of the designers on architectural works and design.

Keywords:

Architectural works and design, copyright, infringement, intellectual property.

1 Introduction

The main objectives of this research is to explore the copyright issues on the protection of architectural works and designs in order to understand how the Copyright Law works in the UK, in order to determine the validity and effectiveness of the UK system in respect to protecting the interests and rights of architects and designers. Furthermore, to investigate the Copyright Law in the UK on the protection of architectural works and designs, and consider the possibility of reform proposed in the UK. Finally, to examine

how effective are the UK and US systems of Copyright Law in protecting the rights and interests of architects and designers on architectural works and designs.

In order to provide solutions to the above objectives, this research sets out to answer the following research question:

“How effective is the UK’s Copyright Law in protecting the rights and interests of architects and designers?”

In considering how effective the UK’s Copyright Law in protecting the rights and interests of architects and designers, the following factors were taken into account:

1. appropriate reward for the authors’ creativity on the design which are original even if they are also functional shall be considered.
2. protection of designs from unauthorised use and copying (exclusive rights)
3. clear infringement standard
4. remedies for infringement

2 United Kingdom Copyright Law

Copyright law protects architectural drawings and plans which are considered as “artistic work” and “graphics” under the Copyright Designs and Patents Act 1988. A work of architecture being a building or a model for a building is an “artistic work” under s 4(1) of the Copyright, Design, and Patent Act 1988 as amended. It is generally agreed that to benefit from copyright protection a work of architecture requires an artistic character in the design. Buildings and architectures can attract protection under the copyright law. They are essentially protected indirectly by copyright in the plans and drawings to construct them. In *Pearce v Ove Arup Partnership Ltd (Copying)*¹, Pearce believed that features of the design of Kunsthal Museum in Rotterdam had been copied from his Docklands plans therefore infringed his UK and Dutch copyrights. It was held that Kunsthal was independently designed with a similar feature to Pearce’s design hence there was no infringement incurred.

The CDPA 1988 s4 defines “artistic work” as:

- (a) a graphic work (including painting, drawing, diagram, map, chart, or plan, engraving, etching, lithograph, woodcut, or similar work), photograph, sculpture, or collage, irrespective of artistic quality,
- (b) a work of architecture being a building or a model for a building, or
- (c) a work of artistic craftsmanship

Whilst copyright law protects architectural drawings and plans, it does not protect the ideas and concepts embodied in them. It can be argued that in this way, the development of architectural ideas and concepts is encouraged without restriction. Copyright only subsists in the form of the expression of ideas, but not taking the ideas and applying in works of architecture. This is to prevent the monopoly of the common stock of architectural ideas. It can be argued that the law is not there to protect the field of architecture from development of ideas evolving from prior concepts. In the area of

¹ (2002) 25(2) I.P.D. 25011; [2002] E.C.D.R. CN2.

music and composed works, a composer may listen to a tune and integrate it into his own composition subconsciously without realising that a similar exists. In *Gomme (E.) Ltd v. Relaxateze*², Walton J, stated that there was no copying, although he accepted that subconscious copying was possible because a tune can remain in human memory for a considerable duration. Similarly, if an architect inspects an architectural plan, he or she may be influenced by the prior viewing. It is possible that he or she may take the architectural concept and style embodied in the plan and applies to his or her own design in an original plan produced by him or her without any infringement.

The main principle is to distinguish whether the plans are an expression of concepts and ideas therefore legitimate to use, or the manifestation of original concepts or styles of the designers therefore constitutes a copyright infringement if they are copied without permission. However, adaptation of an idea is permitted. In a lengthy case of *Jones v London Borough of Tower Hamlets*³, the claimant designed a scheme for 80 new homes for a developer employed by Tower Hamlets Borough Council. The Council had to take over the development when the developer became insolvent. The architect sued for copyright infringement based on his initial design being copied by the council without permission and no payment was made. There were apparent similarities between the designs. The architect's claim failed. The principle was that architects are allowed to adapt an idea provided that there is no copying of the copyright protected works.

Architectural concepts and ideas may be encouraged to be developed but not copied. It can be argued that if the architectural drawings contain "unique" designs, they can be under copyright protection. To ensure their copyright is protected, architects and engineers should grant the permission explicitly for the use of their plans, clearly identify which rights are assignable to the third parties, normally with financial implications.

Copyright is a negative right which inhibits the duplication of the work of the authors by others. Copyright subsists in the form in which ideas are expressed, but not in the ideas themselves (*Jones v London Borough of Tower Hamlets*⁴ [2001] R.P.C. 23). Architects and engineers create particular designs for specific clients. These designs are mostly unique and ought to be protected under the law of copyright. There are clauses concerning copyrights in all professional appointments, but the construction industry seems to pay little attention to intellectual property rights. The copyright in drawings and specifications is normally owned by the designers – architects and engineers. There appears to be a lack of reported cases which have examined these issues as most of disputes are settled either via arbitration or mediation outside the court. They are confidential or negotiated directly and are not normally reported as court cases.

3 Authorship

The author, who creates the work, is the original owner, the first owner as defined in s91 of the Copyright, Designs and Patents Act 1988. There are three main requirements

² [1976] R.P.C. 377.

³ [2001] R.P.C. 23.

⁴ Ibid.

to prove the originality of a design which qualifies for the protection under s213(5) of the CDPA 1988.

1. The design must not be commonplace in the field of design in question under s214(4)
2. The design must originate from the designer; Petersen J, in *University of London Press v University Tutorial Press*⁵ [1916] decided that the word “original” does not necessarily mean that the work must be the expression of inventive thought.
3. The design must not have been copied

The author is a person who creates a work. One of the characteristics of architectural works and designs are their interdisciplinary nature. The law should protect co-authors, and joint authors cannot be liable to one another for copyright infringement. The joint authors should be allowed to exploit the copyright independently and can independently licence the use of the copyright. The joint authors, neither of them should be exclusively responsible for any of the design and the details can be used from one design to another.

There is no requirement for registration on copyright. When the above are fulfilled an unregistered design right will subsist automatically in the design.

Royal Institution of British Architects (R.I.B.A.) conditions of appointment contract has an expressed term stating that unless an alternative agreement has been made the architect retains his copyright in his drawings and the work executed by reference to them. The architect owns the copyright in the drawings and any documentation produced in the course of the project. Therefore, a builder who physically constructs a work of architecture cannot claim copyright even he builds it. In *Meikle v Maufe*⁶, the Court held that the architect should own the copyright in the design documents and the resulting building. Under the CDPA 1988, the architect also enjoys the moral right to be identified as its designer of the building, however these rights have to be asserted. The RIBA conditions of appointment contract express it clearly that the architect remains the author of the architectural works. In situations where a building contractor building for an employer, the commissioning client (*Cyprotex Discovery Ltd v University of Sheffield*⁷) or an employee working for an employer, the employer, in both situations, is normally the author and owner, unless there is an agreement between them stating the contrary (s11(2), CDPA 1988). If a client would like to make use of the calculations, plans, or memoranda, the contract with the designers has to express as provisions. In *Leicester County Council v Michael Faraday & Partners Ltd.*⁸, the local authority claimed that it was the owner of all the documents and plans which were produced by the defendants, who were surveyors contracted to undertake the valuations of the properties. The court held that documents were owned by the defendants because the documents were their expert works, unless there was a contract indicating these provisions.

⁵ [1916] 2 Ch 601 242, 260.

⁶ [1941] 3 All ER 144.

⁷ [2004] RPC 887.

⁸ [1941] 2 KB 205.

4 Assignment and Licence

Architects and engineers as designers normally have two ways of dealing with their copyright: they are assignment and licence. The first method is to expressly assign their copyright to the employer in terms of engagement which must be recorded in writing (ss 90-92, CDPA 1988). This copyright assignment must be stated clearly in the contract signed by both parties. Provided the architect or engineer does not replicate the whole design in another architecture s 64, CDPA 1988 allows the designer to retain the rights to reproduce some parts of the design for development, to be reused in another project. In general, architects prefer licensing than assigning copyright of architectural plans. Licensing provides architects freedom of future use and modification of their design for development. Clause 6.2.1 of the Royal Institute of British Architects, Conditions of Engagement (CE/99) allows Architect to suspend the licence of copyright materials owned by him or her if the client is in default of payment of their fees. The Architect needs to give a seven-day notice for the suspension. The architect can also authorise a third party to use his copyright materials but he continues to retain the copyright. If there is no expressed agreement on copyright, the employer, that is, the commissioning client will have an implied licence to use the drawings and plans for him according to the fee paid to the architect. However, a licence to use the architectural plans may not be implied in every contract with the architects, especially when the architects have charged only a nominal fee for the drawings. In *Stovin-Bradford v Volpoint Properties Ltd*⁹, a licence to use the drawings to construct a building was not implied because the architect had only been paid for the work for the purpose of planning permission. Salmon LJ said that even if the architects were to charge the full scale fee for preparing drawings for a planning permission they would not necessarily imply a licence to use the drawings for erecting the building, if they were not supervising the construction of the work. In *Blair v Osborne and Tomkins*¹⁰, Jacobs J. lays down the principle:

“it seems to me that the principle involved is this; that the engagement for reward of a person to produce material of a nature which is capable of being the subject of copyright implies a permission, or consent, or licence in the person giving the engagement to use the material in the manner and for the purpose in which and for which it was contemplated between the parties that it would be used at the time of the engagement.”

The normal use of the plans may be for the alteration by a different architect, the printing of a developer's brochure for enquiring potential buyers (*Robert Allan & Partners v Scottish Ideal Homes*¹¹). Unless agreement is made, certain unique distinctive features in the design cannot be used in an extension of the original building (*Meikle v Maufe*¹²).

Both the Royal Institute of British Architects, Conditions of Engagement (CE/99) Clause 6.2 and the Association of Consulting Engineers Conditions of Engagement 1995 (2nd ed, 1998) Agreement D have expressly licensed the commissioning client to

⁹ [1971] Ch 1007, [1971] 3 All ER 570, [1971] 3 WLR 256, CA.

¹⁰ [1971] 2 QB 78, [1971] 1 All ER 468, CA at p235.

¹¹ [1972] SLT (Sh Ct) 32.

¹² [1941] 3 All ER 144.

copy and use, and also allowing other consultants and contractors to copy and use the drawings and documents for the purposes related to the project. This is under the condition that the fees have been paid according to the agreement, and the architect has completed the scheme design, detailed design and production information.

The reason behind an implied licence is explained below. The client who commissions an architect to design for him does not entitle to use the designs. It is unlikely to settle on an agreement between the client commercial efficacy and the architect's interest. Hence, if the architect has not have an expressed copyright on the design, a licence to the commissioning client will normally be implied. Most architects and engineers favour the granting of a licence than assignment of their copyright.

In *Glengate-KG Properties Ltd v Norwich Union Fire Insurance Society Ltd and others*¹³, although a licence was granted to Glengate to reproduce the designs after completion of work, the architects had reserved the copyright in the drawings and designs, as explained in Clause 15 of the Royal Institute of British Architects (RIBA) Explaining an architect's service quoted below:

15. Copyright

Unless an alternative agreement has been made, the architect owns the copyright in the drawings and documentation produced in the course of the project. The client will usually have a licence to copy and use them only for purposes related to the project. If stage payments have been made to the architect, and the stage involving the planning submission has been completed and paid for, then the client has the right to use the relevant drawings for the purpose for which they were prepared (i.e. the planning application), regardless of any subsequent dispute which results in the withholding of fees owed for any following project stages¹⁴.

5 Infringement

The CDPA 1988 s16 provides exclusive rights to owners of copyright to do or authorise others to do or to act in the UK in relations to those works. These rights include copying the work, issuing copies of the work to the public, performing, showing or playing the work in public, broadcasting the work and making adaptation of the work. In fact, the adaptation work is an original copyright work in its own rights. For an infringement to take place, the infringed work should be copyrighted, a right should be infringed, and there was no defence. The copyright owners do not need to prove that the defendant had knowledge that the work was copyright protected, it is strict liability. There are two forms of infringement: primary infringement and secondary infringement. Primary infringements are restricted acts carried out by a person himself or herself or by a person being authorised to do so, without the permission of the copyright owner no matter what mental state the defendant is in. The intention of the defendant is

¹³ [1996] 2 All ER 487.

¹⁴ Clause 15 of the *Royal Institute of British Architects (RIBA) Explaining an architect's service*.

irrelevant¹⁵. Secondary infringements are dealings with infringing copies of a work. They are concerned with more serious infringements of larger scale with actual and constructive knowledge.

5.1 Primary Infringements

In *Francis, Day and Hunter v Bron*¹⁶, the Court of Appeal suggested a two-pronged test for infringing reproduction. Firstly, there is a sufficient degree of objective similarity between the two works; secondly, there is some casual connection between the original and the allegedly infringing work. For objective similarity, there is no requirement for the works to be identical before they are considered as infringement. If a substantial part has been copied, it can constitute an infringement. “Substantial part” is a qualitative test rather than a quantitative one as in *Ladbroke v William Hill*¹⁷. The court will identify sufficient degree of similarity with the original work by an objective test as in *Designers Guild Ltd v Russell Williams (Textiles) Ltd*¹⁸. However, it is not considered to be an infringement if the similarity is devolved from a casual connection with the earlier work. The later work can be qualified to obtain separate copyright even though there is an infringement involved in the earlier work.

Copying according to CDPA 1988 s17(2) includes storage of the work in any medium which includes digital storage. According to CDPA 1988 s17(3), making a copy in three dimensions from a two-dimensional work and vice versa is a primary infringement.

5.2 Secondary Infringements

Secondary infringement usually involves commercial infringement of reproducing copies of copyrighted work including possessing, selling, exhibiting or distributing infringing copies¹⁹ with knowledge of the infringement or at least hold reason to believe that there is an infringement. By “reason to believe” it means the knowledge of facts that a reasonable man would come up with the relevant belief as long as he or she has sufficient time to acquire the reasonable belief after investigation of the facts²⁰.

Copyright infringement is an unauthorised violation of the exclusive rights of the author and owner of an intellectual property. The copying of architectural plans by duplicating from the original plans is an example of copyright infringement. A perfect replica is not necessary to prove infringement, as long as an ordinary average layman would realise that there is an appropriation from the original work. Angela Adrian (2008) suggested that

¹⁵ CDPA 1988 ss16-18.

¹⁶ (1963) CA [1963] Ch 587, [1963] 2 All ER 16.

¹⁷ *Ladbroke Football Ltd v William Hill Football Ltd* [1964] 1 WLR 273.

¹⁸ [2001] 1 WLR 2416.

¹⁹ CDPA 1988 s27(2).

²⁰ *Sillitoe v McGraw Hill Book Co (UK) Ltd* [1983] FSR 545.

“The test is whether the plan or building was a copy of the concept or style and therefore legitimate or a copy of the author’s original manifestations of that concept or style and therefore an infringement.” (Adrian, 2008, p529).

5.3 Remedies for Infringements

Under s62 of the CDPA 1988 the remedies for copyright infringement are damages, injunctions, accounts or otherwise. For architectural works, it is unlikely that the court would order the demolition of the building because of the cost involved; hence, the remedies are likely to be in monetary compensation. Another remedy for architectural works would be to register a caution under the Land Registration Act 1925, s54(1) against the property which is alleged to be infringement of copyright of architectural plans owned by an architect (Arnold, 1997).

5.4 Defences

There are a number of exceptions available as defences under the CDPA 1988. These include fair dealing, incidental inclusion and artistic works on public display. Under Copyright, Designs and Patents Act 1988, ss 28-76, the copying of copyright works is allowed for fair dealing if the purpose of using the materials is for research and private study, criticism or review, reporting current events or incidental inclusion, provided it is not for substantial financial gain.

The CDPA 1988 s31(1) allows the incidental inclusion of copyright materials not to constitute an infringement, for example, an informal photographic portrait of painting or sculpture or in a television broadcast. This rule also permits certain acts in respect of buildings, models of buildings and works of artistic craftsmanship which are located permanently in a public place or in premises open to the public. Architecture in public location may be made the subject of a graphic work, a photograph or video, or broadcast as a visual image without copyright infringement²¹.

6 Copyrights Protection Measures

In the United Kingdom, Copyright law protects the creators of designs from having the drawings published, copied or broadcast in any way unless permission is given. Copyright protection is an automatic right, however, if there is any dispute, evidence is necessary to support any claim. In order to deter infringements, architects and designers can protect their drawings by using copyright notices for example including © or using the word “copyright” and ensure they are properly marked in the drawings as shown in Figure 1. They can register their design with verifiable proof of date and content. Registration can be submitted online or by post. Other supporting evidence on the evolution of ideas and development of the design would make a stronger case in the event of a dispute. In the design of architecture, it is quite often that it is a joint venture among the professional designers – architects, structural engineers, building services engineers and other designers. Agreement between the co-authors is an effective way to protect the copyright of individual professionals.

²¹ CDPA 1988 s 62.

Rev	Date	Description	Amended
A	05 06 2006	Design of ETFE atrium roof amended	DL






The Open University
Jennie Lee Building

SECTIONS
1:200 (A1)
05 JUNE 2006
© Swanke Hayden Connell International Limited

2189-P0301A

Figure 1. Example of copyright notices using © or the word “copyright” marked in the drawings

How effective is the UK’s Copyright Law in protecting the rights and interests of architects and designers? The main protection in UK Copyright Law is the written form of the design in recorded drawings, but not the ideas contained in the architectural design. The design has to be created by a qualifying author, such as an architect or an engineer, and is original and tangible. An architect can licence a client or other design professions to use the drawings, but the copyright normally remains with the architect who creates the drawings. The Royal Institution of British Architects (RIBA) has RIBA standard form of appointment contract which protects architects. If fees due to them are not paid, the licence would not be valid. It is an infringement of copyright if the drawings are copied, distributed or rented to other designers or to make alterations based on the original drawings without the author’s permission in the form of a valid licence. If a third party uses the plans for construction knowingly that it is an infringement of copyright, they are in breach of copyright law. Apart from copyright there are other forms of protection for architects’ work such as unregistered design rights, registered nationally as registered designs and European design rights. These issues are beyond the scope of this article.

The legislation in the UK is not clear concerning whether architectural designs are eligible to be registered as there is no specific exclusion or prohibition in doing so. It can be argued that certain modern buildings are likely to be industrial or handicraft items, but it is uncertain whether designs of buildings constructed by assembling prefabricated units can be registered as “three-dimensional products”. It was reported that the registration of concrete shelter was rejected. It remains unclear whether the design of a building can be registered or not.

One controversial issue in relation to the protection of architectural work is in regarding whether it should be afforded the same protection as a sculpture which is classified in the CDPA 1988 s 1 as an artistic work. Most architectural works are artistic works designed for a purpose; they are usually functional and utilitarian rather than purely artistic. Hence, it would be irrational to treat them as sculptures for copyright protection. A better scheme of protection of architectural works is to treat them as a sui generis category separated from sculptures and any other artistic works.

7 Proposed Law Reform

In the UK, under the Copyright Act 1911, architectural drawings were considered as literary works including maps, charts and plans. However, under the Copyright Acts 1956 and 1988, architectural drawings are afforded protection under artistic works. Neither of them provides satisfactory protection to architects and engineers as designers.

It is proposed by the author of this research that the designers and authors of architectural plans and works of architecture should be more adequately protected in the copyright for:

1. Two-dimensional representations such as architectural plans and drawing
2. Three-dimensional full scaled representations such as physical buildings including fixed structures
3. Three-dimensional physical scaled models
4. Digital representations such as digital Computer-aided design (CAD) models, photographs

7.1 Two-Dimensional Representations

The scope of afforded protection should realise that design in architecture normally incorporates selection, coordination or arrangement of elements. Although some of the design elements are not protectable, the integration and arrangement of these elements into an original new design should be allowed to be protected. In *Shine v Childs*²² the court held that the Olympic Tower had met the criteria for “copyrightability” realising an “unmistakable dash of originality” need be demonstrated, high standards of uniqueness in creativity are dispensed with”. The combinations of the architectural design elements of twisting towers, diamond-windowed facades, set-backs and support grids are “dash of originality”. Furthermore, architects and engineers may include novel protectable elements into a standard common features, which may be not be able to be protected, but the protectable novel elements should still remain protected. The philosophy behind these proposals is in line to what were reported in the U.S. Congressional Report for Architectural Works Copyright Protection Bill in 1990.

7.2 Three-Dimensional (3D) Full Scale Representations

The current UK copyright law considered copyright to be infringed if the elevation of the building with artistic characteristics is similar to the original plans. The scope of protection of works of architecture should extend to overall form, arrangement and composition of spaces and elements in the architectural design. Apart from the external visual characteristics the interior of the architecture should also be adequately protected.

The architect should be adequately protected if the 3-D full scaled building is obtained from the 2-D architectural plans as in *Hunt v Pasternack*²³.

²² 382 F.Supp at 610.

²³ 192 F.3d 877, 879-80 (9th Cir. 1999).

7.3 Three Dimensional Physical Scaled Models

The tradition of representation of an architectural works at the design stage is by three-dimensional physical scaled models. The CDPA 1988 s62, models of building were allowed for unauthorised but legitimate copying such as making a graphic representation, making a photograph or filming, broadcasting, and these can be distributed to the public copies of the representation of the models. The copyright owners of these models should be given as a sui generis protection under architectural works.

7.4 Digital Representations

When the CDPA 1988 was enacted, the Internet was not widely used; the problem of pirating architectural designs did not exist on any appreciable scale. Nowadays, as more individuals use personal computers and the Internet, architectural plans and designs can easily be copied and traded by many users. Nowadays, architectural practices produce drawings for their designs with CAD tools. AutoCAD® is an American CAD software application which supports 2D and 3D design and drafting. Architectural designs can be represented in 2D and can be modelled in virtual reality three-dimension models for visualization. It is important to protect architectural works and designs in the form of electronic models and Computer-Aided Design (CAD) drawings such as in AutoCAD® format when they can be downloaded, copied, stored and modified easily.

Newman²⁴ had raised issues on problems with architectural works stored in computer systems. Most clients acquire the properties in the copy of a design from the architect. Newman argued that the final form of design represents a document which becomes the property of the client, even though they may not be printed out in hardcopy. Although the clients and the designers can include in to the contractual documents copyright or licence clauses, it is proposed that the designers should be able to be protected by legislation automatically. They have to rely on implied terms creating uncertainties if they had not agreed on a granted licence.

It is difficult to claim for infringement of architectural works protected by copyright, when it is digitised and made available over the Internet. It is impossible to consolidate claims for infringement in one location, as the infringers may be under different jurisdictions, and they may be unknown to the copyright owner. It is proposed to apply a single universal law in dealing with infringements over the Internet.

The EU Copyright Directive has focused on the harmonisation of copyright laws in the information society²⁵. It extends the rights granted to copyright holders, in the following broad, technology-neutral exclusive rights including the right to authorise or prohibit reproduction including temporary reproduction, the right to authorise or prohibit the

²⁴ Newman, P (2002) "It's mine, not yours: Intellectual Copyright", *Construction Law*, (2002) 13 6 Cons. Law 14, 1 July 2002.

²⁵ *Copyright Directive 2001/29/EC of the European parliament and of the council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society*

communication, or the making available to the public, of a work or other subject matter; and the right to distribute copies of the work.

8 Reform Proposal

The principles of providing appropriate copyright protection to designers such as architects and engineers are to encourage creativity in design. However, such protection should not hinder creative design and discourage future development in works of architecture. The copyrighted materials should be allowed to be used in fair dealing such as research or private study, review and criticism, reporting current issues and knowledge management, provided the main purpose is not for commercial financial gains.

The author's preliminary proposal on the protection of architectural works and designs in the U.K. takes its cues from the AW PCA 1990. From the above discussion it is apparent that the most effective way to protect the rights and interests of architects and designers of their architectural works and designs is by sui generis protection similar to the proposal in this section.

(a) Definitions

- (1) "Architectural works" means "the design of a building as embodied in any tangible medium of expression, including buildings, three-dimensional models, architectural plans, or drawings and electronic representations. The work includes the overall form as well as the arrangement and composition of spaces and elements in the design, including novel functional features."
- (2) "Design" means "the design of any aspect of the shape or configuration (whether internal or external) of the whole or part of an article, other than surface decoration"²⁶.
- (3) A design is "original" if the architectural works is the result of the designer's creativity that provides a distinguishable works over prior works, and is not the result of directly copied from other sources.
- (4) "Design document" means "any record of a design, whether in the form of a drawing, a written description, a photograph, data stored in a computer or otherwise"²⁷.

(b) Copyright Protection

- (1) Architectural works and designs shall be a sui generis statutory subject matter of copyright protection.
- (2) Architectural works and designs shall be protected in its own category separated from "artistic works"²⁸ or "pictorial, graphic and sculptural works"²⁹ and in compliance with the requirements of the Berne Convention.

²⁶ Wording in SI2003/2498, reg 2(1), Sch 1, Pt 1, para 1, 8(3).

²⁷ Ibid.

²⁸ CDPA 1988 s 1.

²⁹ US Copyright Act 1976.

- (3) Architectural works and designs that are independently created and that express high degree of creativity shall be protected. Architectural works and designs that are not new or that do not significantly differ from existing designs shall be protected if there is evidence that they were independently created from prior work.
- (4) The copyright owner of protected architectural works and designs shall have the right to prevent third parties not having the owner's consent from constructing or transferring the works or designs which is a copy, or a substantially similar copy, of the protected works or design, especially when the acts are undertaken for commercial gain purposes.
- (5) Protection shall vest at the time of fixation, when it is embodied in a medium of expression with sufficient permanence, even when it is in a transient form in the memory of a computer, especially for the works and design which are registered with the Copyright Office.

9 Conclusion

Adrian³⁰ said recently that architecture has its special place in the law of copyright. While an architectural drawing can be protected by the UK copyright law, the law does not restrict the development of creative ideas and concepts but prevents the copying of plans.

Having reviewed the evidence and the relevant law, it can be concluded that both the jurisdictions of the U.S. and the U.K. do not adequately protect the rights and interests of architects and designers. There is a need for reform in respect to the protection of the rights and interests of designers on architectural works and designs. The authors' creativity on the architectural designs which are original even if they are functional shall attract appropriate rewards. Architecture should be protected in its own category separated from "artistic works", "graphic works" or the American category of "pictorial, graphic and sculptural works". Because of their nature and characteristics, architectural works and designs should have *sui generis* protection. Most buildings are designed to be functional and practical rather than simply for attaining an artistic purpose. Because of their integrated characteristics of being utilitarian, functional as well as artistic, architectural works and designs should not have to be evaluated under conceptual separability test. There is a need to establish a new test to satisfy their *sui generis* characteristics. Authors should have exclusive rights on architectural works and designs preventing them from unauthorised use and copying. If artistic and functional architectural works were protected, architects and designers might be motivated for greater creativity, and they would be rewarded for their creative works and designs.

There should be clear infringement standards. The authors of architectural works and designs should be more adequately protected in the copyright for the architectural drawings and plans, the full scaled representation in physical buildings and the digital representations in computer models of the works of architecture. In the UK copyright

³⁰ Adrian, A (2008) "Architecture and copyright: a quick survey of the law", Journal of Intellectual Property Law and Practice, 2008, Vol.3 No.8, 2008; pp524-529 Oxford University Press.

law, the physical buildings are considered as artistic works. Copyright protection should not only covers characteristics and original features of the buildings but also covers the methods or processes of construction. Concerning the copyright protection of computer models of architectural works, a single universal law should be introduced to tackle infringements of digitised works distribution of copyright works of architecture over the Internet without authorisation.

The most effective way to protect the rights and interests of architects and designers of their architectural works and designs is by *sui generis* protection. *Sui generis* protection has its benefits because it can take on the special characteristics of architectural works and designs requiring a different enforcement system with particular royalties relevant to the nature of architecture.

Apart from remedies of actual monetary damages, or issue of injunction reliefs, authors should be given the opportunities to influence the decision on their architectural works and designs. They should be given a chance to be consulted on the alterations and destructions of the constructed buildings.

10 Future Research

The advances in digital technologies have brought great challenge to copyright owners. There is no exception for copyrights protection for architectural works and designs. It would be worthwhile to further investigate whether the EU Copyright Directive³¹ and the UK implementation cover the protection of architectural works and designs effectively; and whether it is necessary to propose a reform to cover the architectural works and designs satisfactorily.

11 References

- Adrian, A (2008) "Architecture and copyright: a quick survey of the law", *Journal of Intellectual Property Law and Practice*, 2008, Vol.3 No.8, 2008; pp524-529 Oxford University Press
- Arnold, A (1997) "A New Remedy for Copyright Infringement" [1997] 12 EIPR 689
- Association of Consulting Engineers Conditions of Engagement 1995 (2nd ed., 1998) Agreement D
- Columbia Law Review Association, Inc "Copyright: Protection of Architect's Plans and Architectural Works" *Columbia Law Review*, Vol. 42, No. 2 (Feb., 1942), Columbia Law Review Association, Inc.) pp. 290-295
- Cook, T, Brazell, L, Charlton, S and Smyth, C (2004) "The Copyright Directive UK Implementation", Bristol: Jordans.

³¹ *Copyright Directive 2001/29/EC of the European parliament and of the council of 22 May 2001 on the harmonisation of certain aspects of copyright and related rights in the information society*

- Cook, T, Brazell, L, Charlton, S and Smyth, C (2004) *The Copyright Directive UK Implementation*, Bristol: Jordans)
- Dietz, B C (2001) "Case Comment - United States: copyright - copyright infringement for architectural works", *European Intellectual Property Review*, E.I.P.R. 2001, 23(5), N66
- Garnett, K., James, J. and Davies, G (1999) *Copinger and Skone James — On Copyright*, London: Sweet & Maxwell, 14th edn, pp.3–44.
- Krikke, J (2004) "Case Comment: Netherlands: Copyright—No Moral Rights Protection against Destruction" [2004] EIPR 155
- MacKenzie, C (2007) "Case Comment: Concrete Pty Ltd v Parramatta Design & Developments Pty Ltd" (2007) *Conveyancer and Property Lawyer* 461
- Mann, P and Denoncourt, J (2009) Copyright issues on the protection of architectural works and designs. In: Dainty, A (Ed) *Proceedings 25th Annual ARCOM Conference*, 7-9 September 2009, Nottingham, Association of Researchers in Construction, 707-16.
- Newman, P (2002) "It's mine, not yours: Intellectual Copyright", *Construction Law*, (2002) 13 6 Cons. Law 14, 1 July 2002
- Royal Institute of British Architects, Conditions of Engagement (CE/99) Clause 6.2
- Torremans, P (2008) *Holyoak and Torremans Intellectual Property Law*. (London: Oxford University Press, 5th. edn,)
- Weinberg, P (2004) Architectural Copycats – "Whose house are you building, anyway?" in Thomas, T (Editor) *Zoning and Planning Law Report*, West Thomson Publisher, Vol.27, No. 4, April 2004

Identity as a manifestation of organisational culture

Sai On Cheung¹, Hoi Yan Pang¹, So Yan Tam¹ and Hoi Yan Chan¹

¹Construction Dispute Resolution Research Unit

Department of Building and Construction

City University of Hong Kong

Email: bcsoc@cityu.edu.hk; hypang@cityu.edu.hk; soyantam2@student.cityu.edu.hk; hoichan5@student.cityu.edu.hk

Abstract:

Conducive organisational culture has been identified as an essential organisational stabiliser and growth driver. Organisational culture represents the assumptions, beliefs and norms of an organisation shared by members of the organisation. However, cultural values of an organisation are difficult to be examined. Organisational scholars conceived that organisational culture can be expressed in a number of ways described as manifestations. Identity is one of the informing artifacts. This paper examines the notion of identity in the context of organisational culture. How identity affects the propensity to leave the organisation is illustrated by a study with quantity surveyors.

Keywords:

Organisational culture, identity, performance, propensity to leave.

1 Introduction

Organisational culture (OC) has been used to explain why some companies are more successful than others (Cameron and Quinn, 1999). OC has a strong relationship with individual and organisational performance. It accounts for the existence of inter-group comparison, competition and productivity of an organisation (Schein, 1992). In these regards, OC is believed to be able to define appropriate behaviour as well as motivating individuals. For example, OC governs the way an organisation processes information (Maloney and Federle, 1991). Sexton and Barrett (2003) also found that OC had guided diversifications by integrating new technologies or other strategies to achieve organisational success.

Organisational scholars advocated that individual and organisational working are the symbolization of organisational values. It is a fact that construction professionals have to work together in a construction project despite cultural difference of the organisations they belong to. It is not uncommon to note that divergence in OC has led to conflict within project coalitions. Empirical studies of OC in construction industry have been carried out in different geographical locations such as China, United Kingdom, Turkey and Singapore. The findings suggested that better understanding of OC could greatly improve organisational performance and effectiveness. Furthermore, Ankrah and Langford (2005) found that there is a marked cultural difference between consultants

and contractors and highlighted the importance of studying OC in project-based industries.

According to the recent surveys conducted by the Hong Kong Institute of Human Resources Management (HKIHRM), the overall turnover rate in the construction sector has dramatically increased from 11.26% at 2006 to 20.46% at 2008. Employees in construction appear to leave their organisations quite readily. In this connection, the factors affecting employees' propensity to leave have attracted increasing interest in the Hong Kong construction community.

Other than economic factors, the propensity to leave is directly related to the psychological stress of employee, sense of self and organisational culture. Based on these theoretical anchors, this paper aims to illustrate the notion of identity in the context of organisational culture and articulate a framework for analysing the leave propensity of quantity surveyors.

2 Literature review

2.1 The concept of organisational culture

The concept of culture is difficult to be expressed explicitly; instead, it can be sensed, felt and perceived. Culture is embedded in groups; hence each group must always be clearly identified due to the unique basic assumptions. Understanding culture is equal to the understanding of basic assumptions and values.

Desphande and Webster (1989) defined OC as “the pattern of shared values and beliefs that help individuals understand organisational functioning and thus provide them with norms for behaviours in the organisation”. Furthermore, Schein (1992) considered OC as “a pattern of shared basic assumptions that the group learned as it solved its problems of external adaptation and internal integration that has worked well enough to be considered valid and therefore to be taught to new members as the correct way to perceive, think and feel in relation to those problems”.

Schein (1992) proposed a three-level organisational culture model; 1) artifacts, 2) values and 3) basic assumptions. This model has widely been adopted and applied in OC studies (Hatch, 1993). Hatch (1993) further developed Schein's ideas (1992) and constructed a Cultural Dynamics Model as shown in Figure 1. Hatch (1993) described that culture is constituted by manifestation, realization, symbolization, and interpretation processes which can start from any element clock-wisely and anti clock-wisely. The model presents the relationship between basic assumptions, values, artifacts and symbols (i.e. organisational image). From Hatch's perception (1993), identity can be examined from the manifestation process between values and assumptions of an organisation. The degree of identity stands for the understanding and adaptation of organisational values and assumptions.

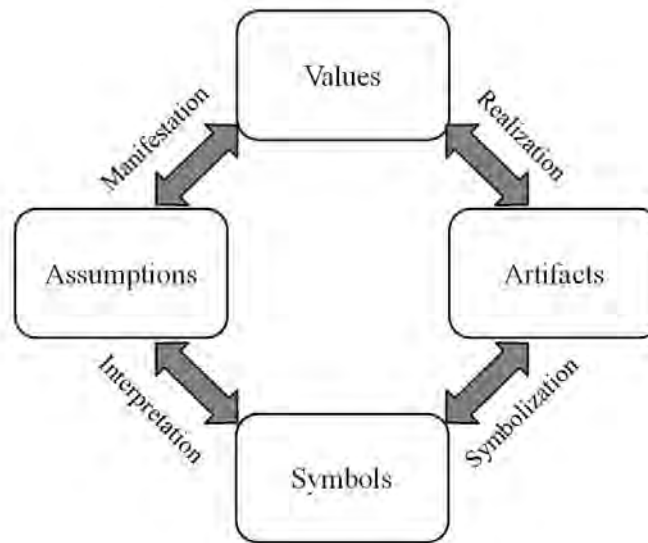


Figure 1. The Cultural Dynamics Model
(Source: Hatch, 1993)

2.2 Measuring organisational culture

Conflicts among organisational members have been identified as one of the principal causes of poor performance of construction organisations. Conflict is the result of disagreements or inconsistencies among intra-groups or inter-groups members (Jehn, 1995). For reducing conflict and improving performance in an organisation, it is important to identify the areas of difference between perceived and preferred OC of organisation members. Majority of studies on OC in construction mostly adopt Cameron and Quinn's (1999) Organisational Culture Assessment Instrument (OCAI) to examine the cultural profile (Zhang and Liu, 2006). Ankrah and Langford (2005) however selected 16 cultural dimensions based on the specific cultural attributes and orientations of the project participants to identify the cultural variability between organisations in project coalition. More recently, Cheung et al. (2010) proposed an organisational culture framework by using artifacts to identify OC in construction.

OC guides individual behaviour in conformance to those acceptable to the organisation. Harassing OC assists managers to devise organisational systems for organisational effectiveness. Measuring OC helps gauging the adherence or otherwise to norms expected.

2.3 The concept of organisational identity

Organisation is structured by groups, groups' norms and categorizes of groups can be described in details by social identity theory. Individual tends to classify himself into social categories or groups. To verify the degrees of individual's identity and sense of self from the organisation, it should pay attention to the performance implications or artifacts which are represented in the individual's mind as social identity.

The idea of identity is proposed by Mead (1934) who suggested that “identity arises in the process of social experience and activity, that is, develops in the given individual as a result of his relations to that process as a whole and to other individuals within the process”. Albert and Whetten (1985) thus proposed three definitions of identity: 1) central character, 2) distinctiveness and 3) temporal continuity.

2.3.1 Social identity and self-categorization theory

Social identity approach is an integrated theoretical perspective on the relationship between self-concept and group behaviour. Tajfel (1972) first introduced the idea of social identity theory by addressing how individuals conceptualize themselves in intergroup contexts. He defined social identity as “the individual’s knowledge that he belongs to certain social group together with some emotional and value significance to him of his group membership” (Tajfel, 1972). A basic human need, self-esteem, motivates social identification and individual’s behaviour. Thus, Crocker and Luhtanen (1990) have developed a self-esteem scale as a measure of social identity. The process for the formulation of identity is named as self-categorization in social identity theory (Tajfel, 1972).

On the other hand, Turner (1985) stated that self-categorization theory as the cognitive dimension of social identity theory. Social categorization of self cognitively assimilates self to the ingroup prototype and thus depersonalizes self-concept. Hogg (2001) further described self categorization in terms of the antecedence and consequence of social categorization of self and others. Individuals thus would have similar prototypes among people in the same group (i.e. ingroup prototypes) and different prototypes between people from different groups (i.e. outgroup prototypes). Self categorization transforms self-conception and individual’s attitudes, feelings and behaviour to ingroup prototypes.

The basic social cognitive elements from the process of self-categorization cause people to identify with group, to utilize themselves and others in group term and manifest group behaviours. Self-categorization in the context of organisation heavily influences group polarization and group cohesiveness which instantly affect the stability of workforce, organisational effectiveness, individual and organisational performance.

2.3.2 Identity in the context of organisation

Self-conception influenced by organisational processes is considered as organisational identification. Organisational identification represents the sense of belonging or a form of employee attachment to organisation. In this connection, Albert and Whetten (1985) concluded that “organisational identity is formed by a process of ordered inter-organisational comparisons and reflections upon them over time”. Organisational identification refers to the psychological linkage between an individual and the organisation by self-defining the affective and cognitive bond with the organisation (Edwards, 2005). Mael and Ashforth (1992) also advocated organisational identification as individual identification in terms of their membership in a particular organisation. While recognizing organisational identification from a cognitive perspective, organisational researchers tend to focus on the sharing of values and goals between an individual and the organisation (Ashforth and Mael, 1989) and the sharing of key characteristics of the organisation (Dutton et al., 1994), which intimately relate to the concept of OC.

The performances of employees are influenced by the degree of identification they attach with the organisation. Employees who identify themselves with the organisation are likely to be more motivated to work and achieve the organisational goals (Ashforth and Mael, 1989; Dutton et al., 1994). Ashforth and Mael (1989) found that organisational identification creates a cohesive work atmosphere, increases the levels of cooperation and satisfies human needs, such as belonging and self-esteem.

2.4 Linking organisational culture to identity

Generalising Mead's identity theory (1934), the meaning of 'I' and 'me' are simultaneously distinguishable and interdependent. The 'I' is "something that is not given in the 'me'"; whereas the 'me' is the self a person is aware of. Jenkins (1996) further articulated Mead's idea (1934) by synthesising identity from the internal (i.e. organisational culture) and external (i.e. organisational image) definitions of the organisation to explain the creation, sustainability and change of organisational identity. The concept of 'I' and 'me' within the processes of the Organisational Identity Dynamics Model (Hatch and Schultz, 2002) is presented as Figure 2.

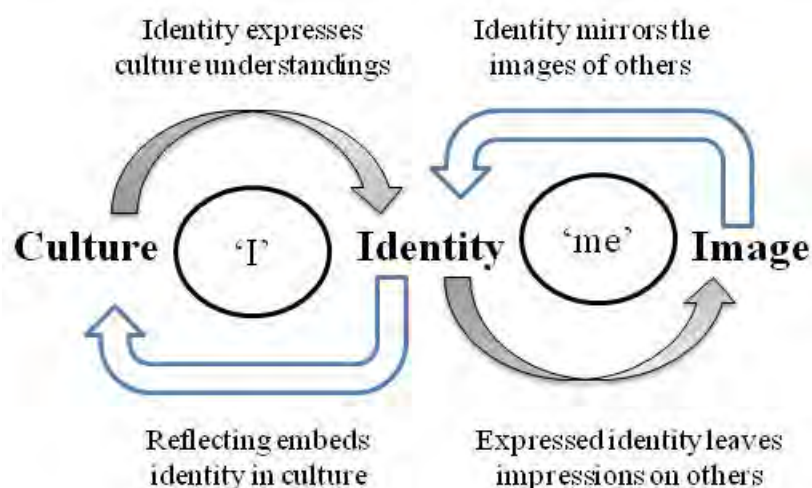


Figure 2. Organisational Identity Dynamics Model
(Source: Hatch and Schultz, 2002)

Culture and identity are closely connected because identity formulates a mindset that governs ways of doing things. Fiol et al. (1998) considered that 'an organisation's identity is the aspect of culturally embedded sense-making that is organisationally self-focused'. Hatch (1993) has a similar argument and addressed that OC provides context for forming organisational identities. He also explained that identity involves how individuals define and experience themselves which is influenced by their activities and beliefs interpreted from cultural assumptions and values. Therefore, organisational members' behaviour reflects their identity with reference to cultural understandings.

Hatch and Schultz (2002) further elaborated the relationship between culture and identity and mentioned that organisational identity needs to be theorized in relation to both organisational culture and organisational image. They thus proposed an Organisational Identity Dynamics Model which links identity, culture and image

together by mirroring (identity mirrors the images of others), reflecting (reflecting embeds identity in culture), expressing (identity expressed cultural understandings) and impressing (expressed identity leaves impressions on others) (Hatch and Schultz, 2002). This helps to explain how organisational identity is created, maintained and changed because identity is the immediate result of conversion between cultural self-expressions and mirrored organisational images.

3 Hypotheses development

3.1 Performance implication

Recent studies on OC in construction industry indicated that various participants (e.g. architects and contractors) bring different ways of thinking, attitudes and work practices. These incompatibilities hinder information sharing, control and coordination. In reality, it may not just affect the performance of project team members. For those who cannot sustain the psychological stress and affirm their sense of belonging, their propensity to leave the organisation increases.

Specific social identity is the salient self-regulation basis in a particular context. Ashforth et al. (2008) formulated identification from the core attributes of identification to the behaviours of identity. Behaviour is thus associated with identity. The levels of identification are directly associated with individuals' attitude, behaviour and thinking. From the aforementioned notions of organisational culture and organisational identity, employees decide to leave their organisations because of the lack of sense of belonging. To face with the high turnover in Hong Kong construction industry, an appropriate mean to comprehend employee's behaviour or propensity to leave is to identify the consequences of organisational identification from a review of the studies on organisational studies.

A hypothetical model of performance implication from organisational culture and organisational identity is shown in Figure 3. From organisational studies, the main reasons for employee choosing to leave are lack of organisational commitment and role stress. Organisational commitment is identified as employee's feeling of obligation to stay with the organisation (Allen and Meyer, 1990), whereas role stress represents role conflict and role ambiguity of an individual who is having incompatible and unclear expectations of his role in the organisation (Rizzo et al., 1970). In this connection, lower level of commitment and perceived high role stress would increase the propensity to leave.

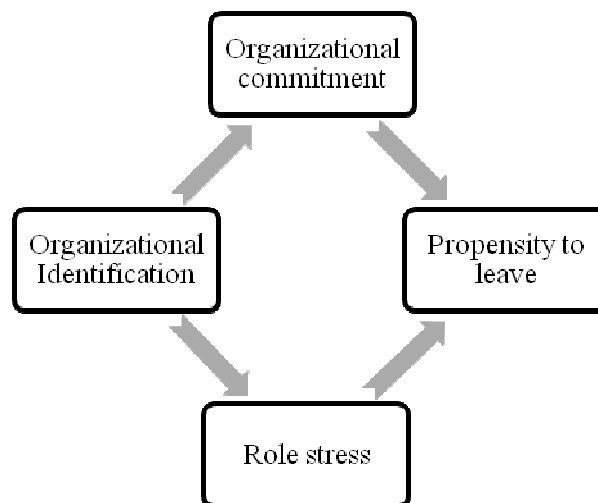


Figure 3. Performance implication from organisational culture and organisational identity

3.2 Organisational identification versus organisational commitment

According to Tompkins and Cheney (1983), organisational identification occurs ‘when, in making a decision, the person in one or more of his or her organisational roles perceives that unit’s values or interests as relevant in evaluating the alternatives of choice’. Gautam et al. (2004) proved that organisational identification is correlated to organisational commitment which represents individuals’ acceptance of the organisation’s set of values and objectives and reflects the willingness of an individual keep working within the organisation. Thus organisational commitment has positive effect on job satisfaction (Cohen, 1993), motivation (Mathieu and Zajac, 1990) and job performance (Mowday et al., 1979). Organisational commitment is especially important to professional organisations; Lui et al. (2001) found that accountants would have lower leaving intention under higher level of organisational commitment. Accordingly, the organisational commitment of quantity surveyors firms is examined at hypothesis 1.

H1: Quantity surveyors’ organisational identification is positively associated with their organisational commitment

3.3 Organisational identification versus role stress

3.3.1 Role conflict

Identification literature highlights identification as a process of self-categorization, named as formalization by Hall (1972). To integrate individuals’ self-concept with organisational goals and values or organisational characteristics, it generates individuals’ affiliation to the organisation and conformance to the organisational norms. At the same time, the formalization of individual with organisational goals also causes negative effect on an individual, such as alienation or uncomfortable feeling; it may increase individuals’ role conflict.

3.3.2 *Role ambiguity*

Employees may have role ambiguity when they have difficulty to be formalized with the organisation. Role ambiguity which refers to the discrepancies between the job-related information available and information needed to perform the job (Kahn et al., 1964). Thus, less formalized organisations with lower level of organisational identification may raise individuals' role ambiguity in their working organisation.

Therefore, hypothesis 2 and 3 are used to examine the implication of role stress in relation to organisational identification.

- H2: Quantity surveyors' organisational identification is negatively associated with their role conflict.
- H3: Quantity surveyors' organisational identification is negatively associated with their role ambiguity

3.3.3 *Propensity to leave*

The effect of role stress on commitment is an increase in the employees' propensity to leave. Their perceived enhancements in psychological well-being mitigate their job satisfaction and exhaust their working emotion (Ngo et al., 2005).

3.4 **Organisational commitment versus propensity to leave**

Furthermore, organisational commitment is negatively related to propensity to leave the organisation of an employee (Bosshoff and Mels, 2000). The affective organisational commitment is likely to reduce the employees' intention for leaving the organisation, which is going to be examined by hypothesis 4.

- H4: Quantity surveyors' organisational commitment is negatively associated with their propensity to leave their working organisation

3.5 **Role stress versus propensity to leave**

A stressful employee is more likely to quit his or her job. When an employee is having conflicts with organisational norms and values, or perceiving ambiguous during their works, it will lead to depression, lower job involvement and result in psychological withdrawal from the work group (Brief and Aldag, 1976). These implications of role conflict and role ambiguity are examined by hypothesis 5 and 6.

- H5: Quantity surveyors' role conflict is positively associated with their propensity to leave the working organisation
- H6: Quantity surveyors' role ambiguity is positively associated with their propensity to leave the working organisation

4 Method

4.1 Sample

Data were obtained from a questionnaire survey with 216 quantity surveyors in Hong Kong. A total of 66 responses were returned, representing a response rate of 30.56%. 56% of the respondents are male and 44% are female. Nearly 90% are working in quantity surveying consultancy firms. Close to 83% are below 29 years old and approximate 64% have less than 3 years working experience. 35% of the respondents have changed their employing organisations previously.

4.2 Measures

Respondents expressed their subjective opinions against the questionnaire statements on a seven-point Likert-type scale (1 = 'strongly disagree', 7 = 'strongly agree'). Organisational identification was measured by a five-item scale developed by Mael and Ashforth (1992). Organisational commitment was measured using a five-item scale adapted from Porter et al. (1974). The eight-item role conflict developed by Rizzo et al. (1970) was used; meanwhile, Rizzo et al.'s (1970) six-item role ambiguity scale was adapted for use in this study also. Propensity to leave the working organisation within the next few years was measured.

4.3 Analysis

Descriptive analyses summarise the attributes of the sample. The reliability of the measures was assessed by the correlation among the variables. Confirmatory factor analysis was then used to evaluate the reliability and validity of the collected data by using the AMOS programme. Regression analysis was employed to test the hypotheses.

5 Results

5.1 Descriptive and reliability analysis

From the aforementioned hypothetical model, the outcomes of organisational identification illustrate individual's propensity to leave their working organisations. Table 1 illustrates the descriptive statistics and inter-item correlation of the study's variables. Seven-point Likert scales give a midpoint scale as 3.5 and a maximum value as 7. The measured variables have a high mean value significantly greater than the midpoint scale. Role ambiguity gets the highest mean value (mean = 4.84); meanwhile, role conflict (mean = 3.94) is quite low from the perspectives of quantity surveyors. Regarding to the outcome, propensity to leave (mean = 4.32) is relatively significant in quantity surveying organisations. From the descriptive analysis, it is highlighted that the importance of organisational commitment and elimination of role ambiguity.

Organisational identification and role ambiguity have strong positive correlation of correlation coefficient of 0.598 ($p < 0.01$). Similarly, organisational commitment proved that positively correlated with organisational identification. Organisational identification is not significantly correlated with role conflict. However, it is noted that role conflict is positively associated with role ambiguity ($r = 0.515$, $p < 0.01$). In case of turnover,

employees' leave propensity is positively correlated to role ambiguity ($r = 0.371$, $p < 0.01$) and negatively related to organisational commitment ($r = -0.370$, $p < 0.01$). The reliability is determined by the value of Cronbach's Alpha with threshold value of 0.70 to 1.00 (Nunnally, 1978). The Cronbach's Alpha for the hypothesised model is 0.823 that achieves the criteria threshold.

Table 1. Descriptive statistics and inter-item correlations of study's variables

Variables	Mean	S.D.	OI	RC	RA	OC	LO
Organisational identification (OI)	4.64	1.14	1				
Organisational commitment (OC)	4.57	0.99	0.586**	1			
Role conflict (RC)	3.94	1.38	-0.055	0.021	1		
Role ambiguity (RA)	4.84	1.10	0.598**	0.515**	-0.149	1	
Propensity to leave the organization (LO)	4.32	1.62	-0.054	-0.178	0.371**	-0.370**	1

Note: N=66, * $p < 0.05$; ** $p < 0.01$

5.2 Confirmatory factor analysis

Confirmatory factor analysis is used to study the relationships between the dependent observed variables and continuous latent variables, where dependent observed variables are referred to as factor and continuous latent variables are referred to as factors. According to the confirmatory factor analysis, the initial model did not have an acceptable fit; modification indices were applied to improve the model. Table 2 provides the fit indices of the confirmatory factor analysis for the hypothesised model.

According to Browne and Cudeck (1993), RMSEA value ranging between 0.05 and 0.08 represents a reasonable fit. The hypothesised model is considered as an acceptable-fitting model which obtains threshold acceptable fit indices except GFI and AGFI. The model is not easy to achieve an acceptable GFI and AGFI because of the implication of sample size (Sharma, 1996).

5.3 Regression analysis

Regression is used to analyse the relationship between the dependent variable and independent variable of the hypothesised model that outlines the effect of organisational identification with association of organisational commitment and role stress. Table 3 shows the regression analysis. Hypothesised consequences for organisational identification are statistically significant. Organisational identification facilitates the generation of organisational commitment in quantity surveying organisations. Quantity surveyors' sense of belonging to their organisation directly influences their role ambiguity. Hypothesised consequences of role conflict and role ambiguity are

supported. Quantity surveyors are more likely to leave their organisation when they feel stressed. However, role conflict is not significantly related to organisational identification and organisational commitment is not the catalyst stimulating the leaving intention of quantity surveyors. Consequently, organisational identification associated with role ambiguity increases quantity surveyors' intention to leave.

The small sample size influences the model fit of the hypothesised model by confirmatory factor analysis. Nevertheless, the hypothesized models obtain an acceptable fit with regression. In this study, the results verify that quantity surveyors in Hong Kong have relatively high intimacy with their working organisations. Comparatively low role conflict within the respondents is the main cause inducing employees' propensity to leave. Besides, it is found that quantity surveyors in Hong Kong are under high role ambiguity, which is the root cause of quantity surveyors choosing to quit.

Table 2. Fit indices from confirmatory factor analysis

Fit indices	Before modification	After modification
CMIN/ DF (<2.00)	1.811	1.161
Root mean square error of approximation, RMSEA (0.05-0.08)	0.112	0.050
Goodness of fit index, GFI (>0.90)	0.658	0.787
Adjusted goodness of fit index, AGFI (>0.90)	0.593	0.705
Comparative fit index, CFI (>0.90)	0.724	0.953
Incremental fit index, IFI (>0.90)	0.733	0.956

Table 3. Regression Analysis

Independent variable	Dependent variable	Unstandardised Coefficients		Standardised Coefficients	<i>t</i> -test
		<i>B</i>	<i>Std. Error</i>	β	
OI	OC	0.519	0.090	0.586	5.783***
OI	RC	-0.064	0.145	-0.055	-0.438
OI	RA	0.573	0.096	0.598	5.970***
RC	LO	0.608	0.190	0.371	3.200**
RA	LO	-0.735	0.231	-0.370	-3.188**
OC	LO	-0.383	0.264	-0.178	-1.451

Note: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

6 Conclusion

Understanding OC is fundamental to understand what underpins an organisation's operation. Organisation culture is also considered as the glue holding individual and organisation together. Cultural profile thus provides a comparative benchmark for an

organisation that governs how people in an organisation behave, what assumptions govern their behaviour and what organisational systems influence the change process.

This paper indicates the relationship between organisational culture and organisational identification. Organisational identification can be distinguished from the manifestations of basic organisational assumptions because organisational assumptions underline values, and individuals infer their assumptions from known values. In this connection, OC gives identity of an organisation, and the sense of belonging so derived is the result of organisational culture and organisational identification.

7 Acknowledgements

The work described in this paper was fully supported by a grant from the Research Grant Council of the HKSAR (Project No. 112008).

8 References

- Albert, S. and Whetten, D.A. (1985), 'Organisational identity', In Cummings, L.L. and Staw, B.M. (Eds.), *Research in organisational behaviour*, Greenwich, CT: JAI Press, pp 263-295.
- Allen, N.J. and Meyer, J.P. (1990), 'The measurement and antecedents of affective, continuance and normative commitment to the organisation', *Journal of Occupational Psychology*, 63, 1-18.
- Ankrah, N.A. and Langford, D.A. (2005), 'Architects and contractors: a comparative study of organisational cultures', *Construction Management and Economics*, 23 (6), pp 595-607.
- Ashforth, B.E. and Mael, F.A. (1989), 'Social identity theory and the organisation', *Academy of Management Review*, 14, pp 20-39.
- Ashforth, B.E., Harrison, S.H. and Corley, K.G. (2008), 'Identification in organisations: An examination of four fundamental questions', *Journal of Management*, 34 (3), pp 325-374.
- Boschhoff, C. and Mels, G. (2000), 'The impact of multiple commitments on intentions to resign: An empirical assessment', *British Journal of Management*, 11 (3), pp 255-272.
- Brief, A. and Aldag, R. (1976), 'Correlates of role indices', *Journal of Applied Psychology*, 61, pp 468-472.
- Browne, M.W. and Cudeck, R. (1993), 'Alternative ways of assessing model fit', In Bollen, K.A. and Long, J.S. (Eds.), *Testing structural equation models*, Newbury Park, CA: Sage, pp 136-162.
- Cameron, K.S. and Quinn, R.E. (1999), *Diagnosing and changing organisational culture: Based on the Competing Values Framework*, Addison-Wesley, New York.
- Cheung, S.O., Wong, P.S.P. and Wu, A.W.Y. (2010), 'Towards an organisational culture framework in construction', *International Journal of Project Management*, in press.
- Cohen, A. (1993), 'Organisational commitment and turnover: A meta-analysis', *Academy of Management Journal*, 36 (5), pp 1140-1157.
- Crocker, J. and Luhtanen, R. (1990), 'Collective self-esteem and ingroup bias', *Journal of Personality and Social Psychology*, 58, pp 60-67.

- Desphande, R. and Webster, F.E. (1989), 'Organisational culture and marketing: Defining the research agenda', *Journal of Marketing*, 53 (1), pp 3-15.
- Dutton, J.E., Dukerish, J.M. and Harquail, C.V. (1994), 'Organisational images and member identification', *Administrative Science Quarterly*, 39, pp 239-263.
- Edwards, M.R. (2005), 'Organisational identification: A conceptual and operational review', *International Journal of Management Reviews*, 7 (4), pp 207-230.
- Fiol, C.M., Hatch, M.J. and Golden-Biddle, K. (1998), 'Organisational culture and identity: What's the difference anyway?', In Whetten, D. and Godfrey, P. (Eds.), *Identity in organisations. Building theory through conversation*, Thousand Oaks, CA: Sage.
- Gautam, T., Van Dick, R. and Vagner, U. (2004), 'Organisational identification and organisational commitment: Distinct aspects of two related concepts', *Asian Journal of Social Psychology*, 7, pp 301-315.
- Hall, R.H. (1972), *Organisations: Structure and Process*, Englewood Cliffs, N.J.: Prentice.
- Hatch, M.J. (1993), 'The dynamics of organisational culture', *Academy of Management Review*, 18, pp 657-693.
- Hatch, M.J. and Schultz, M. (2002), 'The dynamics of organisational identity', *Human Relations*, 55 (8), pp 989-1018.
- Hogg, M.A. (2001), 'Social categorization, depersonalization, and group behaviour', In Hogg, M.A. and Tindale, R.S. (Eds.), *Blackwell handbook of social psychology: Group Processes*, Oxford, England: Blackwell, pp 56-85.
- Jehn, K.A. (1995), 'A multimethod examination of the benefits and detriments of intragroup conflict', *Administrative Science Quarterly*, 40, pp 256-282.
- Jenkins, R. (1996), *Social identity*, London: Routledge, 1996.
- Kahn, R.L., Wolfe, P.M., Quinn, R.P., Snoek, J.E. and Rosenthal, R.A. (1964), *Organisational stress: Studies in role conflict and role ambiguity*, New York, John Wiley & Sons.
- Lui, S.S., Ngo, H.Y. and Tsang, A.W.N. (2001), 'Interrole conflict as a predictor of job satisfaction and propensity to leave: A study of professional accountants', *Journal of Managerial Psychology*, 16 (6), 469-484.
- Mael, F.A. and Ashforth, B.E. (1992), 'Alumni and their alma mater: A partial test of the reformulated model of organisational identification', *Journal of Organisational Behaviour*, 13, pp 103-123.
- Maloney, W.F. and Federle, M.O. (1991), 'Organizational culture and management', *Journal of Management in Engineering*, 7 (1), pp 43-58.
- Mathieu, J.E. and Zajac, D.M. (1990), 'A review and meta-analysis of the antecedents' correlates, and consequences of organisational commitment', *Psychological Bulletin*, 108, pp 171-194.
- Mead, G.H. (1934), *Mind, self and society*, Chicago, IL: University of Chicago Press.
- Mowday R.T., Steers, R.M. and Porter L.W. (1979), 'The measurement of organisational commitment', *Journal of Vocational Behaviour*, 14, pp 224-247.
- Ngo, H.Y., Foley, S. and Loi, R. (2005), 'Work role stressors and turnover intentions: a study of professional clergy in Hong Kong', *The International Journal of Human Resource Management*, 16 (11), pp 2133-2146.
- Nunnally, J.D. (1978), *Psychometric theory*, New York: McGraw-Hill.

- Porter, L.W., Steers, R.M., Mowday, R.T. and Boulian, P.V. (1974), 'Organizational Commitment, Job Satisfaction and Turnover among Psychiatric Technicians', *Journal of Applied Psychology*, 59, pp 603-609.
- Rizzo, J. R., House, R.J. and Lirtzman, S.I. (1970), 'Role conflict and ambiguity in complex organizations', *Administrative Science Quarterly*, 66, pp 150-163.
- Schein, E.H. (1992), *Organisational culture and leadership*, Jossey-Bass Publisher, San Francisco.
- Sexton, M. and Barrett, P. (2003), 'Appropriate innovation in small construction firms', *Construction Management and Economics*, 21 (6), pp 623-633.
- Sharma, S. (1996), 'An empirical investigation into the influence of managerial cognitions and organizational context on corporate environmental responsiveness', In Logsdon, J. and Rehbein, K. (Eds.), *Proceedings of the Seventh Annual Conference of the International Association of Business and Society*, Santa Fe, NM: March.
- Tajfel, H. (1972), 'La catégorisation sociale', In Moscovici, S. (Ed.), *Introduction à la psychologie sociale*, Paris, Larousse.
- Tompkins, P.K. and Cheney, G. (1983), 'Account analysis of organisations: Decision making and identification', In Putnam, L.L. and Pacanwsky, M.E. (Eds.), *Communication and organisations: An interpretive approach*, Beverly Hills, CA: Sage, pp 123-146.
- Turner, J.C. (1985), 'Social categorization and the self-concept: A social-cognitive theory of group behaviour', In Lawler, E.J. (Ed.) *Advances in group processes: Theory and research*, Greenwich, CT: JAI, Volume 2, pp. 47-74.
- Zhang, S.B. and Liu, A.M.M. (2006), 'Organisational culture profiles of construction enterprises in China', *Construction Management and Economics*, 24 (8), pp 817-828.

A Single Ethical Code for the Construction Industry – What do Young Professionals Think about It?

Jim Mason, Department of Construction and Property University of the West of England, Bristol, jim.mason@uwe.ac.uk

ABSTRACT

The theme of engagement with the next generation of industry leaders was spelt out recently in the Wolstenhome Report “Never Waste a Good Crisis”. An important element in any engagement is listening to the other’s views. This paper uses the topical issue of ethical improvement within the construction industry as an opportunity to take a sounding from thirty Masters’ level Construction students. The students were specifically asked for their views in response to the proposition that the industry needs a single ethical code. The responses address this question and some wider issues. The responses presented in this paper give a valuable insight into how young professionals see the industry developing over the coming decade.

The Single Ethical Code for Construction Professionals has been promoted in recent years by the Society of Construction Law. The views of the students deliver a sense of the need for change in the industry and there is broad agreement that a Single Code could help deliver this. The paper concludes with the key recommendations to take forward from the students’ views and provides further discussion points around the state of the industry as the new decade dawns.

Keywords: professional, ethics, enforcement, code, unethical behaviour, young professionals

1. Introduction

This paper returns to the subject of a Single Ethical Code (SEC) which the author last wrote about in 2008 for the COBRA conference in Dublin (Mason, 2008).

In that paper the reader was presented with the SEC promoted by the Society of Construction Law (SCL)¹. The Code has been described as “the first line of defence against corruption” (Uff, 2003). The advantages to the construction industry in adopting the code would lead to “an appreciable reduction in poor designs, shoddy workmanship, delays, claims, excessive charging, cost overruns and disputed claims”(Thornton, 2004). Unfortunately, there has not been much in the way of recent progress with the code. With a considerable degree of foresight, both writers seemed to appreciate this risk and said that the task they set themselves was not an easy one. His Honour Judge Thornton described the task of creating the code as “bold, difficult and somewhat presumptuous” as well as “brave, pioneering and highly desirable.”

The SEC set about the task of writing the code in 2003. The code was set out in the following terms:

¹ Available at scl.org.uk/ethics accessed on 14 June 2010

Statement of Ethical Principles

Ethical conduct is the compliance with the following ethical principles:

1. **Honesty** - act with honesty and avoid conduct likely to result, directly or indirectly, in the deception of others.
2. **Fairness** - do not seek to obtain a benefit which arises directly or indirectly from the unfair treatment of other people.
3. **Fair reward** - avoid acts which are likely to result in another party being deprived of a fair reward for their work.
4. **Reliability** - maintain up to date skills and provide services only within your area of competence.
5. **Integrity** - have regard for the interests of the public, particularly people who will make use of or obtain an interest in the project in the future.
6. **Objectivity** - identify any potential conflicts of interest and disclose the conflict to any person who would be adversely affected by it.
7. **Accountability** - provide information and warning of matters within your knowledge which are of potential detriment to others who may be adversely affected by them. Warning must be given in sufficient time to allow the taking of effective action to avoid detriment.

These principles were written to apply to the work of all professionals working in the construction industry, whatever their original qualification or affiliation and to individuals, whether they work for or on behalf of an independent professional or as a partner, associate, director or employee of a firm or company.

It was intended that the statement of good practice is in addition to any other professional code that may apply. Further, the statement of good practice was seen as part of and additional to contractual and other duties taken on under the civil law and potential breaches of the criminal law.

The Code identified unethical conduct as deliberate or reckless disregard for the ethical principles, as they would apply to the ordinary standards applicable to the activity being undertaken by reference to the recognised practice in that profession.

In my original paper I opined that the SEC would not make any difference where individuals were set on consciously behaving unethically. However, it did appear likely that it could make a difference in the situations where there is unintentional ethical transgression and people lacked the necessary guidance

In the intervening years since the launch of the code it is probably fair to say any clamour for its wide scale adoption has died away. In my view a contributing factor to this result was the fact that the SCL did not feel able to take any responsibility for enforcing the code. The SEC therefore appears destined for the same fate as many good and bold initiatives in the industry – big plaudits but little practical take-up. Notwithstanding this, the sentiment behind the code and the desire to improve things within the construction industry remain alive and well.

2. The Study

As part of our Masters provision at the University we run law themed modules one of which is Construction Contract Law. This module is delivered to students on the Masters in Construction Project Management, Masters in Law for Construction Professionals, Masters in Quantity Surveying and Graduate Diploma students in Quantity Surveying. In a seen examination the students were set a question asking them for their views on the prospects for the SEC. Their answers contained some valuable insight into the views and profiles of the Masters' students. I have set out their views anonymously in this paper with their assent.

The students are all in full time employment within the construction industry. They are therefore well placed to give their frank and honest appraisal of the industry they work in. A large proportion of the students are on the Graduate Diploma programme having completed a first degree in a non-cognate subject. The comparison between what they are experiencing in the construction industry and how this sits with their previous working experience is another interesting angle underlying their responses.

There are naturally limitations on the study. We discussed the SEC during contact teaching sessions and there is a risk that the students' views were influenced by my own. Whilst we take steps to engender and encourage free thinking in our students there remains the possibility that they are merely repeating back my own ideas. I have sought to remove the content where this did occur from the quotations set out below. I have concentrated instead on their views.

Whilst the examination question related solely to the SEC, many of the quotation selected below touch on other issues and underlying concerns about the state of the construction industry in 2010. It is these views as much as anything directly on the SEC which make this paper relevant for this conference and policy makers within the industry.

A further limitation of the study is that there is a risk that students were quoting from unreferenced sources. The assessment from which the data has been gleaned is a seen examination where the students are encouraged to prepare their answers prior to writing them out under exam conditions. The usual rules around strict referencing are relaxed slightly and several sources may therefore have been missed.

3. The Literature Updated

In my original paper I defined ethics as the moral principles by which a person is guided.² Professionalism has been described as being in possession and autonomous control of a body of specialised knowledge, which when combined with honorific status, confers power upon its holders (Greenhalgh, 1997). The exercise of this control by the respective professional bodies is often manifested in the promotion and enforcement of an ethical code³.

I rehearsed the argument that there were no such things as "business ethics" and that such a term was an oxymoron (Ferguson, 1994). The better position appeared to be that general concepts of ethics must be applicable to business, on the grounds that business exists not solely to suit certain individuals, but to serves society and, in addition, meets collective and social needs such as the promotion of sustainability and environmental aspects of the construction industry's activity (Cohen et al, 1998).

² <http://dictionary.oed.com> accessed on 14 March 2008

³ See for example, Royal Institute of Chartered Surveyors (2007) *Rules of Conduct for Members* available at rics.org

There have been developments in the Office of Fair Trading (OFT) investigation into bid rigging within the Construction industry. The efforts of the OFT to combat corruption culminated in fines totalling £129.2 million on 103 construction firms in England which were found to have colluded with competitors on building contracts. The OFT has concluded that the firms engaged in illegal anti-competitive bid-rigging activities on 199 tenders from 2000 to 2006, mostly in the form of 'cover pricing'.⁴ The response from the industry has been to take measures to ensure tendering procedures now comply with the governing regulations. For example, the United Kingdom's Construction Group which includes many of the bigger construction companies launched a Code of Competition Compliance on 20 August 2009⁵.

The dust has now settled on what was a difficult period for the industry. The lingering feeling inside the industry is that the OFT did not properly understand the way the industry works (Bingham, 2008). Another popular sentiment is that contractors had no choice but to put in high bids for work because clients would remove them from tendering lists if they did not. For their part, the impression is that clients have taken this on in part as fair criticism.

Pronouncements from the SCL on ethics have included a review by Professor Lavers of the position in eight European Countries on selected ethical issues (Lavers, 2008). The outcome of the study was to highlight areas of commonality on issues such as retention and termination of the contract whilst diversity appears on other issues such as the duty to warn and the role of the contract administration.

Elsewhere, there have been words of caution against the increasing trend for companies to promote their value statements in client facing literature (Zhang et al, 2008). This paper quotes an earlier paper (Lencioni, 2002) claiming that many value statements are "bland, toothless, or just plain dishonest" which "create cynical and dispirited employees, alienate customers, and undermine managerial credibility." There are obvious parallels with the pitfalls that could await an overly ambitious ethical code. This point is taken up by my colleague Peter Fewings in his book "Ethics for the Built Environment" (Fewings, 2009) "in the context of the built environment, professional ethics are often quoted in the form of professional codes of practice, but there is some criticism as to the effectiveness of these codes and some public suspicion that they might be self-serving."

The major development in terms of government sponsored initiatives has come in the shape of *Never Waste a Good Crisis* (Wolstenhome 2009). Subtitled "A report into progress since *Rethinking Construction* and thoughts for our future" this document has clear pedigree as the inheritor of the movement for improvement in the industry in the last seventeen years since the Latham report (Latham, 1994). The forewords to the report feature the major players during this period in the form of Sir Michael Latham, Sir John Egan and Nick Raynsford, MP.

The report primarily reinforces the best practice advice set out by *Rethinking Construction* (Egan, 1998). The report also contains some stark warnings about the

⁴ <http://www.offt.gov.uk/news-and-updates/press/2009/114-09>

⁵ <http://www.ukcg.org.uk/UK-Contractors-Group-UKCG-Code-of-Competition-Compliance-cbb744f>

financially difficult times to come and the limited progress made on the improvement agenda set. The report concludes with a challenge to the present leaders of the construction industry to move aside in favour of the next generation. The executive summary concludes: “If our present leaders do not feel up to the task, they should at least support the development of the next generation, who appear to understand very clearly what is needed.”

This statement that the next generation understand what is needed is cryptic. The purpose of this paper is to gain some insight into what this “clear understanding” amounts to, within the framework of ethical improvement within the industry.

4. Findings

As previously mentioned, the exam question answers being reviewed here relate solely to the SCL’s SEC. The scope of the responses go much wider than the code itself and bring in other points including: the general health of the industry, its attractiveness to recent graduates and some of the less savoury practices encountered.

The views of this group of students are particularly interesting and relevant given that the students did not come into the industry via the standard undergraduate route. These students have come through the Graduate Diploma route whereby their first degree in any related or unrelated subject gives the student academic credit and a shorter route to achieving the diploma qualification and opening the route to the student to accreditation with the Royal Institute of Chartered Surveyors. This route has been extremely popular at the University of the West of England. These newcomers to the industry bring diversity and fresh perceptions on established practice. It is therefore intriguing to ask “what do they think of the industry?” and “do they accept the status quo?” The importance of “customer relations” cannot be overstated to a business. Here is a group of extremely valuable “customers” who certainly merit being listened to most carefully.

The responses have been grouped together the themes set out below. The themes were identified as capturing the essence of the comments voiced by the students and through which the accompanying narrative could weave together their responses. All the student quotes appear in italics and speech marks.

- Definition
- Context and Views on the Current State of the Industry
- Recent Ethical improvements and Opportunities
- The Obstacles to Further Improvement
- Purpose/Limitations of the Code
- The Positive Steps Required
- The Timetable for Action
- Other issues/Last Thoughts

All good legal topics start with a definition and several students attempted one:

4.1 Definition

“Professional ethics involve professionals gaining a deeper understanding of duties owed to the public, each other and themselves in the practice of their profession.”

“Ethics involves doing the right thing and the degree of integrity and trustworthiness.”

“Ethics should be considered to develop society and not just the individuals involved”

These first attempts at a definition are along familiar lines. The aspirational language used is recognisable as standard ethical definitions. One definition which exposed the limitations of the standard approach to ethics was the comment that:

“Ethics and the question of an ethical code conjure up images of fair trade coffee and perhaps fair trade tee-shirts which are sold on a very limited scale and adopted by only a few retail outlets.”

This is nearer to another common perception of ethics, that they are a token effort at making a difference and though well intentioned are of little practical effect. This comment provides a useful touchstone for rest of the discussion and amounts to a contemporary equivalency of how ethics are viewed.

4.2 Context and Views on the Current State of the Industry

The students gave their thoughts on the current ethical climate within the construction industry. The drivers for improving ethical standards were well understood, evidenced by such statements as:

“Ethics has become a hot topic in recent years as professional bodies like RICS and CIOB endorse ethics as professional behaviour in an attempt to eradicate the preconception of the building industry being ruthless. This image has been built up after years of unethical practice.”

“The construction industry has not had a very enlightened past when concerning ethics. Be this building practices, waste management fraud, lack of concern for neighbours and bribery, the industry has had a long slog in the past 15 years to change its image.”

Another commentator did not see the problem of unethical behaviour being limited to construction:

“The construction industry is shrouded in corruption. The perception of fat cats in suits accepting bribes and backhanders, agreeing between themselves whether or not to tender is a view held by most lay people, and indeed the film industry. The idea of an honest businessman is an exception rather than the rule and the construction industry is tarred with the same brush.”

One comment relayed a personal experience that had clearly left an impression:

“I have had firsthand experience of a scaffolding subcontractor coming into my office with brown paper bags for meetings he was going to later in the day. When questioned he said simply it was the only way to win work with local firms.”

Another comment drew a made a link between “unregulated” industries (i.e. those without a comprehensive ethical code). The suggestion is that unregulated industries meet similarly unsavoury ends.

“Examples of industries which are not regulated include the financial services and the construction industry. Recently we have seen the consequences –the national bail out of a number of trusted high street banks and the OFT investigation.”

4.3 Recent Ethical Improvements and Opportunities

Some comments picked up on positive signs in the area of ethical improvement and sought to borrow ideas from other industries.

“The Office of Fair Trading have made progress in that it has banished any “window dressing” and now we move on to consider the commitment to compliance.”

“Contracts such as PPC2000 NEC3 and JCT CE have helped to bridge the gap from the highly complicated legal behemoth that is the JCT SFC towards a more ethical agreement.”

“Promotion and regulation of sustainable and environmental aspects of construction activity are examples of what is achievable through a new ethical approach.”

“The Construction Industry Council has its Independent Appeals Tribunal which is also thought of as an Ethics courtroom.”

“A single ethical code has worked well for the General Medical Council covering the various medical professions using the code to bring together the fragmented construction industry would benefit the image and integrity of those professionals.”

“Hewlett-Packard, Dell and IBM got together and introduced an Electronics Industry Ethical Code which they would do business under – they are trying to move to a more ethical business environment.”

“The industry should follow the example of the Canadian Society of Engineers in Ontario which holds public hearings and publishes its findings. This will enhance the credibility and ensure ethical behaviour is understood by all.”

“The SEC would be a good move for the industry giving a uniform approach and enabling the industry to set minimum standards for ethics.”

“Bring in standardisation and it may change attitudes.”

“A new charter of ethics would boost the prestige of the industry and serve as a recruitment marker ensuring the best graduates and individuals entered the industry and drove up standards in the process.”

4.4 The Obstacles to Further Improvement

Many of the views expressed here are again familiar from the literature and are common takes on the problems facing those who would create a more ethical industry.

“Even though almost all the bigger companies nowadays have their own ethical codes, still there is a huge amount of unethical behaviour in the industry.”

“The tender process on large scale jobs lends itself to cover pricing. Firms do not want to appear work shy; rather they want to stay on “preferred tenders” lists helping them to secure future work.”

“There is also the problem of companies merely paying lip service. Our firm includes mutual dependency and openness in its core values but this does not affect the way they work at all.”

“The closer to the front line of the construction project the more likely the ethical standards of the person are going to be lower.”

“Individual beliefs are very often smothered by the employing company’s views which can force the employee to act as the company requires.”

“The answer is simple - money. Some parties are taking the risk to gain the profit from unethical behaviour.”

“There is a fine line between what some people perceived to be unethical and what others in the industry see as “fair game”.”

“When it comes to the business world personal ethics seem to be pushed aside in favour of profits and financial gain.”

“Another difficulty is the international application – bribes are considered common practice in some areas.”

“Each profession has its own adjudication process normally held privately using self administered rules. In addition, not all professionals are subject to registration or licensing. It is more difficult to enforce ethical standards on generic “engineers” or “surveyors” where there is no tangible threat to remove their right to operate.”

“Unless construction industry professionals are required to carry a licence, like doctors, they can’t be stopped from acting unethically.”

Some new angles on the obstacles to promoting an ethical code did arise in the form of a criticism of specific professional bodies and an apportionment of blame at the door of the specialist contractors. This latter point is particularly interesting in that it

demonstrates that main contractors and/or clients consider that the lack of competition for certain aspects of the work creates unethical behaviour further down the supply chain. Conversely, the last quote puts the blame for unethical behaviour at the door of the clients who attempt to extract too much from their supply chains in terms of cost savings.

“Another difficulty is the reluctance from professional bodies RICS and CIOB who obtain significant competitive advantage through their own ethical stances and codes.”

“Here lies one of the main obstacles – fragmentation. The lack of a leading body, excluding the government, makes tackling the issue a daunting prospect.”

“Perhaps it is the nature of specialist contractors and the very few number of any one type that is the cause of the dishonesty.”

“Forcing contractors to work under their profit margins does not enforce “fair reward”.”

4.5 Purpose/Limitations of the Code

The students next considered how the SEC would work in practice and the practices that would need to be established for its successful operation and implementation.

“The code should sit above and not seek to replace existing codes such as RICS as these are bespoke to the professionals involved in that sector and have been developed over a number of years. A SEC could not hope to capture all the elements of each niche and should not attempt to do so.”

“The question remains how the code would interact with the law and what would be the consequence of a breach be? The OFT penalty of up to 10% seems to have created a suitably horrified reaction.”

“The main difficulty with ethics in construction is that there are no laws and therefore no punishment to enforce their value or use.”

“Most disappointingly although understandably the Society of Construction Law felt unable to issue and regulate the masses. Here lies one of the main obstacles – fragmentation. The lack of a leading body, excluding the government, makes tackling the issue a daunting prospect.”

“The problem with ethics is mainly down to their standing in English law. Law likes to remain black and white, however ethics is largely grey. The problem lies in providing a level of sanction for this grey area.”

“The problem the industry faces is each professional body may produce their own standards and there is no recognised ethical principles that all should follow and be regulated on.”

“There is some argument that ethical codes do not help much when faced with complicated ethical issues and less woolly advice is required – perhaps a work book would assist here.”

This last point is a further good idea in terms of making the idea of ethics less “fair trade” and more practically useful.

4.6 The Positive Steps Required

The students brought some previously raised indications of positive steps required from the literature as well as some creative new suggestions for improvement:

“The SEC will help but other ways to improve need to be looked at to prevent the increase in unethical actions in order to recoup as much from a contract as possible.”

“If a prosecution system could be established it would definitely reduce the number of infringements.”

“For further improvement one must look to the adoption of more collaborative systems in the construction industry; the respondents in longer term trust based relationships find it easier to sign up to a statement of ethical principles.”

“There is no reason why the code could not be included in consultant’s appointments.”

“The industry could try and include the code in the Standard Form of Building Contract. In 1944 JCT included a provision for fair wages which was not law but which the industry agreed to abide by, a similar thing may happen now.”

“A legal entity is essential to develop case law and guidance.”

“Having a SEC may sound great but until it is enforced and policed it will remain a “buzz phrase” in the industry, Imagine every government project where every construction company wishing to bid for work must sign up to Single Ethical Code and comply with it via the Office of Government Commerce. A bit radical but it’s what is needed. If it was a success then the private sector could copy and enforce/police. The UK could pioneer this radical move.”

“The SEC is a good start but needs to be picked up and followed by the key players in the industry. Key larger companies need to ensure they lead the way.”

“If leadership from top down stamps hard on moral and ethical ambiguity then the whole industry will benefit.”

4.7 The Timetable for Action

Several comments identified the need for urgent action for improvement on the ethical front.

“The wake of the OFT report has left a number questioning what is ethically wrong or right – therefore a perfect opportunity for government to take control in clarifying and then regulating.”

“It will take a generation of estimators in order to fully stamp out this (cover pricing) behaviour.”

“A new regulatory body will be required regulating an entire industry from large companies to single workmen - this is a significant feat.”

“The task of the student generation to revolutionise an industry of entrenched attitudes is a large and slow task.”

4.8 Other Issues/Last Thoughts

Some of the questions and issue here raise some provocative and controversial issues. Several students signed off their work with indictments about the ethical state of the industry. Others with observations and a reminder of what is at stake from allowing unethical behaviour to go on unchecked.

“Why are bid rigging companies still members of the CIOB?”

“Why did the SCL decline to regulate the scheme when it was their brainchild?”

“The Considerate Contractor scheme has tackled outward appearances only.”

“The reality is in my view that acts of immoral behaviour will continue.”

“The SEC cannot mitigate against a human’s instinct for self-preservation i.e. winning work for his company and so I fear a SEC would not improve the morals.”

“The reason ethical improvement seems such a long way off is that corruption and dishonesty are as entrenched in the construction industry as concrete and steel.”

“The answer is yes we should have the SEC. This should be an overarching code inclusive of all those working in the industry after all the actions of a sole trader repairing a boiler could well result in death as much as a negligent design.”

“A single ethical code will bring the rules to the forefront.”

“Is a professional code enough to guide us or simply a limit to excess?”

“Just because it is hard does not mean it should not be done. The potential rewards are too great to ignore.”

The last word should be given to the student drawing the analogy with fair trade coffee and the ethical agenda. He goes on to say *“fair trade construction seems a long way off”*.

5. Conclusion and Recommendations

In my submission there is a good deal to take away from the above exercise in terms of fuel for debate on the further implementation of ethical improvement within the industry. The following list is a distillation of the key ideas presented:

1. Make ethics less woolly and give more practical guidance -a workbook or ethics hotline would improve matters.
2. A government backed pan-industry regulatory body capable of enforcing ethical abuses is required.
3. The OFT investigation has provided further impetus for change.
4. Ethical improvement of the industry is a long term undertaking requiring at least a generation of industry professionals to see it through.
5. If the code is to be adopted, inclusion of the code in standard form contracts and forms of professional appointments would be a positive step.
6. A code on its own does not address the deep rooted causes of ethical abuse which include fragmentation, lowest priced tendering and the claims culture.
7. Ethical improvement is a side benefit of longer term collaborative arrangements within the industry such as partnering.
8. Other industries and countries have important lessons for the United Kingdom, including the electronics industry and professional bodies publishing their ethical deliberations and hearings.
9. The benefits for the industry of ethical improvement include presenting a cleaner image to the world and the recruitment of more talented and dedicated individuals.

6. Acknowledgments

I would like to thank the students whose work I have reproduced here and to wish them well with their career development.

7. References

- Bingham, A (2008), Cover Pricing? Baloney! *Building Magazine* Issue 16
- Cohen, S., Grace, D., (1998), *Business ethics: Australian problems and cases*, Oxford University Press, Australia
- Ferguson, WC (1994) Building a solid ethical foundation in business, *The Executive Speaker*, 9(1), Aug-Sep, 33-9
- Egan, J. (1998) *Rethinking Construction*. London, Department of the Environments Transport and Regions. (<http://www.rethinkingconstruction.org/>).
- Greenhalgh, B (1997) *Practice Management for Land, Construction and Property Professionals*, Chapman and Hall, New York

Latham, Sir M. (1994) *Constructing the Team: Joint Review of Procurement and Contractual Arrangements in the United Kingdom Construction Industry*, HMSO, London

Lavers, A (2008) Ethics in Construction Law - a Review of the Position in Eight European Countries, *Society of Construction Law*, paper no. 144 April 2008 available at **www.scl.org.uk**

Lencioni, P.M (2002) Make your Values Mean Something *Harvard Business Review*, 80(7), 113-17

Mason, J (2009) Ethics in the construction industry: the prospects for a single professional code *International Journal of Law in Built Environment* vol 1(3) 194-204

Thornton, A (2004) *Ethics and Construction Law: Where to Start?* Lecture given to Society of Construction Law available at **www.scl.org.uk**

Uff, J (2003) *Duties at the Legal Fringe: Ethics in Construction Law* Fourth Michael Brown Foundation lecture available at www.scl.org.uk

Vee, C. and Skitmore, R.M. (2003) Professional Ethics in the Construction Industry, *Engineering Construction and Architectural Management* 10(2) pages 117-127

Wolstenhome, A (2009) *Never Waste a Good Crisis* Constructing Excellence Report available at: **<http://www.constructingexcellence.org.uk>**

Zhang, X, Austin, S, Glass, J and Mills, G (2008) Towards Collective Organisational Values: a Case Study in UK Construction *Construction Management and Economics* 26, 1009-1028

Word count 5,041

Constructing a Professional Credentialing System for the Construction Industry

Jenny Casmus, Melissa Daigneault

Department of Construction Science
Texas A&M University, College Station, Texas
United States of America

Email: jennycasmus@tamu.edu; mdaigneault@archmail.tamu.edu

Abstract:

The evolution of the construction industry in the past century has produced projects of ever increasing complexity that demand higher degrees of professionalism from all parties involved. As the industry has evolved, so too have the participants on a project team, perhaps none more so than the contractors. Contractors, who at one time did little more than actually perform construction, are now coordinating a plethora of subcontractors and material suppliers, preparing and coordinating shop drawings, value-engineering, and coordinating, if not, leading green building efforts. Additionally, with the increased use of electronic documentation and even BIM, contractors are routinely involved in projects from the earliest phases actively participating in project definition and design. Yet, these same contractors are the only members of the project team who are not generally considered *professionals*, either by the public or the law. Which begs the question, would the construction industry benefit from a unified professional credentialing process, similar to those typically available to architects and engineers? This paper describes the current state of the construction profession as it relates to certification and licensing requirements in the United States; while touching on related legal and public relations implications. Next, it will discuss some of the different credentialing options available, and their potential effects on the profession both legally and in terms of public perception. The paper concludes with the authors' recommendation that the construction industry would benefit from active efforts to increase certification of individuals. The authors find that certification offers the greatest benefit to the industry, without the economic and bureaucratic costs, as well as potential restraints, that other credentialing options may impose.

Key words:

Construction Industry, Professionalism, Credentials, Licensing, Certification

Introduction

Over the past century, the construction industry has grown dramatically, with regard to the scale and complexity of projects. With this growth, each member of the project team has adapted to meet the increasing need for professionalism and expertise. Although contractors currently lack the professional credentials of their design team partners, their own transformation has been no less dramatic. For example, it is contractors who manage the coordinated shop drawing process

between various specialty trade subcontractors- long before designers are asked to approve them. Additionally, it is the installation savvy contracting community performing detailed constructability analyses that are the linchpins around which the value engineering process revolves. To support leading edge green building efforts, progressive builders today employ multiple LEED accredited professionals. Contractors are also taking more design responsibility and relying on electronic documentation, such as BIM, more frequently than ever (Beck, 2001). However, contractors remain the only members of the project team not generally considered “professionals” by the public or the law.

The Oxford English Dictionary (2010) defines a professional as “as person having impressive competence in a particular activity.” This definition, though broad, can be further refined by Jim Mason (2009) of the University of the West of England as he states that “the core of professionalism has been described as the possession and autonomous control of a body of specialised knowledge.” This definition applies to constructors in that, as stated previously, they are continuously performing a myriad of complex and integrated activities. Constructors must be proficient in many varying technologies and processes in order to achieve success on any project, and they must control the implementation of construction methods in order to ensure that all aspects of a project are constructed safely, efficiently, successfully, and in accordance with all appropriate building codes. In spite of the expertise and specialised knowledge held by constructors, as well as current licensing and certification efforts, the general perception is that contractors are not *professionals*, and they are not treated by the law as *professionals*. This might change, however, with the implementation and support of a formal credentialing system for construction professionals. Three credentialing options available to the construction industry are registration, certification, and licensing. In this study, each of these options will be analyzed by the authors and recommendations are made as to which could potentially bring the most benefit to builders everywhere. While this study is specific to credentialing options available in the United States, it is important to note that constructors around the world deserve to be recognized as the professionals they are.

Current State of the Construction Profession

Currently, each governing body has its own requirements for contractors. Some regions only require that constructors obey all safety and environmental laws, as well as behaving in a legal manner in keeping with federal business law. Some regions require that all constructors register with the governing body in order to perform work of a specified minimum monetary value (California Solar Energy Industries Association, 2009). Other regions require that after earning a specified amount of monetary revenue per year, contractors become licensed in order to be eligible to work within that area (Contractor’s License Reference Cite, 2010). Almost all regions require that construction fields with the highest potential of affecting public safety become licensed; these fields include fire protection sprinkler systems, electrical systems, mechanical systems, plumbing, welding, and equipment operation. In addition to required licensing, professional organizations offer optional certification programs and endorsements to recognize and ensure the overall competency of constructors (Construction Certification Commission, 2003).

Licensing

As previously stated, many of the fifty states require formal licensing of contractors working in specialized fields that have the highest potential of directly affecting public safety (Contractor's License Reference Cite, 2010). This requirement is due to the fact that constructors performing such work must possess the necessary knowledge and expertise in order to ensure the safety of the public; allowing under-qualified constructors to perform dangerous, possibly life affecting work could potentially put a community in the threat of serious harm.

In order to obtain a license for a specific field in construction, such as plumbing or welding, constructors must pass an examination administered by the governing body after attending required training for these processes (Contractor's License Reference Cite, 2010). Furthermore, they must log a certain number of hours working directly under a constructor already licensed in the aforementioned field (Contractor's License Reference Cite, 2010).

In some regions, even contractors and builders not specializing in a field of construction mentioned above are required to obtain licenses (Contractors State License Board, 2010). Generally, these contractors are required to obtain licenses once they begin to generate a specific amount of monetary revenue per year from their construction work. In order to obtain these general licenses, constructors are usually required to pass an examination covering a number of topics, including general building practices, safety, and local building codes (California Solar Energy Industries Association, 2009). In addition, licensing boards usually require that a contractor have a specified number of years experience working in construction in order to qualify for the licensing exam (Contractor's License Reference Cite, 2010).

Certification

Either as a supplement to, or replacement for, required licensing, professional certification is offered worldwide. Certification is an optional process meant to act as an assurance of competence and ethical standards for individuals (American Institute of Constructors, 2010). Along with many professional organizations, the Royal Institution of Chartered Surveyors (RICS) (Royal Institute of Chartered Surveyors, 2010) and the American Institute of Constructors (AIC) (American Institute of Constructors, 2010) offer professional certification for participants in the construction industry.

Royal Institution of Chartered Surveyors

The Royal Institution of Chartered Surveyors administers the Assessment of Professional Competence (APC) which is "the practical training and experience which... leads to RICS membership" (Royal Institution of Chartered Surveyors, 2010). Any construction professional has the opportunity to become certified with RICS, as long as he or she follows one of the many routes to membership available. As a prerequisite to most available routes, constructors must have earned a degree in a construction related higher education degree program, preferably through a degree program that is RICS accredited. Having earned a degree, constructors must then gain experience working in the construction industry or attend specified APC structured training. After a specified number of years' experience, constructors submit their final assessment materials, which consist of a myriad of qualifying items, such as a log book, a report, professional progress statements, learning records, and achievement records. Finally, constructors must undergo an interview with the membership board (Royal Institution of

Chartered Surveyors, 2009). Once all qualifications have been met and criteria are approved, a constructor is awarded RICS membership.

Once awarded membership, constructors are eligible to receive all the benefits available to them through RICS, which include professional training, continued career development, a competitive edge, and the status and recognition of being a certified constructor with RICS qualification.

American Institute of Constructors

The American Institute of Constructors (AIC) offers certification of constructors at two levels, the Accredited Constructor (AC) and the Certified Professional Constructor (CPC). According to the AIC, “the certification process is peer developed with the goal of setting high standards for the skills, knowledge, education, and conduct of the certified constructor” (Construction Certification Commission, 2003).

To obtain AC certification, anyone graduating in an accredited degree program in the field of construction, as well as those with experience in the field, may qualify to take the AC exam covering construction materials and methods, project management requirements and processes, and safety. To obtain CPC credentials, a constructor must first qualify for the CPC examination by having worked in the construction industry for a specified number of years that exceeds those requisite for the AC exam. The applicant must then pass the CPC exam, which tests his or her in depth competency in construction materials and methods, project management, and safety (Construction Certification Commission, 2003).

Legal Ramifications and Responsibilities

Liability

When a doctor performs negligently, thus causing injury to a patient, there can be severe consequences. The injured party could sue the physician for malpractice, the physician could be suspended from the hospital or medical centre he or she is working in or, in some severe cases, his or her license to practice medicine may be revoked either temporarily or permanently. Regardless of the outcome, the legal process is sure to be very costly, and it could cause detrimental and irreparable damage to the professional. Lawyers face similar threats; if they behave in such a way that violates their official code of conduct, or cause further harm to another through their negligence, they could be fined, sued, suspended, or have their license to practice law revoked completely. This information is then made public for all consumers.

Construction professionals also face a great deal of liability and risk for work they perform. Through negligence, compliance issues, and preventable injuries on a jobsite, constructors are frequently faced with severe consequences in the form of steep fines, costly lawsuit settlements, and irrevocable marks on their reputations. However, because many areas do not mandate contractor licensing, the possible revocation of a license is not one of the consequences faced; they will usually still be legally able to continue performing construction work. Furthermore, there is often no public record of negligent work.

On the flip side, because doctors, lawyers, engineers and architects are considered “professionals” in the eyes of the law, in order to be sued, a plaintiff must provide an expert

witness who can testify that the professional violated the requisite “standard of care,” which is a greater showing than the simple negligence required for non-“professionals.” In other words, for recognized “professions,” the law requires that other “professionals” be involved in the law suit to instruct the “court” and any jurors. In fact, in some jurisdictions before a plaintiff can even file a suit against a “professional” (like an architect) they must obtain a “certificate of merit.” (ARIZ. REV. STAT. sections 12.2601 and 12-26021 (2004), CAL. CIV. PROC. CODE section 411.35 (Deering 2004), COLO. REV. STAT. section 13-20-602(1)(a) (2003), GA. Code ANN. section 9-11-9.1(a)(2002), MD. Code ANN., CTS. & JUD. PROC. section 3-2C-02(2003), MINN. STAT. section 544.42 subds. 1 and 2 (2003), NEV. REV. STAT. ANN. 40.6884 (LEXIS 2004), N.J. STAT. ANN. sections 2A:53A-26 to 2A:53A-29 (West 2004), OR. REV. STAT. ANN. section 31.300 (2003), PA. R. Civ. P. Nos. 1042.1 to 1042.8 et seq. (2003), TEX. CIV. PRA. & REM. CODE ANN. SECTIONS 150.001-.002 (Vernon 2004)).

Codes of Conduct

Many industries manifest their professionalism “in the promotion and enforcement of an ethical code” (Mason, 2009), and the construction industry is no exception. Each individual firm in the construction industry creates its own code of conduct and ethical standards which are usually meant to supplement those prepared by professional organizations. All professional organizations, such as the American Institute of Constructors (AIC), the Associated Builders and Contractors (ABC), the Associated General Contractors (AGC), and the Mechanical Contractors Association (MCA), just to name a few, prepare formal codes of conduct that are expected to be adhered to by all members. These codes are considered to be a set of guidelines. Neglecting these guidelines can result in many informal consequences, such as injury to a constructor’s reputation and a decline in future opportunities with clients and colleagues; however, there are very few formal consequences for such dissention. In formally licensed professions, however, adherence to codes of conduct is strictly mandatory and could result in the informal consequences previously listed, as well as possible revocation of license, suspension from organizations and notification¹ to the public.

Public Perception

As construction projects have become more complicated, so too have the industries and professionals working on them. “In the medieval period, the Master Builder headed a construction team and served in two capacities at once, both architect and contractor” (Ngowi, 2000). As a result of the industrial revolution, the fields of architecture, engineering, and construction split, creating three distinct fields with different specialties. Since that time, each entity of the project team has been considered as a professional with the exception of the constructor. Contractors are frequently perceived as coarse, uneducated, and unethical, and today’s media does little to alleviate this perception (Pratt, 2002). However, those who believe this stereotype do not realize the true level of expertise and ethical values that are necessary to work in the construction industry, especially today. As Dave Hall, the Executive Vice President of Geupel DeMars Hagerman said, “with construction projects and project relationships becoming more complex, and with the project delivery processes often being driven by the

¹ Notification to the public of violations of codes of conduct by “professionals,” usually occurs through publication in the official periodical of the profession, maintained by the quasi-governmental entity that oversees the profession. However, this information is also kept and archived by the professional association or governing board, and can be accessed and searched by the public at large.

construction entity, there is a need to recognize the constructor for what he or she is... a truly professional member of the project delivery team” (Construction Certification Commission, 2003).

Credentialing Options

In an effort to formally endorse the expertise of construction professionals, three basic options of formal credentialing are available to consider: registration, certification, and licensing.

Registration

One credentialing option that may be beneficial to pursue is that of registration. Registration would require that once an individual or entity working in the construction industry begins to generate a set amount of revenue per year, they would be required to register with their state or local government as a professional constructor (Barron, 1966). Requisite with registration would be a mandatory examination determining the overall competency of an individual in construction practices. The examination would be administered by the ruling body, but assembled through collaboration among experts in the construction industry (Barron, 1966). No minimum score on this examination would be necessary, but each registrant would be required to make his or her registration certificate readily available to all potential customers. This certificate, issued by the registration board, would include information regarding the constructor’s qualifications as determined by the board, such as his or her education, years of experience, and examination scores.

The system of mandatory registration would create minimal barriers of entry into the construction industry, as a set of specific qualifications would not be required to legally be able to perform construction work. The only barrier caused by registration would be the cost of a registration fee, which should not be set too high to preclude individuals from working.

According to J. F. Barron (1966), a former Professor of Economics at the University of California in Los Angeles, requiring all professionals to register with the state or local government would shift the liability of contractor selection from the governing body to the client. It becomes each buyer’s responsibility to base the choice between prospective contractors on the qualifications and certifications available to the client (Barron, 1966). If a contractor has no relevant education, minimal experience, and performs poorly on the examination, a client should note these shortcomings, which would be displayed readily on the registration certificate, and look for a new contractor. The governing body registering the constructor could not be held liable for allowing such an unqualified business to perform work, because there are no minimum requirements, thus, it becomes the duty of the owner, client, or consumer to ensure that an entity is qualified.

Under registration, the public perception of the construction industry may improve slightly, but would most likely remain stagnant. With no barriers for entry, formally accredited training and education would not be necessary; however, a required examination and mandatory posting of qualifications might act as a step towards an improvement in public perception. By giving clients the opportunity to view and analyze education, experience, and examination qualifications, it may become more apparent to the public that the contractors building our skyscrapers, bridges,

hospitals, and stadiums have the expertise necessary to complete the complex and sophisticated work of construction.

Certification

The certification of construction professionals would be a voluntary option of credentialing individuals, implemented by professional organizations. In order to become certified, constructors would be required to pass a field specific examination compiled and administered by experienced construction professionals within the professional organizations. Additionally, constructors would be required to meet a certain set of qualifications before attempting to pass the exam.

As mentioned before, when discussing RICS membership and certification, to be qualified to take the certification exam, constructors would be required to prove that they are trained in the methods and processes of the industry through their education and experience (Royal Institution of Chartered Surveyors, 2009). Students enrolled in a construction related accredited degree program, such as construction management, would be eligible to take the exam for a lower level of construction professional certification during their last term before graduation. After graduating and working in the construction industry for a set number of years, as determined by the certification board, a lower level certified professional may qualify for the examination to become a fully-fledged certified professional constructor. Constructors choosing to work up through the construction industry trades, rather than pursue higher education, would be able to apply to take the certification examination after a set number of years as specified by the certification board, at which time he or she would be required to provide structured training certificates or acknowledgements of their expertise in their field of work.

After proving his or her qualifications based on education and experience, a constructor seeking certification would then be required to take the certification examination which would be compiled and administered by experienced construction experts within the professional organization. This exam, much like the AC exam administered by the American Institute of Constructors (2010), would cover a myriad of topics relating to the construction industry including materials and methods, surveying, specifications and plan reading, construction law, communication, green building, emerging technology, structural stability, soil conditions, mechanical systems, electrical systems, safety, equipment, estimating, scheduling, and project management. Constructors that pass the exam would receive professional constructor certification.

As a professionally certified constructor, individuals have the benefit of being recognized throughout the construction industry and beyond as a professionally qualified constructor that has “achieved high standards of competence and ethics” (Royal Institution of Chartered Surveyors, 2010). This distinction creates a competitive edge for the individual with both potential employers and clients. In addition, certified members of professional organizations are offered the opportunity to participate in advanced training for current and emerging trends in technology in the construction industry, thus enabling certified constructors to retain their competitive edge throughout their careers.

Because certification of constructors is optional, it poses no barriers of entry to the construction industry; constructors are legally be able to work in their field of choice without becoming certified by professional organizations. Furthermore, clients are free to choose among a myriad of contractors of varying degrees of expertise, thus not precluding anyone from free competition; however, clients who want a guarantee of quality and expertise can limit their selections to professionally certified constructors (Barron, 1966). Additionally, because certified members of the professional organization would be obligated to adhere to the official codes of conduct set forth by the organization, clients could feel confident of their choice in a constructor that adheres to high ethical standards and values. These codes of conduct are meant to protect clients, competitors, and the general public from unethical or unprofessional practices that could possibly result in lower quality work, unsafe working environments, injured reputations, or unfair competition situations. Straying from these codes of conduct could result in a constructor's expulsion from the organization as well as the revocation of his or her certified status; however, there are no legal repercussions, no "public" notice, and nothing that would prevent the offender from continuing to offer those same services in the same locale or in any other.

Through a system of certification, the public perception of the construction industry could improve dramatically, but only if certification is widely publicized, widely accepted, and high standards are maintained by the certifying organizations. Because a constructor must demonstrate his or her competency in the construction industry, successfully completing an examination designed to challenge those seeking certification would prove the individual's ability to handle the complexities and adapt to the technology of an ever growing field. For this reason, those outside the construction industry may be able to see that constructors are fully trained experts in their field who are wholly qualified to manage and build a project while ensuring the safety of the public and the quality of the structure. Certification could "enhance the professionalism of the constructor to the public, client, and employer" (Construction Certification Commission, 2003). However, because certification would provide no barriers of entry into the construction industry, there would continue to be dissent and doubt from the general populace regarding the professionalism of those constructors who, while surely qualified and capable, have not taken the opportunity to prove their expertise through professional qualification.

As stated previously, the Royal Institution of Chartered Surveyors and the American Institute of Constructors currently offer professional certification, but in order to improve the public perception of the construction industry to its deserved level, more employers, colleagues, and clients need to encourage constructors to become professionally certified.

Licensing

Licensing of construction professionals would be a mandatory process of credentialing constructors. Licensing would require that all construction professionals, namely project managers and superintendents, undergo specific licensing processes in order to be able to legally work in the construction industry. Professionals could qualify for a license through education, experience, and examinations.

As there currently are a number of accredited construction management programs available at many universities, professional licensing would require that in order to be licensed, a constructor must first graduate from a construction management degree program that is accredited by the governing body or an approved third party licensing board. Additionally, these degree programs would be required to periodically renew their accreditation status with the appropriate institution. At the completion of the 4-year degree program, each constructor seeking a license would then be required to pass an examination administered by the licensing board. This examination would be similar to that used for certification, although it would be compiled and administered either by the licensing board or an approved third party made up of experts in the construction industry.

Upon graduation, a constructor would then go to work for a firm in the construction industry to gain experience and training. During this apprenticeship period, each constructor would work directly with a licensed professional in order to gain the expertise necessary to become licensed. After a specific number of years, as determined by the licensing board, a constructor could then apply to take the final qualification exam. With his or her application, each constructor would be required to include progress reports and recommendations from his or her company of apprenticeship. The final examination would cover the same areas as in the preliminary examination, but at a much greater depth, requiring a deeper understanding of the construction industry. After passing this final examination, a constructor would be issued a license to practice construction management in his or her state. Furthermore, constructors would be required to renew their licenses periodically throughout their careers.

If approved, an alternate licensing process may be available to those who have experience working in construction, but do not obtain a degree from an accredited construction program. These constructors would be required to provide evidence of their experience and qualifications, provide recommendations from superiors, and attend structured training programs. They would also be required to pass the same examinations as constructors on the academic path.

The credentialing option of licensing would restrict entry into the construction industry drastically in an effort to protect the interest of the public. By only allowing constructors that have formally proven themselves to be qualified construction professionals to work in construction, clients would be more confident that their projects would be built in accordance with plans and specifications, and to the highest standard of quality and safety. There would be, however, many constructors that would be fully qualified and experienced enough to ensure the highest standards of quality, safety, and ethics, but would be unable to complete the licensing process for various reasons. Furthermore, formal licensing would pose specific requirements on construction professionals. Licensed individuals would be required to adhere to a specific code of conduct, enforced by the governing body and licensing board, that set the standards of ethics, quality, and professionalism that would need to be maintained at all times, for fear of rebuke from the licensing board, revocation of license, or involvement in malpractice suits. Many professionals do not agree with mandated codes of ethics, and in a 2004 survey of architects, engineers, construction managers, general contractors, and subcontractors in America, only 30 per cent agreed with proposed mandatory ethical regulations (Mason, 2009).

Formal licensing of construction professionals would increase the public perception of the construction industry drastically. By adopting many of the same criteria and requirements that

architects and engineers themselves must adhere to in becoming licensed, the perceived gap of professionalism between contractors and the rest of the project team could close. By knowing that all builders and constructors receive formal education and training, gain experience by working in their field of choice, and are tested on overall competency and ethics, the general public, clients, and other professionals would come to realize what has been true for a very long time; those working in the construction industry are truly professionals.

Recommendation

The construction profession would benefit from the active efforts to increase the formal certification of individuals. Because certification involves limited or no government involvement, and is a voluntary process, the costs and restraints on the industry would be the lowest of the options discussed in this paper; therefore, certification would offer the greatest benefit to the industry at the present time.

Certification would benefit the construction industry in that it would act as an assurance that certified constructors are fully capable of successfully completing complex projects. Individuals certified as professional constructors would also have a competitive edge over others, and being a certified member of a professional organization would give individuals the opportunity to further their education in emerging technology through available training. However, certification can only be an effective step towards increasing the professional reputation of the industry if it is widely accepted by the profession and the public at large.

In addition, certification would not involve as many costs to the construction industry as mandatory licensing would. First, licensing would create a higher monetary cost, in that more administrative personnel would be necessary for the licensing process. Second, licensing would create more barriers of entry to the construction industry, in that a specific path of education and experience must be followed. Many of the most talented constructors did not follow a path of higher education, and not all constructors would have the opportunity to attend the structured training required for licensing. Third, with such barriers, free competition between companies may be hindered, giving clients fewer options of whom to work with. Finally, licensing would create a higher liability and greater consequences for construction companies, thus increasing the overall cost and fees of construction with no guarantee of increased quality within the profession.

Conclusion

While the construction industry has evolved over the last century, so too have the members of the project team, as projects have become increasingly more complex and integrated. As a result, members of the project team have seen higher levels of professionalism from all involved, including that of the contractor. While current licensing and certification measures have attempted to improve the public perception of contractors, and the quality of the industry, to that of the true professionals they are, only widely acceptance, use, and endorsement of a credentialing system for constructors will be able to achieve such a lofty goal. Three credentialing processes are available for the construction industry: registration, certification, and licensing. Of these three, increased efforts towards certification of constructors would add the greatest benefit to the construction industry, by ensuring quality and competency, without the high costs of licensing, which include the monetary cost and barriers of entry and advancement in construction.

References

- American Institute of Constructors, 2010. *Certification*. Virginia: American Institute of Constructors. Available from <http://www.professionalconstructor.org/professionalconstructor/professionalconstructor/aboutccc/default.aspx> [Accessed 24 June 2010]
- Barron, J.F., 1966. Business and Professional Licensing. California, a Representative Example. *Stanford Law Review*, 18 (4), 640-665. Available from: <http://www.jstor.org> [Accessed 23 March 2010].
- Beck, P., 2001. The AEC Dilemma. *Leadership and Management in Engineering*, 1 (2), 31-36.
- California Solar Energy Industries Association, 2009. *Contractors' License*. Rio Vista, California: California Solar Energy Industries Association. Available from: <http://calseia.org/contractors-license.html> [Accessed 23 March 2010].
- Construction Certification Commission, 2003. *Certification Recognition and Benefits*. Alexandria, Virginia: American Institute of Constructors. Available from: <http://www.constructorcertification.org/crb/index.htm> [Accessed 23 March 2010].
- Construction Certification Commission, 2003. *Benefits of Certification*. Alexandria, Virginia: American Institute of Constructors. Available from: http://www.constructorcertification.org/crb/certification_benefits.htm [Accessed 23 March 2010].
- Contractor's License Reference Cite, 2010. *Find a State's Licensing Information*. United States: Contractor's License Reference Cite. Available from: <http://www.contractors-license.org> [Accessed 24 June 2010].
- Contractors State License Board, 2010. *Contractors State License Board*. California: Department of Consumer Affairs. Available from <http://www.cslb.ca.gov> [Accessed 24 June 2010].
- Mason, J., 2009. Ethics in the construction industry: the prospects for a single professional code. *International Journal of Law in the Built Environment*, 1 (3), 194-204. Available from: www.emeraldinsight.com [Accessed 22 February 2010].
- Ngowi, A.B., 2000. Construction Procurement Based on Concurrent Engineering Principles. *Logistics Information Management*, 13 (6) 361-369. Available from: <http://www.emeraldinsight.com> [Accessed 21 June 2010].
- Pratt, M. K., 2002. Certain Stereotypes Die Hard in the Construction Industry. *Boston Business Journal*, 6 September.
- Royal Institution of Chartered Surveyors, 2010. *Assessment of Professional Competence*. London: Royals Institution of Chartered Surveyors. Available from: <http://www.rics.org/apc> [Accessed 31 March 2010].
- Royal Institution of Chartered Surveyors, 2009. *RICS Routes to Membership*. London: Education and Qualification Standards. Available from: <http://www.rics.org/site/scripts/documents.aspx?categoryID=98> [Accessed 31 March 2010].
- Royal Institution of Chartered Surveyors, 2010. *Join the World of RICS: Why RICS?* London:

Royal Institution of Chartered Surveyors. Available from:

http://www.rics.org/site/scripts/documents_info.aspx?documentID=358&pageNumber=4

[Accessed 31 March 2010]

Spatial planning to the test of property rights in Switzerland: an innovative land management approach to coordinate spatial planning goals with property rights interests.

Philippe Weber¹

¹Institute of land use policies and human environment, University of Lausanne, quartier Sorge – bâtiment Amphipôle, 1015 Lausanne, <http://www.unil.ch/ipteh>.

Email: Philippe.Weber@unil.ch

Abstract:

In Switzerland, the land management regime is characterized by a liberal attitude towards the institution of property rights, which is guaranteed by the Constitution. Under the present Swiss constitutional arrangement, authorities (municipalities) are required to take into account landowners' interests when implementing their spatial planning policy. In other words, the institution of property rights cannot be restricted easily in order to implement zoning plans and planning projects. This situation causes many problems. One of them is the gap between the way land is really used by the landowners and the way land should be used based on zoning plans. In fact, zoning plans only describe how landowners should use their property. There is no sufficient provision for handling cases where the use is not in accordance with zoning plans. In particular, landowners may not be expropriated for a non-conforming use of the land. This situation often leads to the opening of new building areas in greenfields and urban sprawl, which is in contradiction with the goals set into the Federal Law on Spatial Planning. In order to identify legal strategies of intervention to solve the problem, our paper is structured into three main parts. Firstly, we make a short description of the Swiss land management regime. Then, we focus on an innovative land management approach designed to implement zoning plans in accordance with property rights. Finally, we present a case study that shows the usefulness of the presented land management approach in practice. We develop three main results. Firstly, the land management approach brings a mechanism to involve landowners in planning projects. Coordination principle between spatial planning goals and landowners' interests is the cornerstone of all the process. Secondly, the land use is improved both in terms of space and time. Finally, the institution of property rights is not challenged, since there is no expropriation and the market stays free.

Keywords:

Coordination, land management, property rights, spatial planning, Switzerland.

1 Introduction

Switzerland has seen its population grow more strongly than normal during the last decade, this mainly due to the net immigration located in urban and periurban areas (OFS 2009). In terms of land use, population growth is producing obvious impacts that

it is imperative to manage in order to avoid uncontrolled spatial developments, which have negative effects on the environment especially. In this context, land use planning is a necessary condition for the achievement of an ordered territorial development, which can fulfil the objectives set in spatial public policy (spatial planning, energy, environment, etc.). Regarding the State intervention, the fact is that for most of landowners it is difficult to restrict their rights, namely the constitutional guarantee of ownership so that land use match as much as possible with what have been planned. It is thus common to be faced to owners who do not wish - for various reasons¹ - to develop their land according to what has been planned by municipalities.

The concept of land hoarding, or the withdrawal of a buildable plot of the land markets, allows showing that it exists a gap² between a *potential supply* of buildable land (defined by the planes) and a *real supply* of buildable land available for construction (defined by the land markets). The stake is very important since land hoarding is one of the trigger elements of spatial planning problems to which Switzerland is currently confronted (urban sprawl, poorly localized land supply, land prices) (BCV 2007; Rüegg 2000). Regarding public authorities and the interests they represent, it is imperative to consider appropriate measures for striving against this problem to ensure a rational development of our territories and an equal treatment between landowners and other stakeholders. What are the available tools within the Swiss spatial planning system? What is their efficiency?

We prefer to announce immediately the limits of our paper. We do not pretend to write a comprehensive study of the Swiss system, and therefore we will focus on a specific tool developed and implemented in the canton of Vaud, Switzerland: the Land and Improvement (LI) Syndicate in buildable land. The canton of Vaud, in order to restore the balance between potential and real supply of buildable land, and thus to address the problem of land hoarding, is implementing a land management approach based on the coordination between public interests and private interests. The main hypothesis of this paper, which is guiding our reflexion, is that the coordination between public and private interests is a necessary condition but not sufficient to reduce the gap. As much as we know, the tool used by the canton of Vaud has not yet been assessed, partly or fully. To fill this gap, this paper seeks to describe this tool theoretically, and then assess its contributions and limits through a case study.

2 From land markets to land policies

Comby & Renard define the object of “*land science*”³ as that which “*allows to state either in terms of results or in terms of method, how the lands will be allocated among different possible uses*”⁴ (1996, p.3). Inspired by the definition of economic science, the land, as a specific economic resource, obeys not yet to rules and conditions that depend not only on the law of supply and demand because of its nature (unique, irreproducible, nonproduced) and the functioning of the market linked to its exchange

¹ For example, these reasons may include the retention for speculative purposes or heritage.

² It may be more or less important regarding the regional contexts where this phenomenon is analysed.

³ Author's translation from french: “*science foncière*”.

⁴ Author's translation from french: “*de permettre de fixer, soit en termes de résultats, soit en termes de méthode, comment les terrains seront alloués entre les différentes utilisations possibles*”.

(atomicity, transparency). Indeed, there are many other factors that influence land values such as the rights attached to a plot. Accordingly, land markets are working in an imperfect manner, and should be therefore regulated by mechanisms that the State coordinates in order to avoid drifts that are socially unacceptable, economically inefficient, ecologically dangerous (Comby & Renard 1996; Prélaz-Droux 2008; Rüegg 2000).

These mechanisms, which we may call *land policies*, follow two bodies of objectives, physical and financial. From a physical point of view, it is to promote a rational and coherent territorial development with rules and plans. Our hypothesis is that the market is not enough effective to generate a sustainable territorial development. From a financial point of view, it is to inflect the harmful redistributive consequences that land markets generate spontaneously. For example, the fact that an owner sees his parcel gain value under authorities action is something to be regulated (Comby & Renard 1996; Prélaz-Droux 2008).

The following of our paper focus mainly on the physical body of land policies. Into the next chapter, we present briefly the Swiss land regime. This allows us to do necessary linkages between the land hoarding problem identified in the introduction chapter and the necessity of its regulation regarding to the elements identified just above.

3 Land regulation in Switzerland

During the period 1960-1990, the Swiss land management regime has strongly evolved regarding the Federal constitution and law. After the 90's, the land issue seems to have been taken off the political agenda, as demonstrated by the lack or absence of relevant research and publications in this field (Nahrath 2005). Now, under pressure of demographic constraints, urbanisation processes or still environmental challenges, land management issues resurface with in the Swiss scientific and political agenda in becoming more than ever essential in the context of sustainable development.

In Switzerland, the institution of property rights is an individual fundamental liberty that is guaranteed by the Constitution (art. 26 Cst)⁵. In this context, any restriction on property rights must be based on a legal basis, justified by a public interest, and proportionate (art. 36 Cst). Thus, the Swiss Constitution provides a number of bases that legitimate the action of the State in various field related to natural resources (land, environment, forest, water, noise, etc.). Spatial planning (land use), on which we will

⁵ Art. 26 Swiss federal Constitution state that: "*The right to own property is guaranteed. The compulsory purchase of property and any restriction on ownership that is equivalent to compulsory purchase shall be compensated in full*". The doctrine and jurisprudence define the guarantee of ownership along two axes: the subjective guarantee and the objective guarantee. The first is the one understood in the classical sense (explained in the text) that is the respect of private property right by the state in terms of spatial planning. According to this view, the owner may oppose to a measure of planning when it does not respect its rights. The objective guarantee is this more subtle that it requires the maintenance of the institution of property rights in itself. Two elements are constitutive. The first is quantitative, in the sense that the land market must stay free and it should have enough land in the hands of private landowners. The second is qualitative and implies that property rights keep a sense for the landowner. The concept of objective guarantee is relatively recent within the Swiss law and, as much as we know, no decision has yet been taken on this basis.

work more precisely, has a constitutional article (art. 75 Cst)⁶ that gives to the State the possibility to work towards an economical use of the land and its properly ordered settlement. The introduction of this constitutional basis in 1969 has allowed the creation of the Law on Spatial Planning in 1979 (LAT), which specifies the constitutional objectives and defines a number of tools in order to achieve these objectives. The three main tools of the Swiss system are: conception plans and sectorial plans (Swiss Confederation); the cantonal Master Plan (cantons); the zoning plan (municipalities). The two first are strategic documents, which have not force of law for landowners. They are respectively implemented by the Confederation and the cantons. On the contrary, the zoning plan implemented by municipalities is binding on landowners and authorities. It has force of law once it has been accepted democratically. Zoning is in this sense a restriction to property rights' institution that is a legally accepted practice formalised in the LAT. This arrangement explains why municipalities have such an important autonomy in terms of spatial planning. Regarding the implementation and the contexts, municipalities are facing different problems, but especially the one we have presented in introduction: land hoarding. This problem is resulting of a complex articulation between private interests and public interests. Indeed, as we have presented, the constitutional arrangement requires taking into account the guarantee of ownership by the authorities in the implementation of spatial planning. In fact, private interest outweighs public interest. In other words, although the LAT legitimates the practice of zoning, it requires not the landowner to build his property after the entry into force of the plan. Thus, inaction leads not to a penalty and cannot be considered as non conform behaviour regarding the legislation on spatial planning. For example, the fact to not maximise the potential use of a buildable plot do not constitute a sufficient element for expropriation or compulsory purchase. The landowner is thus in situation of quasi monopole once building rights have been distributed after the entry into force of the zoning plan (Rüegg 2000).

The land hoarding phenomenon, which is in our view constituting a market failure generated by the functioning of land markets, should be regulated so that the spatial planning objectives can be operationalized. The regulation of this problem constitutes a big issue for both public authorities and the owners, which are *a priori* forces going in opposite directions. In this context, the land management tools provided are generally pursuing the balancing of these forces in order to achieve a territorial production of good quality.

4 Land management issues in the canton of Vaud, Switzerland

Traditionally, the planning process is made of successive levels with in which planners aims to draw up zoning project for urban production. Land issues are only solved once planning project has been done. This kind of approach, which is common in Switzerland, creates a lot of difficulties and incoherencies, and fails to reconcile conflicting interests between private and public in order to achieve a quality urban production with respect to the LAT. Moreover, it fails to reduce the gap. Thus, it often leads to the postponement, or even the cancellation, of land development projects.

⁶ Art. 75 Swiss federal Constitution states that: “The Confederation shall lay down principles on spatial planning. These principles shall be binding on the Cantons and serve to ensure the appropriate and economical use of the land and its properly ordered settlement”.

Within the Swiss land management regime, the canton of Vaud has developed a specific approach based on the introduction of new tools within its land management regime, which aims to establish coordination between spatial planning and land management issues in order to avoid such problems (land hoarding, diverging interests, etc.)⁷. Thus, in agreement with the federalism principle, Vaud has edited laws both on spatial planning (Canton de Vaud 1987) and on land improvements (Canton de Vaud 1962). Approach advocated by Vaud is based on a major principle: the coordination between spatial planning (public interests) and land management issues (private interests) (art. 55 LATC; art. 4 LAF). We assume that coordination between public and private interests is a necessary condition, but not sufficient, to rebalance the forces. The principle of coordination is what we wish to focus on, and more particularly the process that it entails. The question is how does the coordination between spatial planning and land management issues takes place.

4.1 Feasibility study

In most of the case, zoning plans can be realized through the signature of conventions between landowners. In fact, in complex cases, when public interest is obvious or/and there is a major conflict of interests, authorities (Cantonal Office for spatial development, municipality) or landowners can launch a *feasibility study* made up of a territorial project outline and a preliminary land improvement study (art. 85a-i LAF). Such a study seeks to assess the land development project feasibility both on spatial planning and land management issues. In general, this approach leads to four main results: a territorial project outline and its implementation principles (i.e. a localised zoning plan); an outline of the collective spaces and equipments that need to be provided; a cost estimate of collective spaces and equipments to +/-25% which establish investments to be done by landowners; a land management approach proposition adapted to the case. According to the feasibility study results, concerned landowners can express their opinion, as to whether they agree or not. Landowners' answers can be threefold. In the first scenario, all landowners are agree with the project. The land management approach is essentially based on the signing of conventions. Or some landowners don't agree with the project although a majority who wants to pursue it. In this case, the majority can create a voluntary *Land Improvement Syndicate* (LI syndicate), which comprises the whole perimeter of the project. Here, all landowners are members of the LI syndicate even if they do not agree, because decisions are taken on a majority basis (1 plot = 1 voice). In the third case, landowners oppose the zoning plan. The authorities (municipality) either accept this decision, or create a mandatory LI syndicate in order to implement the zoning plan of municipal jurisdiction (art. 27 LAF; art. 85k LAF), or launch a limit and easement correction process for the purpose of overriding public interest (art. 93a LAF). Within the case study we especially focus on the voluntary LI syndicate.

4.2 Land Improvement Syndicate in buildable areas

In terms of land management, the LI syndicate is the most developed organisational model. By reading articles 702 and 703 of Swiss Civil Code (CC), the LI syndicate is a public corporation, which is a form of property and works as an association. It is

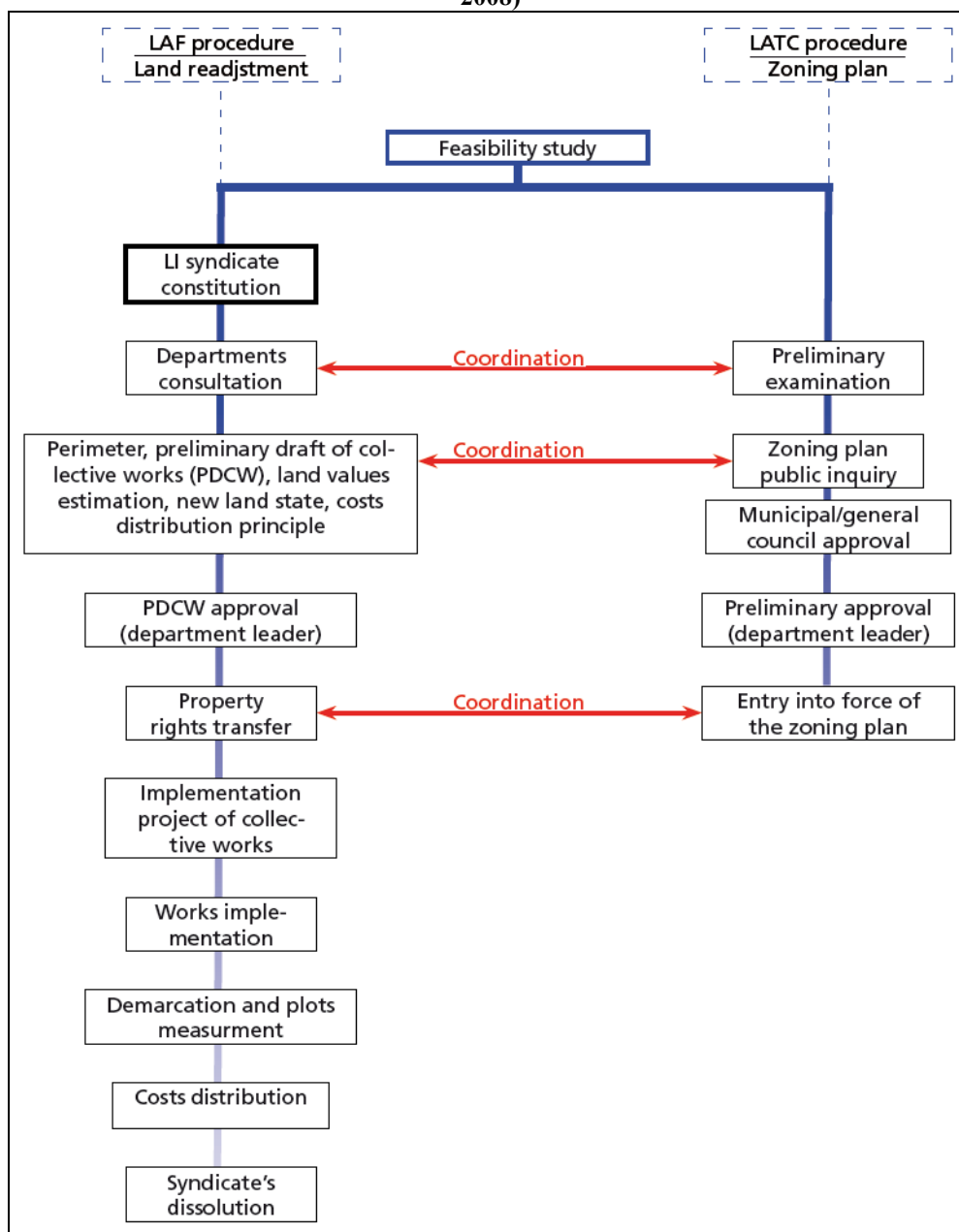
⁷ This chapter is based for its most part on the paper of Denis Leroy, engineer at the Cantonal Office for Spatial Development (SDT) of the Vaud Canton, Switzerland (Leroy 2008).

governed by LAF's dispositions and supervised by the authorities (art. 122 LAF). All landowners affected by the spatial and land development project are members and have the right to vote within the decision process. Decisions are taken on a majority basis, i.e. each owner has one voice. Together, landowners seek to reorganize the land property to allow valuable development of the land, to equip their properties accordingly and to update the building rights according to the chosen development project (Leroy 2008; Prélaz-Droux 2008; Service des améliorations foncières 2003; Service des améliorations foncières 2005). A LI syndicate is divided into five bodies:

- The *general assembly* is the central body, in which all landowners take part (1 landowner = 1 votes).
- The *steering committee* is the executive body, which implements decisions taken by the general assembly. It also has an administrative role.
- The *management commission* is in charge of accounts auditing. It ensures the proper financial management of the steering committee.
- The *classification commission* is a panel of neutral and independent experts who establishes the project and treats complaints in first instance.
- The *technician* provides the realisation and the coordination of the project. It is generally a surveyor.

The lifetime of a LI syndicate is divided into three main steps (see table 1). The first step is dedicated to studying the land property management, costs distribution, changes between old and new land state. On the other hand, the land improvement file and the zoning plan are subjected to a simultaneous public inquiry. The coming into force of the zoning plan and the property rights transfer, simultaneously treated, mark the end of the first phase. The second step is dedicated to the servicing of the land and the building of collective spaces. And finally in the last step cadastre and land registry are updated. Costs are distributed between all members, and the LI syndicate is dissolved (Leroy 2008). We can see that coordination is ensured during the whole process, since the feasibility study until the property rights transfer.

Table 1: Procedures coordination under LI syndicate after the feasibility study (from (Prélaz-Droux 2008))



In financial terms, the whole process is supported by a self-financing system whose content is decided by the general assembly. In fact, all members must pay fees each year proportionally to the size of their plot (art. 43 and 46 LAF). These fees pay for the studies done during the first step (see table 2). Nevertheless, within the second step, LI syndicate must often get bank loans in order to ensure the costly servicing of the land. Bank guarantees are provided through the coming into force of the zoning plan, which ensures the legality, the feasibility and the building of the whole project in a given time. Thus, the entire land value is used to back up the debt (see table 3): agricultural value, capital gain obtained by zoning, capital gain obtained by land improvement (collective spaces and equipments). In other words, the land servicing and the public services are

what need to be financed, while the guarantee covers the whole land value (Leroy 2008; Service des améliorations foncières 2005).

Table 2: LI syndicate implementation's process (from Service des améliorations foncières 2005)

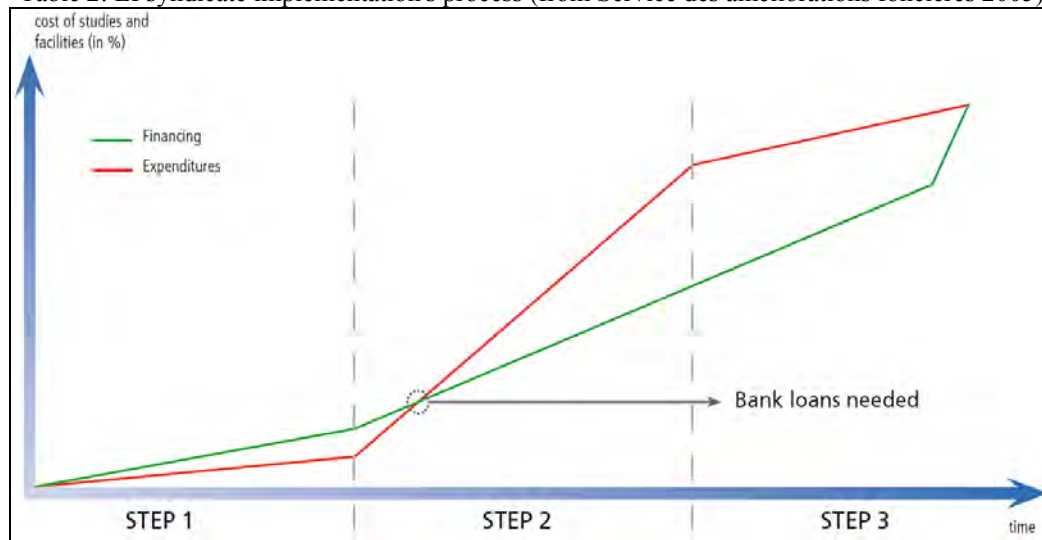
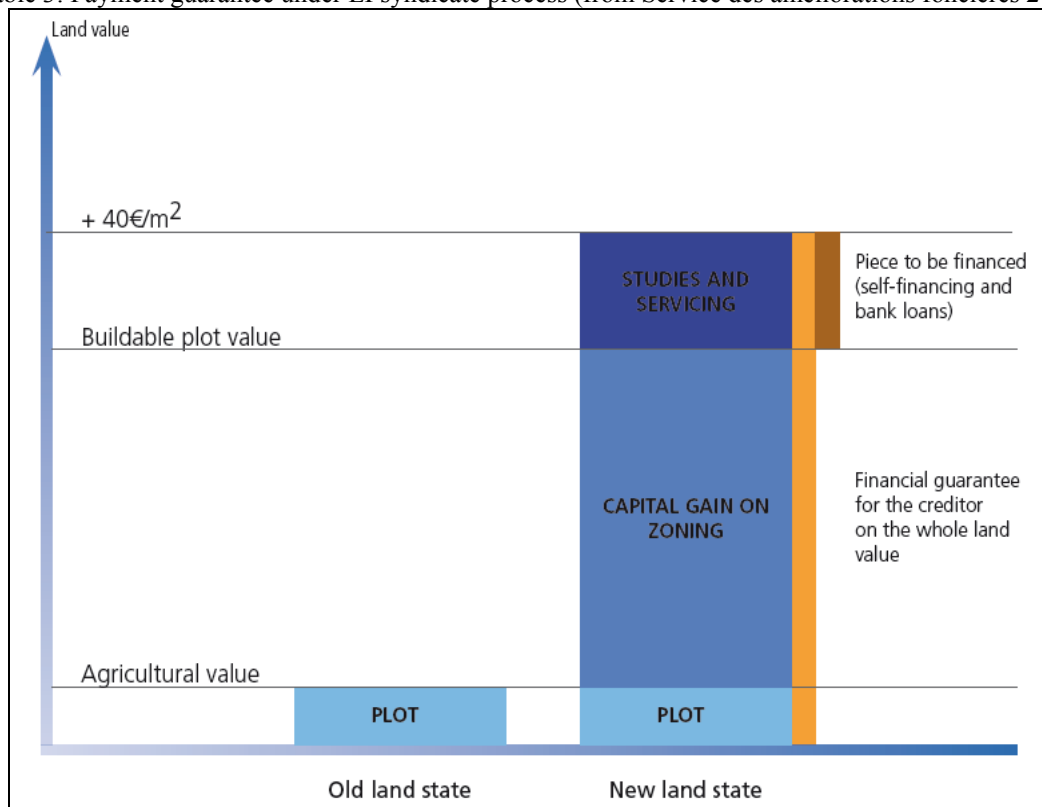


Table 3: Payment guarantee under LI syndicate process (from Service des améliorations foncières 2005)



5 La Boironne case study

In Switzerland, it is common to deal with stalled situations for years. As already said above, the problem comes mainly from owners who hoard their land. Thus,

municipalities facing growth (population) don't dispose the necessary surface at the right place. This leads finally to sprawl which is not desired. In this kind of case, authorities can launch a feasibility study in order to draw up an urban project outline and to propose a land management approach adapted to the case in order to realize the land as it has been zoned before. Our case study addresses explicitly this aspect and allows us to demonstrate the relevance of our starting hypothesis.

5.1 Context and history

Yens is a little rural village of about 1'000 inhabitants. It is located near the cities of Morges and Lausanne, within the Geneva Lake metropolis in Switzerland. Within the village of Yens, the *La Boironne* area has been the subject of many planning studies between 1980 and 2000 (see figure 1). It has firstly been decided to zone this area as a buildable one in 1980. Legally, this area has been managed with a partial zoning plan⁸ since this date, but hasn't been built yet. Different reasons are involved, especially the wish express by the municipality to build a collector road crossing the area, which has delayed the project.



Figure 1: Situation map of Yens and La Boironne area (from www.geoplanet.vd.ch)

In 2000, a new partial zoning plan was drawn up for the area at the initiative of certain landowners. This initiative led the cantonal office in charge of spatial planning to propose the launch of a feasibility study on the area. In fact, from a legal point of view - it has been demonstrated that the plan might have impacts on the state of land and on

⁸ In fact, such a kind of plan allows managing an area in accordance with specific constraints by taking into account local process as well as regional ones.

the property rights and would therefore necessitate a land readjustment - the canton had to launch such a feasibility study (art. 85a LAF and seq.).

The figure 2 shows the area selected and the zoning categories with which the project have to cope. The two first zones are buildable areas. The brown zone describes an area dedicated to village area. The yellow area is dedicated for low-density housing (villas). Then, the area in light-grey is called intermediate zone. Its definition is formulated by article 51 L ATC. Briefly, intermediate zones are in fact agricultural zone in terms of value and can't be built. But, drawing such a zone is useful as it enables to define areas, which will probably be zoned as buildable in the future. In the present case, a part or the whole surface can be zoned in the buildable category only if a real equalization ensures equal treatment between landowners, either by conventions or by land readjustment. Finally, plots and roads already drawn show the new land redistribution between landowners.

Still within the figure 2, we can see that the surface is already divided into plots. They correspond to the new state after the land readjustment. As we can see, the plots size corresponds to the developments, which will be realized in this area. The old state didn't make possible the implementation of the plan due to small vineyard plots, which were too tight to be developed. That was another reason launching a land readjustment.

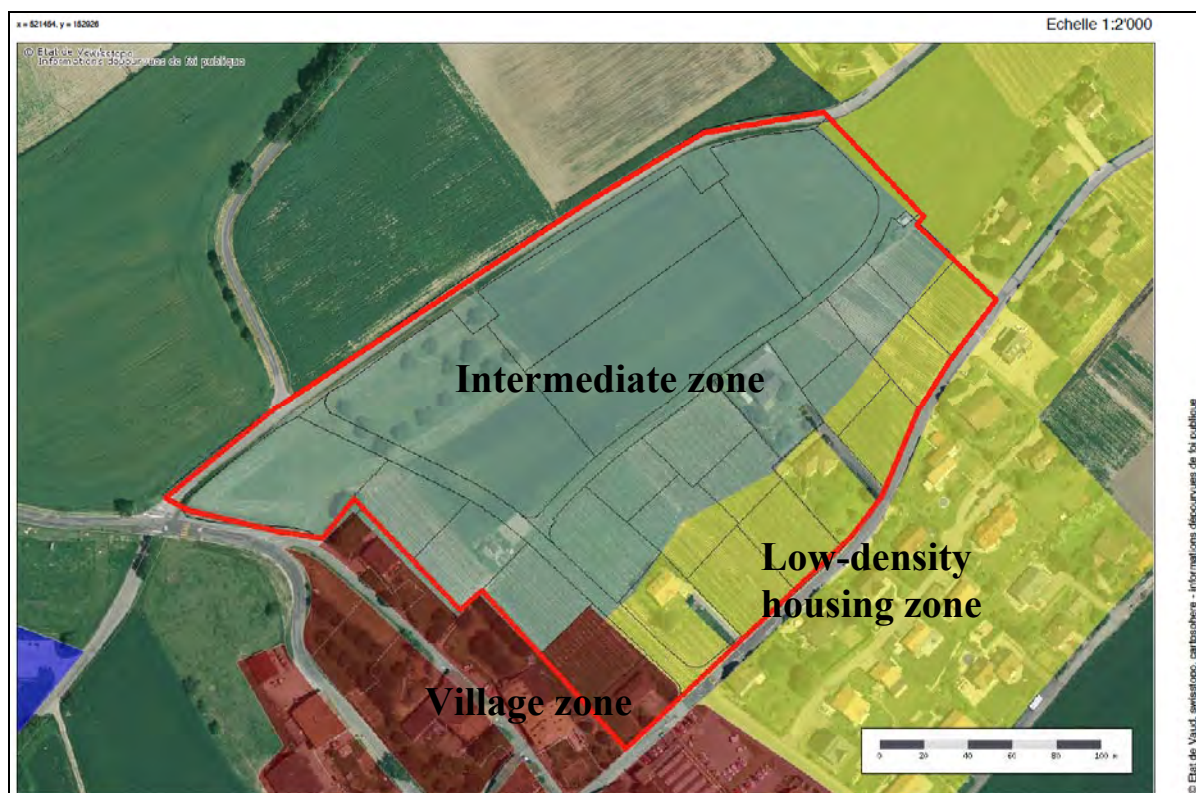


Figure 2: New land state and old zoning categories inside La Boironne area (from www.geoplanet.vd.ch).

5.2 Feasibility study

In order to move forward with this project, the feasibility study was launched the 8th of April 2002 with the following objectives. It had to analyse the urban project outline,

define estimating land value principles between old and new land state, define costs' distribution principles, draw up a first project outline for collective spaces and equipments and the resulting costs, and finally propose a land management approach for further operations (Robyr et al. 2003).

As we have said in theory, such a feasibility study should lead to four main results. What are they in our case? The panel of experts, who was in charge of the feasibility study, has produced a report (Robyr et al. 2003), which has been given to authorities. This first step is the starting point of the coordination process between spatial planning and land management issues. On its content, the report proposes the following results⁹.

About the territorial outline and planning principles, experts conclude to the feasibility of the partial zoning plan. In addition, according to the land use index, the plan capacity is estimated to 160 new inhabitants distributed on three main parts (see figure 3): 16 housing units ($150\text{m}^2/\text{housing}$) within the orange area (village area); 18 housing units (150m^2) within the yellow area (villas); 22 individual housing units (175m^2) and 15 common settlements (150m^2) within the light-grey area (intermediate zone subject to real equalization).

Collective spaces and equipments are planned to provide new roads, collectors for wastewaters and freshwater distribution, electricity, gas and telecommunications networks. Investments, which have to be done, are estimated to approximately 1,400,000€ (25€ per square meter). 85% of this sum concerns collective spaces and equipments. The rest belongs to studies. It is considered to be an average price.

In order to implement the spatial planning project, the adopted land management approach is that of the voluntary LI syndicate, because experiments show that in such a complex cases (many landowners, readjustment, etc.) this way is the only one that can reach the goals within a reasonable period of time. In addition, the report contains land value estimating principles and some more details on real equalization principle.



Figure 3: Areas distribution according to the partial zoning plan (from www.geoplanet.vd.ch).

⁹ Results presentation follows the same structure as described in the chapter 3.1.

5.3 *La Boironne* LI syndicate

Following recommendations of the feasibility study, *La Boironne* LI syndicate is constituted on November 8th 2004. It is made up of 21 landowners and its perimeter contains 34 plots. In reference to the third article of the syndicate statutes, its aims are the approval of the partial zoning plan, the land readjustment that allows the implementation of the partial zoning plan, and the servicing of the land and construction of collective spaces.

As we have seen above, the process takes place in three main phases. The first step, which is very important in terms of coordination, has already been completed. Apart from two oppositions, which have been quickly resolved, the process was conducted according to the model presented above (see table 1). In deed, one of these was the compensation of an orchard situated inside the intermediate zone. In this case, the municipality supplied a parcel outside of the perimeter in order to replant it. Currently, the procedure is not yet complete. It still lacks the transfer of property rights and the coming into force of the partial zoning plan, a crucial step for obtaining the necessary bank loans. Accordingly, the second and the third step need some works more. The servicing of the land has started in November 2009, and is still in progress. From a financial point of view fees paid by landowners are currently sufficient to pay the initial works. Thus, it seems likely that the syndicate will be dissolved during the years 2010/2011. Therefore, since the launch of the feasibility study in April 2002, between 8 and 9 years will have passed.

5.4 Case study results

This case study allows us to highlight the following results. First, the process of land management developed by the canton of Vaud (feasibility study which leads on the choice of the appropriated land management tool) gives to authorities the possibility to integrate all the stakeholders inside the process of territorial development. It is thus possible to compare the functioning of the LI syndicate to that of a municipality. Of course the concerned area is much smaller. Nevertheless, in terms of decision, the process works almost identically, whether a majority of owners that can impose their choice to the rest by having integrated the opinion of the minority. The decision is thus accepted democratically, since the vote has been endorsed.

Another very important element is the capacity of the LI syndicate tool to manage the compartments of urban production (Prélat-Droux 2008) at the same time: zoning (building rights), readjustment (property rights), and collective equipments and spaces. Thus, the implementation of the zoning plan is not only dependent on the goodwill of the landowners. In stead, each compartment is producing effects on the others what we should manage at the best in order to reach a result of quality. For example, in terms of zoning, the LI syndicate allows to plan an urban project (partial zoning plan) by integrating the wishes of the owners directly at the start of the process. In fact, landowners become shareholders of the project. This means that the value that landowners have at the start is proportionally redistributed at the end of the project, thus conforming to the equal treatment principle and to the real equalization principle. This is what might be called the principle of value conservation, by referring to first law of

thermodynamics in a closed system. Goods are not lost, but in constant mutation with reference to land readjustment. In terms of land value, a part of capital gain can be levied by getting land in order to build collective spaces and equipments. Moreover, one very important thing is that land values increased during the whole process¹⁰. Thus, landowners are in any case better off at the end. This is probably why the approach presented works so well in this case. On the other hand, if the landowners had not participated in the process, they would have had difficulties to obtain such a capital gain. Thus, their plot would stay unused increasing the problem of land availability for municipalities. Finally, the procedure guarantees the construction of collective spaces and equipments since the members have the obligation to implement the plan, following the public inquiry. In terms of time frame, we can easily say that the process has allowed the resolution of a stalled situation in a short period of time.

6 Conclusion

In this study, we presented a case study dealing with the canton of Vaud's land management approach. We started from a theoretical point of view by discussing about the structure and the content of the Swiss land management regime. We described its current arrangement by defining two important constitutional articles. This allowed presenting one big problem we have to cope in Switzerland, land hoarding, which can be defined as a measure of the gap between a *potential supply* of buildable land (defined by the planes) and a *real supply* of buildable land available for construction (defined by the land markets). In order to show some of the existing possibilities to reduce this gap, we described one approach used in the canton of Vaud, which combines a feasibility study with a LI syndicate land management approach. We showed these theoretical bases and analyse them through the case of Yens, Switzerland. Finally, we presented specific results about the case study, by which we can say that the proposed process for this case is working very well.

Now, regarding our starting questions and hypothesis, we can point to the following results. According to our hypothesis, we firstly mention that this land management approach brings a very complete mechanism to involve landowners in planning projects so that the zoning plan can be realized. In fact, the coordination principle between spatial planning goals (public interests) and land issues (private interests) seems to be the cornerstone of all the process. Thus, achieving spatial planning goals, and more generally spatial public policies, seems to be strongly conditioned by the necessity to coordinate stakeholders' interests (private and public).

Secondly, from a local point of view, the land use seems to be improved both in terms of space and time. Indeed, the implementation of the process has allowed the resolution of a conflict situation in a short lapse of time. Furthermore, the partial zoning plan has been updated according to the last criteria of spatial planning which will ensure a result of quality.

Thirdly, as we have discussed above, the Swiss Constitution guarantees the institution of property rights. Under this basis property rights can be restricted only under certain

¹⁰ This information comes from a conversation we have had with Denis Leroy December 9th 2009 at the Territorial Development Service of the canton of Vaud.

conditions. In this context, the presented land management approach does not challenge the guarantee of ownership since the owners retain control over their land. Another important element is the fact that the land market stay largely dominated by private landowners. Indeed, this approach does not consider the purchase of lands by the State on the ground of the implementation of the zoning plan.

Finally, by looking at a municipal scale, urban production seems to be well situated from the spatial planning point of view. However, if we look more global (spatial scale), commuters should probably inhabit such an urban production, what has become the norm now. This leads to an increase in terms of mobility and an increased land consumption, which is not desired in a long term. Thus, the future of human activities seems likely to play within the limits of cities. One of the next steps, which is now considered as a big stake, is the implementation of such approaches in brownfields where we can identify a very important source of land reserve for the next fifty years and where developments are currently widely challenged because of a lack of a suitable land management approach to such cases.

7 Acknowledgements

The author wishes to acknowledge the persons that have contributed to the writing of this paper. In alphabetical order: Denis Leroy (Service du Développement Territorial, canton of Vaud), Roland Prélaz-Droux (University of applied science, Yverdon-les-Bains), Jean Ruegg (University of Lausanne), Leah Tillemans (University of Lausanne).

8 References

- BCV, 2007. *Vers une pénurie structurelle de logements?*, Lausanne: Banque Cantonale Vaudoise.
- Canton de Vaud, 1987. *Loi sur l'aménagement du territoire et les constructions (LATC)*.
- Canton de Vaud, 1962. *Loi sur les améliorations foncières (LAF)*.
- Comby, J. & Renard, V., 1996. *Les politiques foncières*, Paris: Presses universitaires de France.
- Leroy, D., 2008. *Coordination. Aménagement du territoire et aspects fonciers: descriptif du système vaudois et exemple d'application*, Lausanne: SDT.
- Nahrath, S., 2005. Le rôle de la propriété foncière dans la genèse et la mise en oeuvre de la politique d'aménagement du territoire: quels enseignements pour la durabilité des aménagements urbains? In *Enjeux du développement urbain durable*. Lausanne: PPUR, p. 299-328.
- OFS, 2009. *Portrait démographique de la Suisse. Edition 2009*, Neuchâtel: Office fédéral de la statistique (OFS).
- Prélaz-Droux, R., 2008. Des instruments fonciers pour faciliter les projets territoriaux.

Urbia: les cahiers du développement urbain durable, (6), 43-64.

Robyr, S., Longchamp, P. & Biner, B., 2003. *Etude préliminaire en amélioration foncière et aménagement du territoire. Secteur de la Boironne*, Morges: Canton de Vaud, Commune de Yens.

Rüegg, J., 2000. *Zonage et propriété foncière*, Paris: ADEF.

Service des améliorations foncières, 2005. *Fonctionnement et financement des syndicats d'améliorations foncières*, Lausanne: Etat de Vaud, Département des infrastructures. [on line] [http://www.publications1-dinf.vd.ch/prod/dinf/publicationdinf1_p.nsf/ref/C7F0BE25D3D5B2A2C125712B0046B0EB/\\$file/SAF-A4.2005.pdf](http://www.publications1-dinf.vd.ch/prod/dinf/publicationdinf1_p.nsf/ref/C7F0BE25D3D5B2A2C125712B0046B0EB/$file/SAF-A4.2005.pdf), consulted January 4th 2010.

Service des améliorations foncières (eds), 2003. *Les démarches foncières en Pays de Vaud. Guide à l'usage des communes et des entreprises AF*, Lausanne: Etat de Vaud, Département des infrastructures.

Implications of Commons Act 2006 for the registration of land as a town or village green

Eur Ing Dr Phebe Mann CEng MICE MRICS MCIOB MCIArb
 School of Construction Management and Engineering, University of Reading, P.O. Box 219, Reading, RG6 6AW, UK
 phone +44 (0) 118 378 8200
 fax +44(0)118 931 3856
 p.mann@reading.ac.uk

1. Introduction

Town and village greens originate in customary law. Baroness Hale of Richmond, in *Oxfordshire County Council v Oxford City Council and another* [2006] UKHL 25, considers them a very present amenity to the communities they serve. Town and village greens are, as defined in s22(1) of the *Commons Registration Act 1965*, land which is defined settlements or geographical areas where local inhabitants have indulged in lawful sports and pastimes including organised or ad-hoc games, picnics, dog walking, fêtes and similar activities for over 20 years as of right. It is presumed to be a village green if the rights to indulge in sports and pastimes were exercised as in *R. v Norfolk CC Ex p. Perry* (1997) 74 P. & C.R. 1 QB D and *R. v Suffolk CC Ex p. Steed* (1998) 75 P. & C.R. 102 CA (Civ Div) cited in *Oxfordshire County Council v Oxford City*. As long as the land owner preserves the open spaces in the public interest, as distinguished in *JA Pye (Oxford) Ltd v United Kingdom* (44302/02) (2006) 43 E.H.R.R. 3 E CHR, he can retain his title and right to use it in any way which did not prohibit the local inhabitants for recreation. The land use can be either continue to be used in this way or have ceased to be used for more than a prescribed period. Some town or village greens can be privately owned or owned by the Parish or Community Councils. The land has to be used “as of right” that is *nec vi* (without force), *nec clam* (without secrecy) and *nec precario* (without permission), for at least 20 years. If a landowner grants permission for use where there was over 20 years use “as of right”, s.15 of the *Commons Act 2006* regarded this “as of right” to continue, no time limit for an application for registration is imposed. A significant provision is contained in s15(7) of the *Commons Act 2006*, where 20 years’ use has accrued, a landowner who then gives permission for the use will not thereby prevent the use being “as of right”. Prior to the Act in the Oxfordshire case cited above, the House of Lords had decided this would prevent any later application for registration.

The *Department for Environment, Food and Rural Affairs (DEFRA)* defines *common land* as open, unfenced and remote land owned by a person, but another person is entitled to exercise rights of common over the owner’s land. The exercise includes grazing animals, cutting bracken for livestock bedding and other similar activities. Such rights are normally exercisable in common with others. A

right of common is defined in Cooke's *Inclosure Acts* as "a right, which one or more persons may have, to take or use some portion of that which another man's soil naturally produces". Devon County Council indicates that the rights of common may include *herbage*, the grazing of live stock such as sheep and cattle, *turbary*, obtaining turf and peat, *estovers*, collecting wood, gorse or furze, *piscary*, fishing and *pannage*, feeding pigs by acorns or beechmast¹

2. Additional two-year grace period

For the purpose of s22(1A) of *Commons Registration Act 1965 (CRA 1965)* the use for sports and pastimes had to continue until the date of the application for registration. One of the main changes brought about by s15 of the *Commons Act 2006* is that it allows applicants an additional two-year to make an application after the town and village green users have been prohibited by the landowner, where the inhabitants' use ceased before the time of the application but after 6 April 2007 (i.e. when s15 came into force), and the application is made within the period of two years after the cessation of use, s15(3). The criteria of satisfying the use of land for lawful sports and pastimes as of right for at least 20 years, remains unchanged². However, the period allowed for developers to object to the application remains to be 6 weeks. It can be argued that the extended period of the 2 years provides the applicants a better chance to be successful hence make it more difficult to develop this category of land which has the potential to register as greens³. This will create a greater extent of uncertainty for developers who would like to develop open land. Furthermore, the period of grace under s15(4) is applicable where the inhabitants' use ceased before 6 April 2007 and the application is made within a period of five years after cessation of use. The exception to this is set out in s15(5) where planning permission was granted before 23 June 2006 in respect of the land and construction works were commenced before that date in accordance with that permission. This exception only applies, however, where the land in question has or will become permanently unusable by members of the public for lawful sports and pastimes.

Developers may have to take greater measures to stop public use of the part of the land they plan to develop in future as soon as possible, so as to minimise the opportunity of any possible applications before the two years end. They have to take measures such as erecting fences and putting up warning sign declaring the land to be private

¹ Devon County Council Local Authority Searches (Land Charges) Common Land and Village Green Searches http://www.devon.gov.uk/index/environmentplanning/planning-system/local_authority_searches/common_land_searches.htm (accessed on 15 July 2009)

² *Inhabitants must use the land or lost it Commons*; Land registration; Village greens. E.G. 2006, 0632, 79

³ *Village greens and the fight against development* K.L.J. 2007, 18(1), 145-157

3. Statutory period of closure

Provisions in *s15* seem to be in favour of the applicants against the developers. The continuity of the 20 years would not be considered broken if there was a statutory period of closure of land, for example, foot and mouth outbreak closure, where access to the land is prohibited to members of the public by an enactment, *s15(6)*. Furthermore, it cannot prevent an application to succeed if there was a permission to use the land after the land had been used as town and village greens for 20 years (*s15(7)*). In both above circumstances, the developers are in a weaker position to object. These new provisions will make it harder for the developer to develop land

4. Correction of mistakes on registers

Another new provision is to allow mistakes on the registers to be corrected. Under *Schedule 2* of the *Commons Act 2006*, for village greens or common land which is mistakenly not registered under the *CRA 1965* but are village greens or common land under older statute, however, applications need to be made before 31st December 2020. This date also applies to waste land of a manor mistakenly not registered first time round. For the deregistration of building and curtilages, such as cottages or garden on or abutting the common or green, which are wrongly registered as common land, this also needs to be made before 31st December 2020. This provision may not make future development easier, but at least, landowners and developers are permitted to correct wrongly registered building and curtilages as common land.

5. Disallowing severance of rights

Under the *Commons Act 2006*, severance of rights can no longer be taken, other than by *Natural England* or *Commons Council* applications. One of the main reasons why the severance of common rights is prohibited is to prevent commoners from selling, leasing or letting their rights away from the property to which rights are attached⁴. Severance reduces the efficient use of land for grazing; it creates problems for local right holders to use the land for grazing⁵. The prohibition of severance of rights may reduce the opportunities for landowners and developers to develop land of this category

6. Voluntary registration

Voluntary registration of a town or village green was unknown under the *CRA 1965* but can now be registered under the *Commons Act 2006* or applied for, there is no time limit. Land owners may apply to register their land as a town or village green for use by local inhabitants voluntarily without showing any previous use of land for sports and pastimes (*s15(8)*). However, the leaseholder or

⁴ *The Dartmoor Commons (Authorised Severance) Order 2008*

⁵ *The Commons (Severance of Rights) (England) Order 2006 (SI 2006/2145)*

chargeholder's consent has to be obtained if the land is leased for over 7 years or subject to a charge or mortgage. Voluntary registration does not seem to add any advantage, if not harder, for future development of this type of land. Previously it was not possible for landowners to legally dedicate their land as a green, as the use would then have been with permission and so not as of right.

7. Replacement land in exchange

Developers can apply to deregister common land for development under s16 of the *Commons Act 2006*, but in most circumstances, an equivalent replacement land in exchange has to be provided and the overall application has to be in the interest of the public. However, the *Commons Act 2006* intends to provide safeguards for common land against deregistration in order to ensure a long term protection; this makes future development on potential greens more difficult⁶.

Some other changes brought about by *Commons Act 2006* include attachments, variation and re-allocation.

Land owners can apply to create and register a new right of common on attachments under s6 of the *Commons Act 2006*. This can be a right over an existing registered common land or unregistered land. These new rights of common must be attached to land in order to be exercisable by the land owner. Land owners of the right of common in gross can apply to attach land for example farm holding under s10 of the *Commons Act 2006*. Obviously, this would not provide more advantage to develop land which has the potential to be registered as a green.

Land owners can also apply to vary an existing right of common which is exercisable over new land under s7 of the *Commons Act 2006*. The variation can be the change of the type of stock which may be grazed. Only the owner of the land to which the right of common is attached or of the right of common held in gross can apply under s7.

Land owners can apply to re-allocate a right of common attached to future non-agricultural land under s11 of the *Commons Act 2006*. They can apply where part of the land has been acquired by any authority in accordance with the *Acquisition of Land Act 1981* by *Compulsory Purchase Order*.

Although the changes concerning attachment, variation and re-allocation do not explicitly make future development more difficult, there is no encouragement to development.

⁶ *How green is my law?* L.S.G. 2007, 104(23), 36

8. Conclusion

The purpose of the *Commons Act 2006* is to protect common land and promote sustainable farming, public access to the countryside and the interests of wildlife. The changes under s15 of the *Commons Act 2006* include the additional 2-year grace period for application, discounting statutory period of closure, correction of mistakes in registers, disallowing severance of rights, voluntary registration, replacement of land in exchange and some other provisions. These new provisions under the *Commons Act 2006* seem to make it more difficult in future to develop land which has the potential to be registered as a town or village green. The DEFRA's guidance stated clearly that "if building work has gone ahead and the land is subsequently registered as town or village green, the local inhabitants will have recognised recreational rights over the land...", yes, the transitional provision contained in s15(4) CA 2006 is particularly a cause for controversy as DEFRA has indicated buildings will have to be taken down where development has gone ahead and a subsequent application to register the land as a green is successful, obliging the developer to return the land to a condition consistent with the exercise by locals of recreational rights, which sums up that it would be harder in future to develop land which has the potential to be registered as a green.

References

- Department for Environment, Food and Rural Affairs, Wildlife & Countryside (2009) *Competing for the same patch of grass Applications for registration; Golf courses; Land use; Village greens*. E.G. 2009, 0915, 97
- Department for Environment, Food and Rural Affairs, Wildlife & Countryside Website <http://www.defra.gov.uk/wildlife-countryside/protected-areas/common-land/tvg.htm>
- Devon County Council Local Authority Searches (Land Charges) Common Land and Village Green Searches http://www.devon.gov.uk/index/environmentplanning/planning-system/local_authority_searches/common_land_searches.htm
- How green is my law? Commons; Land registration; Landowners; Village greens*. L.S.G. 2007, 104(23), 36
- Inhabitants must use the land or lost it Commons; Land registration; Village greens*. E.G. 2006, 0632, 79

Landowners face a snake in the grass Development; Registration; Village greens. E.G. 2007, 0711, 155

New town and village greens Land registration; Land use; Prescription; Protection of property; Village greens. R.W.L.R. 2007, Oct(15.3), 123-127

Oxfordshire County Council v. Oxford City Council Application date; Applications for registration; Land use; Landowners; Village greens. R.W.L.R. 2006, Jul(15.3), 97-105

Meager, R (2009) *A set back for the "village green industry?"* [2009] 68(2) CLJ 281-283

Meager, R (2007) *New Town and Village Greens: The Calm Before the Storm?* Rights of Way Law Review (RWLR), October 2007, Section 15.3, 123-127

Meager, R (2006) *New Town and Village Greens: Back From the Brink* Oxfordshire County Council v Oxford City Council and Another [2006] UKHL 25 [2006] 70 Conv 584-598

Smith, R. J. (2006) *Property Law Cases and Materials*, 3rd Edition, Pearson

Source: <http://www.defra.gov.uk/wildlife-countryside/protected-areas/common-land/about.htm> Access on 13 July 2009

The Law Commission and HM Land Registry (1998), *Land Registration for the Twenty-First Century; a consultative document* CM 4027 (Law Com No 254) London, The Stationary Office, 1998

The Law Commission and HM Land Registry, *Land Registration for the Twenty-First Century a conveyancing revolution* HC114 (Law Com No 254) London, The Stationary Office, 2001

The trap grounds, the sports arena and the Commons Act 1996 Applications for registration; Commons; Land use; Local authorities' powers and duties; Registration; Village greens. J.P.L. 2007, Feb, 183-193

Village greens and the fight against development Legislative intention; Planning control; Registration; Village greens. K.L.J. 2007, 18(1), 145-157

Statutes

Commons Act 1876

Commons Act 1899

Commons Act 2006

Commons Registration Act 1965

Countryside and Rights of Way Act 2000

Inclosure Act 1845

Inclosure Act 1857

Interpretation Act 1978

Land Registration Act 1925

Land Registration Act 2002

Law of Property Act 1925

Law of Property Act 1969

Local Government Act 1972

Local Government, Planning and Land Act 1980 (c.65)

Open Spaces Act 1906

Prescription Act 1832

The Commons (Severance of Rights) (England) Order 2006 (SI 2006/2145)

The Dartmoor Commons (Authorised Severance) Order 2008

Cases

JA Pye (Oxford) Ltd v United Kingdom (44302/02) (2006) 43 E.H.R.R. 3 ECHR

Oxfordshire County Council v Oxford City Council and another [2006] UKHL 25

R. (on the application of Cheltenham Builders Ltd) v South Gloucestershire DC [2003] EWHC 2803

R. (on the application of Lewis) v Redcar and Cleveland BC [2008] EWHC 1813 (Admin); [2008] N.P.C. 97; (QBD (Admin))

R. v Norfolk CC Ex p. Perry (1997) 74 P. & C.R. 1 QBD

R. v Suffolk CC Ex p. Steed (1998) 75 P. & C.R. 102 CA (Civ Div)

The contemporary urban commons – a case study of Darling Harbour, Sydney

Spike Boydell¹ and Glen Searle^{1,2}

¹Asia-Pacific Centre for Complex Real Property Rights,
University of Technology, Sydney, PO Box 123, Broadway, NSW 2007
Australia

²School of Geography, Planning and Environmental Management,
University of Queensland, Brisbane, St Lucia, QLD 4072,
Australia

Email: Spike.Boydell@uts.edu.au ; g.searle@uq.edu.au

Abstract:

This paper reports on research funded by the RICS Education Trust. We offer an analysis of the contested state of contemporary urban commons in an Australian city, through a case study of the Darling Harbour scheme in Sydney. Moral claims of ‘ownership’ often outweigh legal claims of ownership in the scarce public spaces that comprise the contemporary urban commons. We question if these contemporary commons provide a counter-weight to the privatised rigidities of urban capitalist societies.

Our research design incorporated an analytical hierarch process (AHP) to test a ‘mosaic’ of property rights in the contested commons. This mosaic essentially comprises a spectrum of property rights, resulting in a variety of contemporary ‘commons’ that range from spaces with unrestricted public access to private leasehold property with limited public access rights. We identify several distinct types of commons at Darling Harbour. We engaged an Expert Choice analysis of interview data to ascertain the relative importance of each kind of property right for users and lessees of different kinds of spaces in Darling Harbour. This enables Darling Harbour to be mapped as a heterogeneous commons comprising a set of spaces that each has their own distinctive mosaic of predominant perceived property rights.

Keywords: property rights, contemporary commons, AHP, Darling Harbour, Sydney

1 Introduction

The issue of the commons has received increasing attention from an environmental perspective. There has however been a dearth of research into the urban commons. In

this paper, we investigate real property rights in the contemporary urban commons - the perceived public spaces that are found interwoven into the tapestry of the urban environment. The city comprises a milieu of competing and complementary property rights, ranging from the individual to the communal. Whilst property rights provide a coherent legal, economic and social framework for the relationship between people, place and property, they are often misunderstood and misinterpreted by the multiplicity of stakeholders sharing the space that is the contemporary metropolis.

The importance of treating the urban commons as a distinct research area is underlined by McShane (2010, 107), who states

... analysis of urban commons requires distinct conceptual and contextual frameworks, requiring new thinking about sectoral relationships and calling into question the separation of nature and culture that is reproduced in the rubric of old/new commons.

One of the central issues here is the diversity of rights, obligations and restrictions that can apply to an urban commons. This diversity leads to the need to conceive of a mosaic of real property rights that govern these public (and semi-public) spaces. The need for such a perspective has been identified by Kohn (2004, 10):

The term commons . . . erases the distinction between a number of different kinds of collective property. The commons of a gated community is not the same as a Boston Commons. We need a language that helps us distinguish between apparently similar forms of collective ownership that have very different social and political effects.

We provide an overview of property rights and the contrasting legal, economic, and sociological, interpretations and the conceptualisation of rights as being composed of a mosaic. This is used to frame an analysis of the diversity of real property rights in the contemporary urban commons of Darling Harbour, Sydney. The analysis compares perceived property rights within our mosaic and identifies the main typologies of rights in each category of the commons in Darling Harbour.

2 Property Rights and the Urban Commons

As Bollier (2008, 4) identifies, “a commons arises whenever a given community decides that it wishes to manage a resource in a collective manner, with special regards for equitable use, access and sustainability”. The contemporary urban commons, as a particular type of inherently public property, are represented by spaces of public interaction. They are important spaces between buildings where people can meet, eat, play and relax. As Kohn (2004, 194) argues, “These places are not consumed or diminished when others enjoy them. In fact, in many places, their value actually increases when a critical mass of strangers congregates”. Nevertheless, this obscures the existence of a range of rights, obligations and restrictions attributable to each of the stakeholders who use and enjoys such spaces. Given that moral claims of ‘ownership’ often outweigh legal claims of ownership, the identification of the rights that may apply is critical to understanding the contested state of contemporary urban commons.

3 What makes a Property Right?

Discourse on property rights has emerged within a broad range of disciplines. These include, but are not limited to, archaeology, anthropology, ethics, sociology, history, psychology, law, geography, biology, philosophy, economics, and planning. The most influential Western theorising about property is underpinned by what Hann (2007, 290) refers to as the ‘standard liberal model’, yet property is “much broader than the liberal tradition recognizes, and that the political, economic and social functions of property are in continuous flux”. Many of these disciplines draw heavily on legal traditions, and in particular have been influenced by Henry Maine’s (1861) metaphor of ‘a bundle of rights’. The importance of the ‘bundle’ metaphor is that it highlighted the common circumstance that different individuals or groups may hold differing rights, obligations and restrictions over the same parcel (or piece) of land. The understanding and articulation of property rights, obligations and restrictions influence property relations in all human societies.

Anthropological engagement with property issues has often been overtaken by the standard liberal model. These models range from Hardin (1968) wanting to avoid a ‘tragedy of the commons’, the ethnographic cost-benefit analysis Demsetz (1967), the new institutional economics of North (1990), through to the beguiling capital creation propositions of de Soto (2000). Likewise, we find the disciplines often evolve their interpretation of property rights in isolation and concern themselves with different aspects (McEwin, 1993, 35):

Lawyers, at least academic lawyers and judges, rely on principles of jurisprudence to determine ‘fair’ property right allocations on the basis of ethical and moral standards. On the other hand, economists tend to evaluate property-right systems on the basis of whether they ensure resources go to their most valued use.

Table 1: Hohfield's System of Jural Relations
(Source: Hohfield, 1916-1917, 710)

Jural Opposites	right	privilege	power	immunity
	no-right	duty	disability	liability
Jural Correlatives	right	privilege	power	immunity
	duty	no-right	liability	disability

As Cole and Grossman (2002, 318) highlight, “divergent conceptions of property rights can lead to differences in analysis and to confusions in cross-disciplinary scholarship”. Caution needs to be taken when research into property rights, such as that evolving within the Asia-Pacific Centre for Complex Real Property Rights, strives to be transdisciplinary (Nicolescu, 2006, Max-Neef, 2005), and navigates the boundaries of diverse disciplines. As Bromley (1991) highlights, there are few concepts in economics that are more central, or more confused, than those of property, rights and in particular property rights. Loose talk about property, rights and property rights is not new, indeed in the early twentieth century Hohfield (1913-1914, 1916-1917) offered a system of

jural relations (rights and obligations) in an attempt to avoid such misunderstandings (see Table 1).

Whilst Hohfield's system has become an influential and enduring contribution to analytical jurisprudence, we determined to analyse these relationships in the context of the contemporary urban commons in order to discern the level to which the confusion persists. To do this, we engaged a synthesis of property rights, obligations and restrictions developed by Boydell (2007, 111), which drew on a range of contemporary property rights literature (see Crocombe, 1975, Bromley, 1991, Payne, 1997, Rigsby, 1998, Sheehan and Small, 2002, World Bank, 2003, Farran and Paterson, 2004). We then applied these rights, obligations, and restrictions, to the institutional arrangements prevailing in the contemporary urban commons of Darling Harbour, Sydney. On closer inspection (see Table 2), it will be apparent that some of these *rights* (e.g. duration and flexibility) are actually attributes of more fundamental rights, rather than separate individual rights.

The standard libertarian view (that was criticised above by Hann, 2007) regards ownership as the highest level of right to act, and that property rights vest only with the individual. In contrast, Lyons *et al.* (2007) view is that with property rights at any level come roles, obligations and restrictions. This perspective is especially crucial in considering challenges concerning the maintenance and development of the urban commons, where spaces are potentially under threat of conflicts over the roles and obligations of the various stakeholders, who see themselves as asserting certain properties rights over the commons. At one level this can be seen as an issue of information and negotiation transaction costs (Webster and Lai, 2003, Coase, 1960). But at a deeper level such conflicts need to be understood in terms of a more nuanced view of the nature of property rights (Cole and Grossman, 2002), of varying ideological beliefs about the urban commons, and of implicit power relationships between stakeholders (Flyvbjerg, 1998).

Rifkin (2000) highlights an increasing emphasis on the importance of 'access' rather than ownership. Access is key in the urban commons, as "These places are not consumed or diminished when others enjoy them. The magnetism that attracts people to make diverse use of public space is what Kohn (2004, 104) refers to as the 'carnival of the commons'. The *values* are not necessarily economic. Rather, what is valued in urban commons is:

"...places to sit, plenty of shade, 'touchable' water, good food, and well-connected streets and sidewalks. But what a vital commons can offer a community is far from simple. A sense of identity. Belonging. Connectedness. Fun. Tradition. Stories. Great public spaces are what memories of cities are made of" (McGinnis, 2001).

There has been significant analysis of the goals of, and impacts on, the range of stakeholders involved in urban planning schemes that involve consideration of the commons, especially using Lichfield's planning balance sheet / community impact evaluation methodology (Lichfield, 1996). However, the treatment of the commons in such analyses has considered it as a product of the goals of different stakeholders rather than as being a contested site for different property rights, and where the way in which

those rights are resolved in each area of the commons has central implications for urban sustainability.

Table 2: Property Rights Summarised
(Source: Boydell, 2007, 111)

Right	Explanation
Direct use	Rights to plant, harvest, build, access and similar, maybe shared rights
Indirect economic gain	Such as rights to tribute or rental income
Control	Conditions of direct/indirect use, held by persons other than the user
Transfer	Effective power to transmit rights-by will, sale, mortgage, gift, or other conveyance
Residual rights	Remaining rights at the end of a term (such as lease, death, eviction), includes reversionary rights
Rights of identification (symbolic rights)	Associated with psychological or social aspects with no direct economic or material function
Duration	Length of time property right is held, indicating profits and/or savings
Flexibility	Right should cater for modifications and alterations
Exclusivity	Inverse of the number of people with shared or similar rights, more relevant to water property
Quality of title	Level of security that is available as tenure shifts from the optimum of notional freehold
Divisibility	Property right can be shared over territories, according to season, etc.
Access	Entry/ admission onto the land
Withdrawal (extraction)	Extraction of resources by owner despite leasing property
Management	Be able to make decisions on how and by whom a thing shall be used
Exclusion	Disallowing others from entry and use of resources
Alienation	Transfer of an interest (right) in property to another, in perpetuity

In contrast, our systematic analytic framework for understanding property regimes, as with several other projects being undertaken by the Asia-Pacific Centre for Complex Real Property Rights, is influenced by the thinking of von Benda-Beckmann *et al.* (2006) of the Legal Pluralism Research Group, at the Max-Planck Institute for Social Anthropology. Their model distinguishes four layers. The

ideological layer consists in the norms and values of a cultural tradition. The second layer comprises political and legal regulations. The third layer consists of the social relations of property. The final ‘concretised’ layer is the layer of practice, where actors operationalise or reinforce the patterns of the other layers, or strive to initiate changes. Property must be analysed at all four layers.

Having provided some background to property and property rights, in the next section we will provide locational and spatial context for the case study site.

4 Darling Harbour

Sydney’s Darling Harbour provides a lens through which the mixed use contemporary commons can be explored (see Figure 1). Darling Harbour, on the western boundary of the Sydney CBD, has been a port area since the early days of the city. The construction of a new port to the south at Botany Bay in the late 1970s, to relieve congestion in port activities in Sydney Harbour generally, caused an extensive area of land at the southern end of Darling Harbour to become available for redevelopment. In 1984, the state Premier announced the government’s intention to develop Darling Harbour to capture more of the already expanding global tourism numbers coming to Sydney. The redevelopment included a major exhibition and convention centre, a harbourside market / retail precinct, and hotels along the foreshore, as well as a Chinese garden and maritime museum. A casino was also included, with the licence fee to NSW Government intended to recoup much of the overall development costs of the project.

To carry out redevelopment, the state government established a special authority (now merged into the Sydney Harbour Foreshore Authority) with its own planning powers. The large land holdings held by the state government at Darling Harbour vested in the authority. The result of Darling Harbour’s redevelopment has been the formation of a space with a complex and uneven landscape of property rights, in which public ‘rights’ vary from site to site depending on the particular institutionally-driven micro-history of each land ‘parcel’. Up to the 1980s, when Darling Harbour was Sydney’s main port area, the foreshore was predominantly owned and administered by two state government authorities: the Maritime Services Board (port operations land) and the State Rail Authority (goods railway land). Public rights over these areas were essentially limited to access rights associated with carrying out port or rail business. A small proportion of the land around Darling Harbour was in private ownership, mostly warehouses and storage facilities associated with port activities, but also included incidental commercial premises such as hotels. The state government also owned Pyrmont power station, a stand-by facility that was being phased out with the construction of major power stations on the coalfields north and west of Sydney. Again, public access rights to these properties were generally limited to persons employed or doing business there. Thus while most of Darling Harbour land was in public ownership, the general public had few rights to enjoy the use of the space. There were no ‘commons’.

5 The mosaic of property rights

The formation of the Sydney Harbour Foreshore Authority and subsequent redevelopment of Darling Harbour as a form of contemporary commons where people

congregate to relax, walk, play, dine, shop, attend an event or watch a movie, has resulted in what we refer to as a *mosaic* of ‘public’ property rights (Boydell et al., 2007). This mosaic essentially comprises a spectrum of property rights, resulting in a variety of contemporary ‘commons’ that range from spaces with unrestricted public access to private leasehold property with limited public access rights. In detail, we can identify several distinct types of commons at Darling Harbour.

The **first** type of commons is the classic version of a space that is publicly owned and open for general public use. This type is found in the open space of Tumbalong Park (a multi-use grassed area) and the promenades around the harbour. Whilst accepted as commons, it is monitored by a range of security cameras and patrolled by park ‘rangers’.

The **second** type of commons are a shared space on land side of the promenades, where café’s and restaurants ‘spill’ over onto the commons to provide popularised ‘al fresco’ dining. This is achieved through a clear demarcation of space for which the café’s and restaurants pay a significant licence fee.

A **third** type of commons is a public garden with an entry fee – the Chinese Gardens. In this case, the public right to use the commons requires a payment, but the public space is indirectly controlled via the Authority.

A **fourth** type is an extension of the second type. It involves areas set aside for public spectacle that require an entry payment, but which are privately leased and operated. Darling Harbour has several such ‘commons’, including the Aquarium, Maritime Museum, and Australiana Zoo. We can view such areas as a type of commons in that their predecessors in central cities and elsewhere were usually provided by the government or non-profit organizations with free entry. The Darling Harbour versions retain a tenuous link with older public equivalents by dint of their long-term leases of land owned by a public authority.

A **fifth** type of commons are the public spaces in private facilities on land leased from the Authority, such as in the Casino in particular. These are a genuine commons insofar as they are on publicly-owned land that allows, to all outward appearances, free and unrestricted access. However, these spaces are in fact leased by the Casino company and are patrolled by security guards who do not have the right to exclude entry to anyone. They thus represent a version of the privatised commons. The Casino, whilst outside the SHFA boundary, is a key typology of commons as the licence agreement provided revenue that was used to finance the rest of the public development at Darling Harbour.

In addition to the above, but continuing the same logic, two additional types of commons can be identified as follows:

A **sixth** type of commons are the public spaces in private facilities on land leased from the Authority, and for which an entry fee is charged. These include the exhibition spaces and conference facilities that are hired for use and for which individuals typically pay for entry. These are a commons insofar as they physically offer the same level of

common access as the third and fourth types, but are managed by private business, rather than public entities. They provide a social space at a fee, but with common access to the corridors, toilet and eating facilities. To the extent that they provide a social space they are a type of commons, however the measure of their commonality is inversely related to the level of the charge for access. They thus represent a version of the privatised commons.

A **seventh** type of commons are employment spaces in private or public facilities on land leased from the Authority, to which only limited entry is possible. These are perhaps the most limited commons, in that access to them is restricted to certain persons, however the persons admitted are not the lessees, but persons who derive their livelihood from using them as employees, albeit that there is public access to foyers and circulation spaces. For these people, these spaces provide the physical context for the social interaction of employment, and they rely upon its free provision. They represent a valid, though restricted form of the privatised commons.

Table 3: The Mosaic of Property Rights
(Source: Boydell et al., 2007)

	Public land	
	Managed publicly	Managed Privately
Fee for Entry	3	4 2 6
No Fee for Entry	1 7	7 5

The dimensions of these various types of commons are shown in Table 3. This sets out a matrix of commons, some of which are managed privately, some publicly and for access to which fees may or may not be charged.

The rationale underlying the formation of this range of commons is the Authority's assessment of the mix that is needed to maximise development and thus payments to the state government. In particular, the provision of traditional commons in Tumbalong Park and elsewhere is intended to make Darling Harbour an attractive place for visitors. In turn, this will increase the custom for profit-making activities at Darling Harbour. Thus, Darling Harbour illustrates a new role for urban commons in an era of constrained government capital expenditure, with governance replacing government in the public

sphere to meet shortfalls in public provision. In the contemporary era of governance, the creation of a variety of urban commons can leverage private investment in addition to meeting traditional community needs for public spaces.

In this process, the creation of a relatively autonomous agency like the Sydney Harbour Foreshore Authority that has its own planning and development powers greatly facilitates these outcomes. In this sense, the Authority has effectively created a private city precinct in which public access and related property rights are given in ways that maximise the Authority's interests (Glasze et al., 2006). The Authority is vested with the powers normally held by local government to choose the bundle of developments, including various types of commons, that serve to benefit its goals. The transaction costs involved (basically, the operating costs of the Authority) have been kept in check by the granting of strong powers that require minimal negotiation with the rest of the public sector, and by contracting out initial design and development activities. Normal democratic processes for debating the supply of commons have been able to be excluded because of the Authority's powers to by-pass the public consultation processes of the NSW Environmental Planning and Assessment Act, and because of the original economic development and project deadline dimensions of the Authority's planning and development activities.

6 Interpreting the Commons

The various types of commons found in Darling Harbour reflect a hierarchy of social accessibility. Commons of first type are true commons, dedicated to the public for free personal and social use, subject to limitations of acceptable use. They provide the physical context for social action and connection.

The second type provide nominally free social access, but carry a social more for persons using the space to participate in a commercial relationship with the proprietors of the restaurants in order to validate occupation. Their value to the restaurant proprietor comes largely from their connection to the type one commons and the entire surrounding environment in its physical and social dimensions. The rent paid by the restaurant proprietor is based on the business that flows from this connection and the desirability it creates for the humans who use the area.

The third and fourth types of commons are private property owned by the state. They behave similarly to the fifth type, except that they have an element of social identification as public space which connotes a relationship between the owner and user that make the extraction of a profit repugnant. The consequent poor commercial performance of commons of this type deserves close study as an example of a variety of commons, such as public car parks and camping grounds that have in recent times been privatised to improve their commercial performance.

The fifth type illustrates the social dimension of private property. While this land is leased from a public authority, its treatment is private and its management is directed towards commercial profit. It functions as a social space, because commerce in any form is a social activity and the provision of services, such as gambling opportunities, is social.

The final two types of commons have conceptual links to the foregoing. They rely on their character as providing a space for social activity and derive their value from the matrix of surrounding commons. The seventh form of commons recognises that work, as well as leisure, is social, and even when nominally private, employment space is a commons.

7 Analysis

The identification of the variegated public rights surface of Darling Harbour used data from interviews with the multiple stakeholders who use the area for a range of business and recreational activities. Separate sets of property rights were identified for each of the above seven types of commons in Darling Harbour. For the area comprising publicly owned space for general public use, 25 interviews across the general public at different locations throughout Darling Harbour (see Figure 1) were conducted. Some 52% of our interviewees were from outside Sydney, compared with 46% for Darling Harbour as a whole during 2008 (based on the Darling Harbour Nov 2007-June 2008 visitor survey conducted by I-View Pty. for SHFA).

Our research instrument was developed from our summarised property rights (per Table 2). We created a multi-criteria decision making tool using Expert Choice software (Saaty, 1988), to formulate an interactive analytical hierarchy process matrix of property rights for the contemporary urban commons. This enabled us to collect data direct onto laptop computers, enabling respondents to see their judgements as they were recorded in the field. Each interviewee was asked for their perception of what property rights they held in that space (the type one space). For the other six types of space, lessees of each type of space were interviewed about the property rights they perceived themselves to have in each case. One lessee was interviewed in each case, except for Type 4, publicly leased space with an entry fee, where two lessees were interviewed. The interviews asked each person to rank different types of property rights (as set out in Table 2) in terms of the priority each one had compared to each other type of right (pairwise-comparison).

The paired rankings thus obtained were then converted into an overall priority listing of each property rights type perceived by users or lessees of each type of space (see Table 4). The conversion of paired rankings to an overall prioritisation was analysed using Expert Choice and then transferred into Microsoft Excel for refinement and presentation. The methodology generated an inconsistency measure in each case that indicated the extent to which rankings were inconsistent with each other (that is, showing the extent of intransitivity).

Figure 1: Survey locations, Darling Harbour
 (Source: Authors for this research – image courtesy of Sinclair Knight Mertz Pty)

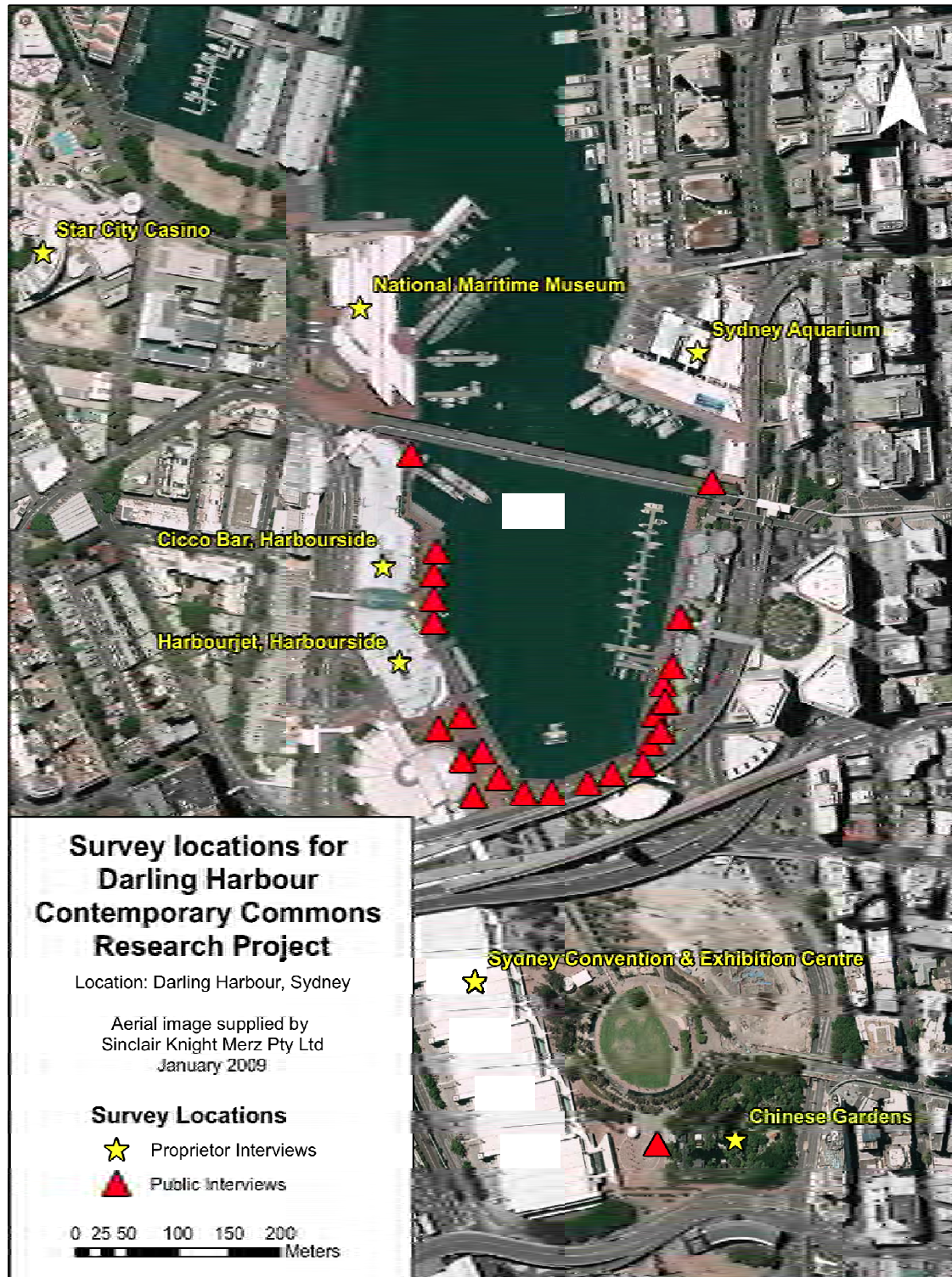


Table 4: Analysis of Property Rights Mosaic for Darling Harbour
(Source: Authors for this research)

Mosaic Type	#1	#2	#3	#4	#5	#6	#7
Property Right	Public Space	al fresco' shared	Public Garden (fee)	Public spectacle	Public in Private	Privatised (conv)	Restricted
Access	0.217	0.226	0.095	0.059	0.098	0.047	0.053
Alienation		0.008	0.024	0.008	0.020	0.008	0.008
Control	0.053	0.050	0.147	0.207	0.134	0.092	0.042
Direct use	0.211	0.233	0.055	0.122	0.082	0.106	0.181
Divisibility	0.033	0.009	0.037	0.038	0.009	0.112	0.008
Duration	0.150	0.159	0.104	0.079	0.159	0.045	0.194
Exclusion		0.017	0.018	0.023	0.023	0.020	0.033
Exclusivity	0.033	0.089	0.030	0.062	0.230	0.044	0.143
Flexibility	0.167	0.036	0.060	0.074	0.017	0.047	0.033
Identification	0.061	0.040	0.009	0.045	0.029	0.013	0.008
Indirect economic gain	0.037	0.015	0.023	0.162	0.071	0.239	0.161
Management		0.070	0.125	0.076	0.092	0.092	0.111
Quality of title		0.009	0.134	0.008	0.014	0.101	0.008
Residual rights	0.040	0.030	0.115	0.009	0.009	0.011	0.007
Transfer		0.010	0.023	0.030	0.015	0.024	0.011
Inconsistency = 0.00827 Inconsistency = 0.18 Inconsistency = 0.21 Inconsistency = 0.16 Inconsistency = .017 Inconsistency = 0.20 Inconsistency = 0.22							

8 Analysis issues

In the context of *Publicly owned open for general public use space* (category #1), five property rights typologies were not included. During the pilot study, it became evident that Alienation, Exclusion, Management, Quality of Title, and Transfer were not rights that members of the public could hold over the urban commons. In the initial interviews, it quickly became apparent that reducing the property rights to ten typologies streamlined the data collection on site and reduced confusion on the part of the respondents (Expert Choice recommends nine as the optimal). In part, the reduced number of typologies and the larger number of respondents resulted in a significantly more consistent rating of the relative importance of different property rights. We retained all 15 typologies in the other categories.

9 Results

The prioritisation of different types of property rights in regard to each type of commons space at Darling Harbour showed a differentiated landscape of perceived property rights. The importance of the different types varied from one type of space to the next. In this section we summarize the findings for each type of space and overview the differences by mapping the property rights priorities for each space.

#1: Publicly owned open space for general public use

The perceived property rights for this space are those shared by the general public. These are marked by a strong prioritization of access and direct use rights (Table 4). This reflects basic reasons as to why the general public visit Darling Harbour – to enjoy walking, resting and recreating in the general public spaces and having the assurance that they can do so when ever they come to visit. Table 4 indicates that rights embodying flexibility and duration are also perceived as being significant for the general public. Flexibility is important because the public places value on visiting Darling Harbour for different purposes, whether it be to promenade, play in the park, sit

down and eat food, and so on. Duration is also important because visitors want to be able to choose how long they wish to use the public spaces each time they visit.

#2: Shared space alongside promenades

The perceived property rights here were those identified in an interview with the proprietor of a café with *al fresco* dining, in a space shared with the general public. The dominance of direct use and access rights (Table 4) mirrors the high priority given to the same rights by the general public. The café proprietor saw direct use rights being significant because of the *al fresco* area's thoroughfare location. Access for staff and customers was seen as important for his business. Duration was also significant because the proprietor considered it important to feel secure in holding the lease for a reasonable period of time. The meaning of duration in this context differs from the duration of access enjoyed by members of the public in category #1. This highlights the challenging semantics of property rights descriptors.

#3: Public garden with an entry fee (Chinese Garden)

This space embodied several types of perceived property rights of relatively similar significance (Table 4). The three most important rights were seen as being Control, Quality of Title, and Management. Complete control of the enclosed garden space was regarded as necessary to retain the Garden's public image. Quality of Title was seen as an important asset that ensured the preservation of the educational and cultural benefits of the gardens. Management rights were also regarded as significant, reliant on the security that these rights provided. This is because of the oversight of the Sydney Harbour Foreshore Authority ensuing from management reporting requirements. Residual rights were also important. There was a sense of defensible space, by virtue of the existence of a wall around the gardens.

#4: Public spectacle leased space with entry fee (Maritime Museum and Sydney Aquarium)

For this type of space, Control was the most important perceived property right. Direct Use, Indirect Economic Gain, Access and Management rights were also seen as being significant (Table 4). Control rights in this space are given effect through the leases, which confer a legally binding status on these rights. It also provides them with a high-level of asset security. Management rights are seen as being closely related to Control rights. Direct Use rights are considered important because these allow the Maritime Museum and Aquarium operators to run the venues as businesses and attract admissions. Indirect Use rights are important because of the extra revenue these can generate, through sub-leases of office space and through spending in retail outlets at the venues. One of the two operators here saw Access rights as being very important because the activity would be impossible to carry out in the absence of these.

#5: Public space in private facilities (Casino)

Exclusivity is the most important perceived property right. Duration and Control rights are also regarded as being important (Table 4). The significance of Exclusivity rights here derives from the Casino's operating rights on the site. The state government guarantees exclusivity (in return for the licence premium), as this is the only casino allowed in the state. Duration rights stem from the length of time to which this exclusive right to operate as a casino applies. This is an intrinsic element of the commercial value of the combined rights of Exclusivity and Duration, and hence its significance perceived by the Casino. The ability to Control access and use of the Casino in order to maximise revenue is fundamental to the investment value that the Casino can extract from the site, and is thus a central component of the property rights the Casino sees as significant.

#6: Privatised commons where entry fee charged (Exhibition and Convention Centre)

In this space, Indirect Economic Gain is the dominant perceived right (Table 4). This arises from the perspective of the Sydney Harbour Foreshore Authority (as they control the Centre's management company), which goes beyond the Centre's conference and exhibition revenue. This perspective incorporates wider metropolitan economic activity generated by the Centre, such as accommodation and associated spending. Other rights for this space also considered important are Divisibility, Direct Use, Quality of Title, Control, and Management. Divisibility is important because a core function of the Centre is the ability to hire space within the Centre on behalf of conference organisers and exhibitors. Quality of Title, Control, and Management rights are seen as interlinked. The first stems from SHFA's ownership of the Centre, which gives SHFA control of the Centre and the right to manage it. In turn, these rights allow SHFA to have direct use of the Centre, which is seen as important because it allows SHFA to extract revenue from this use.

#7: Public access spaces in private owned space for employment

For these spaces, rights of Duration, Direct Use and Indirect Economic Gain are significant. The respondent selected to represent this space operated a marine tourism business. Direct Use rights are perceived as essential by providing premises from which to run the business, while rights of Duration are also important in giving Direct Use rights a degree of security via a lease. Indirect Economic rights are also important because the business uses sales agents to market its services who are located beyond Darling Harbour.

10 General Trends

It is clear from the above analysis that certain typologies of property rights are not significant in the contemporary urban commons of Darling Harbour. The most obvious of these are Alienation, Exclusion, and Transfer. This is understandable, because the commons are not about alienating, excluding or transferring – these are not property

rights that impact in the commons. Property rights in the commons need to be interpreted in a communal way, compared to the private property rights of an individual. We therefore need adapt the summary of property rights that prevail in the literature to a more specific list that is more appropriate to the analysis of property rights in the contemporary urban commons. We therefore contend that only those property rights that apply in the commons should be considered in analysing other urban parkland.

Whilst we excluded Alienation, Exclusion, Management, Quality of Title, and Transfer in the analysis of the Public Spaces (category #1), the Management, Quality of Title typologies remain relevant in the other six categories of our mosaic.

11 Limitations

The way in which the ranking of property rights typologies are identified by the Analytic Hierarchy Process within the Expert Choice software highlights potential inconsistency between the ranking of different rights due to intransitivity. The resulting level of inconsistency for each mosaic category is shown in Table 4. Inconsistency levels of 10% or less are not regarded as being significant. The objective is to make ‘good’ decisions rather than attempt to minimise the inconsistency ratio (EC2000, 2002). What is important is to maintain coherence among the objects of our experience, in this case members of the public analysing hierarchies of property rights that may demonstrate inconsistency amidst lawyers or property professionals. The only mosaic category with an inconsistency of <10% is that of Public Space. In this category, consistency arises both from the reduced number of property rights typologies that were analysed (10 cf. 15) and from the larger number of respondents.

For all other categories the inconsistency level was significant (ranging from 16% to 22%). In an analysis with 15 criteria a level of inconsistency is to be expected, as the human brain has can have difficulty comparing more than nine components. We could have reduced the criteria to 9 or 10 (as we did for the Public space category #1). However, in accordance with the spirit of the research instrument we did not want to moderate or limit the number of categories derived from the literature. Likewise, for the purpose of this data collection we did not want to stifle the potentiality of respondents to manage multiple ideas. We elaborate on the impact of this below.

The Public Garden (category #3) demonstrated an inconsistency of 21%, because the respondent prioritised Exclusivity (a low ranked right) over Direct Use (also low ranked). Similarly, confusion over pairwise comparison between the other low ranked rights typologies also contributed significantly to the overall inconsistency value. The lack of perception about the nature of these low ranked rights could be suggested as a major reason for the inconsistent ranking of these rights.

Public Spectacles Space (category #4) had an inconsistency level of 16%. This is lower than the other non Public Space categories, because there were two respondents compared to one in categories 2, 3, 5, 6 and 7, with the hybrid output demonstrating a more balance perception.

In the Restricted commons (in employment spaces) category (#7), the major inconsistency arose from the rankings of Direct Use and Indirect Economic Gain. Both emerge as being of similar importance, but the pairwise comparison between the two significantly prioritised Direct Use (extreme importance) over Indirect Economic Gain. Whilst we could speculate that this relates to the owners perception the importance of his companies control at Darling Harbour over the external sales agents influence, such interpretation is beyond the scope of this study.

12 Conclusion

We have engaged a novel approach to interpret the contemporary urban commons. We have provided distinct conceptual frameworks (from the literature) and contextual frameworks (from locating the analysis in Darling Harbour, Sydney). As we demonstrated, notions of property rights are confused and contested within the literature. The urban commons literature highlights that access to urban parkland is important for the well-being of city dwellers. Our analysis reinforced the importance of access to public space to relax and recreate from the perspective of our general public respondents. The public space in urban commons prioritises rights of access, from the perspective of both jural opposites and jural correlatives.

The seven categories that comprise our mosaic of property rights are closely aligned to the third layer of social relations of property (per von Benda-Beckmann et al., 2006). Our mosaic categorisation was robust enough to be tested in the context of Darling Harbour. It will be interesting to see if the categorisations render similar responses in other contemporary urban commons, which are spaces of interaction that accommodate and expand the norms and values in a multicultural city like Sydney. The political, legal and indeed economic relationships between people, place and property influenced the responses for our categories #2 through #7. By analysing all four layers, we were able to consider within our approach the ‘concretised’ layer of practice, and how users relate to their engagement with space.

13 Acknowledgements

The authors acknowledge the funding support of the RICS Education Trust (project #349), which combined with the assistance of the UTS: Asia-Pacific Centre for Complex Real Property Rights enabled this research to be undertaken. We also thank the Sydney Harbour Foreshore Authority for their interest and support in this project. Dr Garrick Small was a member of the research team in the formative stage of the project, and his collaboration in the development of the mosaic was invaluable. We also thank Kate Fairlie, a doctoral scholar with the UTS: APCCRPR, who was the research assistant for the project.

14 References

- BOLLIER, D. (2008) *Viral Spiral : How the Commoners Built a Digital Republic of Their Own*, New York, The New Press.
- BOYDELL, S. (2007) Myth, Embeddedness and Tradition: Property Rights Perceptions from the Pacific. IN SMAJGL, A. & LARSON, S. (Eds.) *Sustainable Resource Use: Institutional Dynamics and Economics*. London, Earthscan.
- BOYDELL, S., SEARLE, G. & SMALL, G. (2007) The Contemporary Commons: understanding competing property rights. IN HAMNETT, S. (Ed.) *State of Australian Cities Conference (SOAC 2007) - Growth, Sustainability and Vulnerability of Urban Australia*. Adelaide, South Australia, SOAC/UniSA.
- BROMLEY, D. W. (1991) *Environment and Economy: Property Rights and Public Policy*, Cambridge, Ma., Basil Blackwell, Inc.
- COASE, R. (1960) The Problem of Social Cost. *Journal of Law and Economics*, 15.
- COLE, D. H. & GROSSMAN, P. Z. (2002) The Meaning of Property Rights: Law versus Economics? *Land Economics*, 78, 317-330.
- CROCOMBE, R. G. (1975) An Approach to the Analysis of Land Tenure Systems. IN LUNDGAARDE, H. P. (Ed.) *Land Tenure in Oceania*. Honolulu, University of Hawaii Press.
- DE SOTO, H. (2000) *The mystery of capital: why capitalism triumphs in the West and fails everywhere else*, New York, Basic Books.
- DEMSETZ, H. (1967) Toward a theory of property rights. *American Economic Review*, 57, 347-359.
- EC2000 (2002) Expert Choice 2000 - Advanced Decision Support (AHP) Software.
- FARRAN, S. & PATERSON, D. E. (2004) *South Pacific Property Law*, Sydney, Cavendish Publishing Limited.
- FLYVBJERG, B. (1998) *Rationality and Power*, Chicago, University of Chicago Press.
- GLASZE, G., WEBSTER, C. & FRANTZ, K. (2006) *Private Cities: Global and local perspectives*, London, Routledge.
- HANN, C. M. (2007) A new double movement? Anthropological perspectives on property in the age of neoliberalism. *Socio-Economic Review*, 5, 287-318.
- HARDIN, G. (1968) The Tragedy of the Commons. *Science*, 162, pp.1243-1248.
- HOFFIELD, W. N. (1913-1914) Some Fundamental Legal Conceptions as Applied in Judicial Reasoning. *Yale Law Journal*, 23, 16-59.
- HOFFIELD, W. N. (1916-1917) Fundamental Legal Conceptions as Applied in Judicial Reasoning. *Yale Law Journal*, 26, 710-770.
- KOHN, M. (2004) *Brave New Neighborhoods : The Privatization of Public Space*, New York, Routledge.
- LICHFIELD, N. (1996) *Community Impact Evaluation*, London, UCL Press.
- LYONS, K., DAVIES, K. & COTTRELL, E. (2007) The Need to Consider the Administration of Property Rights and Restrictions Before Creating Them. IN SMAJGL, A. & LARSON, S. (Eds.) *Sustainable Resource Use: Institutional Dynamics and Economics*. London, Earthscan.
- MAINE, H. S. (1861) *Ancient Law: its Connection with the Early History of Society, and its Relation to Modern Ideas*, New York, Charles Scribner.
- MAX-NEEF, M. A. (2005) Foundations of transdisciplinarity. *Ecological Economics*, 53, 5-16.

- MCEWIN, R. I. (1993) Legal Versus Economic Perceptions of Real Property Rights. *The Australian Economic Review*, 3rd Quarter 1993, 35-40.
- MCGINNIS, K. (2001) Great Commons of the World. *yes!* Bainbridge Island, WA, Positive Futures Network.
- MCSHANE, I. (2010) Trojan Horse or Adaptive Institutions? Some Reflections on Urban Commons in Australia. *Urban Policy and Research*, 28, 101 - 116.
- NICOLESCU, B. (2006) Transdisciplinarity - past, present and future. IN HAVERKORT, B. & REIJNTJES, C. (Eds.) *Moving Worldviews: Reshaping sciences, policies and practices for endogenous sustainable development*. Leusden, ETC/COMPAS.
- NORTH, D. C. (1990) *Institutions, Institutional Change and Economic Performance*, Cambridge, Cambridge University Press.
- PAYNE, G. (1997) *Urban Land Tenure and Property Rights in Developing Countries: A Review*, London, Intermediate Technology Publications/Overseas Development Administration (ODA).
- RIFKIN, J. (2000) *The Age of Access : The New Culture of Hypercapitalism, where all of life is a Paid-for Experience.*, New York, Jeremy P. Tarcher / Putnam.
- RIGSBY, B. (1998) A Survey of Property Theory and Tenure Types. IN PETERSON, N. & RIGSBY, B. (Eds.) *Marine Tenure in Australia: Oceania Monograph 48*. Sydney, University of Sydney.
- SAATY, T. L. (1988) *Multi-criteria Decision Making: The Analytical Hierarchy Process*, Pittsburgh, PA, RWS Publications.
- SHEEHAN, J. & SMALL, G. (2002) Towards a Definition of Property Rights. *Working Paper 01.02*. Sydney, UTS Property Research Unit.
- VON BENDA-BECKMANN, F., VON BENDA-BECKMANN, K. & WIBER, M. G. (2006) The Properties of Property. IN VON BENDA-BECKMANN, F., VON BENDA-BECKMANN, K. & WIBER, M. G. (Eds.) *Changing Properties of Property*. New York, Berghahn Books.
- WEBSTER, C. J. & LAI, L. W.-C. (2003) *Property rights, planning, and markets : managing spontaneous cities*, Northampton, MA, Edward Elgar Pub.
- WORLD BANK (2003) Land Policies for Growth and Poverty Reduction. IN DEININGER, K. (Ed.) *World Bank Policy Research Report*. Washington, Copublication of the World Bank and Oxford University Press.

Compensation rights for land use regulation: the findings of cross-national comparative research

Rachelle Alterman

Technion – Israel Institute of Technology

alterman@technion.ac.il

Abstract

In every country where land-use regulations and development controls operate, they often cause changes in the economic value of real property. Where the effect is downwards, do landowners have a right to claim compensation from the planning authorities? This topic addresses an inherent "raw nerve" of planning law and practice, bearing deep economic, social and ethical implications. Yet, this paper will report on the first large-scale systematic comparative research ever conducted on this topic. This full research findings are presented in my recently published book (Alterman 2010)¹ titled *Takings International*. Because the book was published in the USA I used American terminology – "regulatory takings" for what is better known in other countries "injurious planning decisions" or similar terms.

Thirteen advanced-economy countries were selected for analysis – comprising over a third of all OECD member countries. To enable systematic comparison, I prepared a framework of questions to serve as common benchmarks for the experts from each country that I invited to participate in the project. The framework includes: Legal history, rationale; constitutional law, statutory framework, types of compensable injurious decisions (if any), distinction between physical takings, regulatory full takings and partial takings; between direct and indirect impacts; extent of compensation due; procedures, incidence and possible impact on planning practice.

The findings show a surprising high degree of variation among countries – including those with common legal traditions. The 13 countries are grouped into three clusters of countries: Minimal compensation right, moderate or ambiguous rights, and broad compensation rights. The findings are often surprising and counter-intuitive. They shed new light on the laws and practices of each of the participating countries. The most dramatic differences are regarding "partial takings". The

¹ Rachelle Alterman *et al* (2010). Takings International: A Comparative Perspective on Land Use Regulations and Compensation Rights, Chicago: American Bar Association Press.

comparative analysis highlights the opportunities for cross-national learning regarding this fundamental issue.

Introduction

One might have thought that the law of regulatory takings would be a prime topic for cross-national research. Yet there has been little comparative research on this topic and few opportunities for knowledge exchange. This research project is, to the best of our knowledge², the first large-scale comparative research devoted entirely to regulatory takings. The full survey of literature is beyond the scope of this paper (see Alterman, 2010: chapter 1).

Most of the contributions were published in the USA. Considering Europe's quest for a "single market" and the importance of the free movement of capital – including real estate investments - one would have expected that European scholars would study the similarities and differences in regulatory takings and compensation laws across Europe. Yet, as surprising as this may seem, there has been very little comparative research on regulatory takings in Europe³. The major comparative project on planning law and institutions commissioned by the European Commission (1997-2000) does not cover regulatory takings law at any depth⁴.

The purpose of the book on which this paper is based was to tap the universality of the regulatory takings issue and reduce some of the insularity and compartmentalization that has characterized the issue so far. The book offers readers from various parts of the world an initial systematic cross-national vantage point from which to view their own country's debate. Such a perspective may be able to provide a sense of scale not available when one is imbued in a single country's laws, policies, and debates. Because nine of the thirteen countries covered in this research are members of the European Union, an ancillary objective is to stimulate broader cross-national debate within the EU.

1. Scope of the research

I asked each contributing authors to address the following main question:

Under your country's laws, do landowners have the right to claim compensation when a government decision related to planning, zoning or development control causes a reduction in property values? If

² The survey of literature covers publications in the English language and partially in French as well.

³ This assessment is supported by the 13 European authors who participated in this project, who cover a variety of languages. The French author, Vincent Renard, brought my attention to a French language edited collection of short papers on "servitudes d'urbanisme" which, in French, also refers to regulatory takings. See: special issue of *Droit et ville* no. 48/1999 titled: "L'indemnisation des servitudes d'urbanisme en Europe".

⁴ European Union, European Commission (1997, 2000), The EU Compendium of Spatial Planning Systems and Policies, Luxembourg, Office for Official Publications of the European Commission. Schmidt-Eischstaedt's 1995 book is another important comparative analysis of planning laws, but it too does not cover regulatory takings. See: GERD SCHMIDT-EICHSTAEDT, LAND USE PLANNING AND BUILDING PERMISSION IN THE EUROPEAN UNION (Deutscher Gemeindeverlag Verlag W.Kohlhammer, 1995) (in English and German) .

so, what are the legal and factual conditions that a landowner must meet to claim compensation? And how extensive are such claims in practice?

Injuries to property values caused by land-use regulations may fall along a continuum – from no injury at all to a reduction of all or most of the property's value. This entire range is within the scope of this research, but it does not encompass every possible topic related to regulatory takings. The focus is on regulations related to land use planning, not on all types of regulations that may affect property values. The limitation accords with the core subject of regulatory takings law as historically viewed in most countries. Thus, some types of environmental regulations that are not part of land use planning are not systematically covered.

It is important to distinguish the right to compensation for injurious land use regulations from the right to compensation for land taken through eminent domain ("*expropriation*" or "*compulsory purchase*"). In such cases, the ownership rights are compulsorily transferred to an authorized body. Eminent domain does not fall within the direct scope of this research project. However, because expropriation law and regulatory takings law are related to some extent in many countries, the authors of most of the country chapters present a brief introduction to expropriation law in order to link into regulatory takings law⁵.

Another topic not covered here is exactions of land or other amenities through regulation. The reason is that unlike in the USA, in some of the countries analysed this topic is not regarded as part of regulatory takings law. It is the subject of special regulations on topics such as compulsory dedications of land for roads, public buildings, public passage, and – increasingly – environmental mitigation. This topic too undoubtedly merits comparative research⁶.

2. The jurisdictions selected for study

Thirteen countries were chosen for comparative analysis. This is a large number and makes the analysis quite challenging. At this stage of knowledge-buildup, a large sample of countries was necessary in order to dispel assumptions that there is a single dominant approach outside the USA or that there is a "European approach" versus an "American approach" (more on this later). My objective was also to open the cross-national discussion to participants in many countries around the world.

In selecting the set of countries, I did not have the privilege of relying on prior theory from which to derive a typology of regulatory takings laws. The countries chosen should of course not be regarded as a random sample that represents all the world's countries. Recognizing the pioneering character of this research, I sought to span a wide spectrum of countries based on several variables that could

⁵ In many countries represented here, including the USA, the distinction between regulation and de-facto expropriation is not always "a bright line". Situations of "near-expropriation" do occur. Called by various terms in various countries, including "inverse condemnation" and "planning blight", and types of takings *are* included in the scope of this project. See my analysis in Chapter 2 *infra*.

⁶ A comparative analysis of exactions in a few countries is included in my book PRIVATE SUPPLY OF PUBLIC SERVICES: EVALUATION OF REAL ESTATE EXACTIONS, LINKAGE, AND ALTERNATIVE LAND POLICIES (1988).

reasonably be hypothesized as related to differences in approaches to regulatory takings.⁷ The set of countries chosen for analysis therefore includes jurisdictions from the two major Western legal traditions – common law and civil law⁸ –, both unitary and federal jurisdictions, and a broad geographic spread. In view of the important legal and ideological changes occurring in the post-socialist countries, I wanted to include at least one such jurisdiction in the sample. The common denominator is that all countries chosen share a democratic system of government and have an advanced (or fast-emerging) economy. In most countries that do not meet these two criteria, planning laws are often irrelevant (due to corruption or widespread noncompliance), and regulatory takings law is either dormant (no claims filed) or nonexistent. Because Europe encompasses a large portion of the world's democratic, advanced-economy countries, the sample selected includes many European countries. My assumption is that the set of countries chosen does represent the majority of legal approaches among the world's democratic countries with advanced economies.

The thirteen countries chosen for research are (from west to east and south): the USA, Canada, the UK, France, the Netherlands, Sweden, Finland, Germany, Austria, Poland, Greece, Israel and Australia. These countries are located on four continents. Nine are members of the European Union. The set includes five federal countries – the USA, Canada, Germany, Austria and Australia. There are variations among the federal jurisdictions in the degree of importance of the sub-federal law on regulatory takings. In the USA, Canada, and Germany, the federal level is the most important for regulatory takings law. Therefore, the analysis of the USA, Canada and Germany is on the federal level. Austria does not have any overarching federal law on regulatory takings – neither a significant body of constitutional case law nor statutory law. Because each of the nine states in this small country has its own statutory law and they differ significantly from each other. Finally, in Australia, in addition to federal-level constitutional law, state statutes are very important in takings law.

The set of countries chosen represent the two major legal "families". Five share the common-law tradition: the UK, Canada, Australia, the USA, and Israel⁹. The remaining eight countries – all located on the European continent – are civil-law countries. My assumption was that in the field of planning law the two families of countries have enough in common to enable comparative analysis. The findings corroborate this assumption and show that the differences within these families of laws are not lesser than between the two groups.

⁷ I of course limited my search to countries where I had located suitable scholars in the planning law field. See "Research Method" below.

⁸ For a general discussion of legal tradition in comparative research see: MARY ANN GLENDON, PAOLO G. CAROZZA AND COLIN B. PICKER, *COMPARATIVE LEGAL TRADITIONS: TEXT, MATERIALS AND CASES OF WESTERN LAW*, 23-49 (Thomson-West – American Casebook Series. 2007)..

⁹ Israel is regarded by comparative-law scholars as a mixed system, but with strong attributes of the common law tradition. On the one hand, precedence is a major source of the law and common law is still prominent in a few areas, on the other hand, statutory law is dominant in most fields of law today. See GLENDON, CAROZZA & PICKER, *supra* note 52, at 948 (about mixed jurisdictions in general) and 976-982 (about Israel). In the field of planning law, the British influence dates back to the British Mandate on Palestine from 1921 to 1948. The planning law is a direct derivative of legislation enacted during that time.

3. The research method

Analysis of the laws of thirteen countries in many languages is beyond the capacity of a single researcher. To carry out this research project, for each of the candidate countries, it was necessary to build a team. I sought out a leading expert (or experts) on planning law who are well versed in English. All seventeen contributing authors are highly regarded scholars in their country. The contributing authors were asked to provide an in-depth analysis in English of their country's laws and practices on regulatory takings according to a set of guidelines that I developed (described below).

The challenge of bringing the parallel analysis of all the countries onto a common platform was not easy. The details of the takings law in each country are complex and nuanced, and require in-depth knowledge of each country's law, jurisprudence and practices. Often, what a particular author assumed to be similar in other countries and therefore easily understandable to readers from other countries was in fact quite opaque. I worked with the authors to provide enough contextual information so that readers from other countries would understand the implications of a particular law or institution. There were also language barriers: in every country in the set, court decisions on land use law are delivered in the local language only. No non-English speaking country in our sample offers translations into English of court decisions in the planning area. In some countries even the statutes have not been translated.

To enable comparative analysis, I developed a common framework and a set of guidelines for the entire team. Prior to developing the guidelines, I read the literature available (in English) on land use law and practice in each of the countries. Next, I conducted a set of preliminary interviews with each of the prospective authors. To ensure that each of the participating authors would understand the guidelines in a similar way, I prepared a preliminary set of scenarios composed of types of regulations, types of potential injuries to property values, and types of contextual conditions. Through a "rolling" strategy I expanded or refined the scenarios and guidelines until we were satisfied that these would encompass all the laws and practices in the sample countries. The effort of coordinating and editing the set of papers nevertheless proved to be a major challenge. In most cases, the papers underwent several rounds of editing by me and rewriting by the authors, supplemented by clarifications through personal meetings, phone or email communication.

The comparative analysis was carried out by me in consultation with the contributing authors. To digest all thirteen reports, I first "dissected" them back according to the topics of the guidelines, appropriately revised to reflect the findings and supplemented by new "variables" that emerged from the analysis. From among these, I selected four major topics to serve as the main anchors for the comparative analysis. These are: constitutional law, major ("categorical") takings, direct partial takings, and indirect partial takings. These topics are explained in Chapter 2 of Alterman (2010). These four headings were selected because they could capture the main similarities and differences, but they by no means cover all aspects of regulatory takings law. There are many minor topics that remain for the readers to review in each country chapter. Once the comparative analysis was finished, I sent each contributed author the relevant parts for his or her comments.

4. The comparative findings in a nutshell

The comparative findings are analysed in depth in the book (Alterman, 2010: chapter 2). Here I shall present some of the highlights.

4.1 The scale of degree of compensation rights

The findings show that there is no universally consensual approach, nor even a dominant approach. Different countries at different times have adopted varying approaches to dealing with the property-values dilemmas. The diversity is great: no two countries have the same law on regulatory takings – not even countries with ostensibly similar legal and administrative traditions. The differences among the countries are significant, and often unpredictable on the basis of other attributes known about these countries.

If one imagines a hypothetical scale of degrees of compensation rights, the countries in the sample seem to represent as full a spectrum of compensation rights as one can conceive (of course, there may be more variations among other countries not represented here). At one extreme are two countries whose laws and practices have said a clear "yes" to compensation rights for a broad range of "downzoning" situations, in many ways rivaling Oregon's Measure 37. At the other extreme are a few countries whose legal doctrine does not grant compensation rights for any types of regulatory takings (except some extreme situations). In between are several countries whose laws grant compensation rights for regulatory takings, but these are limited in time or in subject. The matrix of specific grounds and conditions is rather complex. On the whole, each country's set of laws and policies differs significantly from every other's equivalent set. There are striking differences even among the nine countries that are members of the EU, despite the fact that they all also fall under the jurisprudence of the European Court of Human Rights.

Perhaps the most interesting finding – one that is difficult to accept - is that any attempt to guess a given country's position on regulatory takings law based on some other attributes that characterize it, is likely to fail. Careful reading of the full set of papers shows that presumptions or hypotheses based on legal "family", legal-institutional structure, geographic proximity, shared language and cultural backgrounds, or planning needs (such as density, size or resources scarcity), do not hold in the face of the comparative findings.

It is our hope that the wide variety of laws and practices identified in this team project on comparative research will enable the readers to learn from other countries' experiences and thus to gain a new perspective on their own country's laws and policies.

4.2 Viewing US regulatory takings law from the outside

When viewed from a cross-national perspective, the most striking finding about US regulatory takings law is the glaring disparity between the intensity of the US "property rights debate" and the factual positioning of US takings law midway along the "scale" of degree of compensation rights for regulatory takings. The comparative perspective can offer both sides in the debate a way of thinking "out of the box". Both the proponents and the opponents of "property rights" may gain by looking at US takings law from the outside. Both sides can also look to other countries, either for alternative

models to support their own position (with appropriate adjustments) or for middle-of-the road approaches that may contribute to a rapprochement in this long-raging contest.

Alongside its moderate degree of protection of property rights, US law on regulatory takings exhibits some unique characteristics that place the US apart from other countries.

Perhaps the most prominent feature is the intensity of the property-rights debate itself.¹⁰ In no other country in the sample, and at no time in recent decades, has the issue of regulatory takings occupied a similarly prominent position in public opinion and action as in the USA. In no other country has the issue of regulatory takings become a major topic in national (or state) elections. In no other country has public opinion led to a legislative saga such as Oregon's extremist Measure 37 and its quick demise within only three years. Interestingly, the relatively docile status of the takings issue in most other countries exists regardless of the position occupied by that country's takings laws on the compensation rights scale: whether on the very restrictive side regarding compensation rights (Canada, Australia, the UK, France or Greece); or on the side that offers broad rights (Poland, Germany, Sweden, the Netherlands, and Israel).

Another key difference between the US and the other countries is the extremely prominent role played by constitutional law. In most other jurisdictions in this study, statutory law (whether on the national or sub-national levels) is a key "player" in takings law. Only in the USA is takings law decided largely by direct application of the Constitution (in most states where there are "regulatory takings statutes", these laws do not add substantive causes of action beyond constitutional law)¹¹.

In interpreting the constitution, the US Supreme Court has refrained from making "bright line" rules, leaving many legal issues to be decided through case-by-case determination¹². The result is that US takings law is characterized by a high degree of uncertainty that both landowners and government agencies face whenever regulatory takings are challenged in the courts¹³. In two more countries in the sample - Finland and Austria - a high degree of legal uncertainty still prevails. However, in these two countries, the reason for the uncertainty is that there have been few claims and no jurisprudence to interpret the language of the statute. American society is much more litigious. The unique feature of

¹⁰ For a discussion of the debate and literature sources see Chapter 1.

¹¹ Most of the state statutes only required government agencies to conduct a takings assessment prior to adopting a regulation, or to institute conflict resolution measures. The only exceptions are Florida and Oregon - and the record there too is disappointing for the proponents of property rights. See: Stacy M. White, *State Property Rights Laws: Recent Impacts and Future Implications*, LAND USE LAW AND ZONING DIGEST (July: 3-9, 2000).; Hannah Jacobs (2007); and JOHN D. ECHEVERRIA AND THEKLA HANSEN-YOUNG. *THE TRACK RECORD ON TAKINGS LEGISLATION: LESSONS FROM DEMOCRACY'S LABORATORIES*. 1-2 (Georgetown Environmental Law & Policy Institute, 2008). Regarding Florida see: Joni Armstrong Coffey *High Hopes, Hollow Harvest: State Remedies for Partial Regulatory Taking*, 39(3). THE URBAN LAWYER 619-632 (2007).

¹² See the US chapter by Roberts.

¹³ Many American authors make a similar point regarding insufficient clarity and inconsistencies. See for example: DAVID L. CALLIES, ROBERT H. FREILICH AND THOMAS E ROBERTS. *CASES AND MATERIALS ON LAND USE* 380 (Fourth Edition; 2004) (henceforth "Callies, Freilich and Roberts"). See also: Edward J. Sullivan and Kelly D. Connor, *Making the Continent Safe for Investors - NAFTA and the Takings Clause of the Fifth Amendment*, Chapter 4, pp. 47-83 in *CURRENT TRENDS AND PRACTICAL STRATEGIES IN LAND USE LAW AND ZONING* (Patricia E. Salkin, ed., American Bar Association, 2004). At p. 67 the authors argue that the degree of certainty and uniformity intended by the Federal Constitution has not been accomplished in the field of takings law.

US takings law is that high legal uncertainty persists despite a large body of Supreme Court jurisprudence extending over almost nine decades, alongside many decisions by the lower courts.

The combination of intensive public debate, a constitutional focus and considerable legal uncertainty has produced what is by far the largest body of scholarly research and publications on regulatory takings anywhere in the world. the Lexis Nexis search I conducted using the terms "regulatory takings" together with "land use" yielded a number of items beyond what the display would carry. Likely, Thousands of scholarly papers and hundreds of books discuss the "takings issue". Every new Supreme Court decision generates scores, sometimes hundreds of scholarly publications. Yet this huge body of knowledge – which I hypothesize, is several times larger than all the scholarly writing on the topic in all other countries and languages combined – has not produced greater convergence on the property rights debate. The Oregon turbulence of 2004-2007, which left no side satisfied, is unlikely to be the last time legislative initiatives are brought up.

4.3 The Mutual Images of Americans and Europeans

In academic or professional discussions – and even in some academic publications – one sometimes encounters Americans who refer to the "European approach" to regulatory takings, as contrasted with the "American approach". The image is that Europe has a unitary approach that offers less protection of property rights than the USA. There seems to be an image that European countries as a group do not recognize the concept of regulatory takings or do not regard these as entailing compensation by the public purse. On the opposite side, one encounters non-American practitioners and scholars – even from English-speaking countries – whose image of US takings law is that it offers landowners extensive protection from downzoning and generous compensation rights¹⁴. This view is part of a broader image of the USA as offering extensive protection of property in general.

The evidence from the thirteen-country study shows that both images are far from correct. There is no "European approach" to regulatory takings. The nine European countries exhibit the full scale of legal (and public-policy) approaches to regulatory takings, almost to the very extremes. This is so even though all European countries are bound by the European Convention on Human Rights and Fundamental Freedoms (ECHR). Article 1 of the First Protocol of ECHR provides for property protection, but qualifies it "with the general interest"¹⁵. Furthermore, all European countries analysed are also members of the EU. Yet the laws and practices of the nine European countries differ so greatly from each other that a "Euro-blind" reader may not have guessed their joint affiliation.

The canopy of the ECHR has to date shown high tolerance for the variety of interpretations of both property rights and the public interest. The effect of ECHR jurisprudence so far has been modest: It

¹⁴ Thomas Roberts, the author of the US chapter, is well aware of this false image, and points it out.

¹⁵ See also the discussion in Chapter 2. Protocol 1, Article 1 says: "Every natural or legal person is entitled to the peaceful enjoyment of his possessions. No one shall be deprived of his possessions except in the public interest and subject to the conditions provided for by law and by the general principles of international law. The preceding provisions shall not, in any way, impair the right of a State to enforce such laws as it deems necessary to control the use of property in accordance with the general interest or to secure the payment of taxes or other contributions or penalties." Protocol 1 to the Convention for the Protection of Human Rights and Fundamental Freedoms as amended by Protocol No. 11, art. 1, Mar. 20, 1952, Europe. T.S. No. 9, available at <http://conventions.coe.int/treaty/en/Treaties/Html/009.htm>.

has pared down only the extremities on the non-compensation side, but has not influenced the countries whose laws fall on the other side of the scale with extensive compensation rights. The European countries at the very extreme edges of the scale are France and the Netherlands. On the one extreme one finds France where the planning statute says explicitly that land use regulation is not to be compensated, and at the other extreme is the Netherlands, whose law grants compensation even for regulatory takings of minute degree and of indirect effect. The other seven European countries – each with a different set of laws on regulatory takings – fall in-between on the scale: some are closer to France (the UK and Greece), others are closer to the Netherlands (Poland, Germany and Sweden), and two have unclear positions (Finland and Austria). In total, four of the nine European states offer landowners higher degrees of protection of property than US law (though not uniformly in all factual situations); two offer lesser rights; and regarding the last two, there is not enough jurisprudence to know.

The comparative findings also show that there is no unitary "British approach" to contrast with the US approach. The four countries with British law in their background in the past century – the UK, Canada, Australia and Israel – span the two extremes on takings law: Canada on one side (extremely restrictive) and Israel on the other (excessive compensation rights). Today, there are not many similarities among these countries' laws on takings.

US takings law holds a middle seat both on major takings (called "categorical" in the USA) and on partial takings. On major takings, US law is in line with the majority of countries analysed (with a few exceptions to be noted below): it recognizes a taking when the economic value of a property has been totally or almost totally extinguished. But to win a categorical-taking challenge, US jurisprudence has set a number of conditions that are difficult to satisfy. On partial takings too, US law is mid-scale. It joins half of the set of countries where partial injuries are compensable. But unlike its image overseas, US law (except for some state-law exceptions) places a rather high quantitative threshold as well as various preconditions that make it difficult for American landowner to win challenges for partial takings. The third type of taking – indirect injuries – are not recognized in US law, thus placing US law among the majority of countries; only a few countries grant extensive compensation rights for such injuries.

At the same time, the comparative analysis also indicates that on several important counts, US regulatory takings law is more generous to landowners than the laws of most other countries in the set. First, in all the other countries in our set, takings claims can only be made when a government body *changes* an existing regulation to a *more restrictive* category. Owners of farmland or even vacant land that generates no income usually do not have the right to expect a rezoning. The US is the only country among the set where *refusals to upzone* or to grant development permission can conceivably serve as grounds for a taking challenge¹⁶. Although a challenge on these grounds is difficult to win, the threat of facing one lurks in the background when US policymakers decide on farmland and other environmental protection regulations. Second, in most countries (with few exceptions), regulatory takings – especially partial takings – are not an open-ended concept; a statute usually defines a limited set of government decisions that may entail compensation. The historic as well as the current core of compensable decisions in most countries revolves around "classic" land-use planning and zoning (not

¹⁶ See *Tahoe-Sierra Preservation Council, Inc. v. Tahoe Regional Planning Agency* 535 U.S. 302, 122 S.Ct. 1465, 152 L.Ed.2d 517 (2002).

even all types of potentially injurious decisions are necessarily included). For example, some environmental regulations may not be compensable.

5. Possible statutory models for cross-learning

Proponents of the no-compensation doctrine can find an assortment of approaches among the countries surveyed. The cluster of countries on the no-compensation side of the spectrum includes Canada, Australia, the UK, France, and Greece. France and Greece, however, would probably not be suitable models. Greece is unsuitable because its law on regulatory takings lacks internal consistency, and poor administrative practices have made them dysfunctional. France may be unsuitable because its no-compensation ideology is too extreme to withstand US constitutional challenges (the French planning statute disallows payment of compensation for any land use regulation).

Canada, next door, presents at the federal level, another extreme no-compensation doctrine that is much at odds with US constitutional protection of property. However, some of the Canadian provinces have enacted more moderate statutes or administrative practices. These may well merit further evaluation. Australia is somewhat less extreme in its no-compensation doctrine. Unlike Canada's provinces, the Australian state statutes are more consistent in granting compensation rights for major ("categorical") takings. Especially interesting would be to look at the differences among the states and evaluate their legal and public impacts. In 2007-8 some Australian states¹⁷ began a reassessment of their regulatory takings laws and the outcomes are worthy of follow-up.

In my view, on the no-compensation side, the UK is the most interesting model (see also Mandelker's Afterword). UK law is the most coherent on the no-compensation side, where the pieces fit together into a consistent whole. The UK system is well worth further study by those who seek a legal system where there would be minimal companion rights, yet where landowners would have a reasonable degree of protection in extreme situations. In the UK this balance is achieved with a greater degree of legal certainty than offered by US law.

UK law is able to strike this balance between private and public interests by bypassing the very notion of development rights. Thus UK law avoids most situations in which land-use decisions can cause a partial regulatory taking. A dramatic 1947 reform of the planning law removed all then-existing, unbuilt development rights. A one-time compensation fund was set up to cover claims.¹⁸ From then on, statutory plans – equivalent to US zoning – would no longer grant development rights and thus there could not be a "downzoning". Land-use plans retained their importance – but would have an advisory role. The right to develop (called "planning permission") would be granted case by case on a discretionary basis and would be valid for five years only¹⁹. If planning permission is withdrawn before the five years are up, the landowner has the right to full compensation for the depreciation in

¹⁷ See for example: *Compensation for Injurious Affection*: Discussion paper. (Government of Western Australia - Land Reform Commission of Western Australia October, 2007). The state of Victoria also commenced on a reform track in 2008-9. Communication with Rebecca Leshinsky, Manager, Planning legislation review, Office of Planning, Heritage & Urban Design, Department of Planning and Community Development, Government of Victoria, Australia. Nov. 2008.

¹⁸ Malcolm Grant, *Compensation and Betterment*, in BRITISH PLANNING 62-76 (Barry Cullingworth ed., 1999)

¹⁹ PHILIP BOOTH PLANNING BY CONSENT: THE ORIGINS AND NATURE OF BRITISH DEVELOPMENT CONTROL (2003) ; Victore Moore, A PRACTICAL APPROACH TO PLANNING LAW. (Blackstone Press, 2005).

property value as well as to indemnification for specific out-of-pocket costs. In practice, revocations are made only when there is an overwhelming public consideration for a policy change. This is extremely rare because under the UK system, the timing of the development permission is very close to the development initiative and government has little uncertainty about its future policies. Decades of practice show that the UK system "works" without over-burdening the public purse.

Where major ("categorical") takings are concerned, UK law provides more certainty for landowner than US law. Recognizing that property values might be diminished when land use plans – though advisory – designate land for some types of uses, UK law gives landowners two optional causes of action for making inverse-condemnation claims. One procedure, called "planning blight" is available when a local plan designates private land for a public use. The plan does not have to be officially approved and may even be diagrammatic. The property may still retain some beneficial use. The landowner only needs to prove that if sold, the property would obtain a price significantly below what it would have obtained without the designation for public use. The second procedure, called "purchase notice", is available for any land use designation, not necessarily a public use. The threshold condition, however, is more difficult to prove than for planning blight. The landowner must show that the property has no beneficial use at all and that the owners' request for "planning permission" had been refused. Planning authorities try to avoid blighting property, so the number of claims for major takings is very small.

Now to the other side of the debate. What can proponents of property rights take from this study? They should first heed the lessons from the Measure 37 experience, as detailed above. If new statutes are proposed, they must contain adequate internal checks and balances. Proponents of such statutes have much to learn from the international experience. The survey showed that seven countries other than the USA have statutes that grant compensation rights for some types of partial takings, not only major ones (Finland, several of Austria's states, Poland, Germany Sweden, and Israel). American proponents of property rights should, however, note that none of these countries recognize takings claims where there were no prior development rights. The underlying notion of all compensation laws is *reliance* on government decisions, not reliance on private wishful thinking.

Which among the seven countries can serve as useful models? The experiences of Finland and Austria leave too much legal uncertainty to be useful (although the great assortment of laws offered by the nine Austrian states may stimulate some ideas). Among the five remaining countries, the Netherlands and Israel are models that can help to foresee what mistakes to avoid. These countries too – though to a lesser extent than Measure 37 – have over-burdened the public purse with a disproportionate number of claims. Poland's law is still embryonic in practice. The most interesting models, in my view, are the remaining two countries – Germany and Sweden.

German and Swedish laws on regulatory takings are of the same vintage, with minor but interesting differences. Both laws draw a clear distinction between major and partial takings and provide high certainty on both. When property is designated for a public-type use (one that falls among a long pre-defined list), the landowner has a statutory right to full compensation by means of a 'transfer of title' claim that can be made at any time. There are no preconditions.

In cases of partial takings too there are rights to full compensation (beyond a *de minimis* level) but there is a set of preconditions. Unlike US law on partial takings, where the precondition of showing

"investment backed expectations" has no preset criteria and is to be determined case by case, the German-Swedish preconditions are predefined and easy to determine. The pivotal concept is a time frame (aptly called "implementation time" in Sweden). Compensation rights last for seven years in Germany and for five to fifteen years in Sweden (usually closer to fifteen). These time frames are counted from the time the development rights were initially granted, not from the date of approval of the injurious amendment (so this should not be confused with a regular statute of termination, which is additional. The idea behind the implementation time is to create a sharing of risk between landowners and the government body. The Dutch model too is based on the idea of a shared risk, but it has no preset time frame, thereby leaving considerable uncertainty for both sides and more need for litigation.

By means of the implementation period, the concept of reliance on government decisions – a notion that underlies the regulatory takings laws of most countries, including the US - becomes transparent to both sides. Unlike UK planning law, the German and Swedish laws do recognize development rights and do grant compensation rights when government changes its mind and downzones. However, the German and Swedish models are based on the rationale that the public purse is not a timelessly open-ended insurance policy against a change in public decision. If landowners wish to be ensured that the development rights will not be restricted without compensation, the landowners should apply for a development permit before the preset time frame is over. The development rights do not self-terminate – but if the landowners procrastinate they take the risk of a downzoning without compensation.

The concept of time-limited compensation rights has a potential ancillary benefit as a growth-management tool in high growth areas. "Normal" planning regulations across the world are notoriously bad at controlling the timing of private development decisions and planners everywhere seek ways to either regulate or to incentivize developers. In high-growth areas in the USA the implementation time frame can serve as a growth-management tool to encourage landowners to channel their development decisions into a specified time frame. The public authorities can thus better manage infrastructure investments, school thresholds, housing mix, or versatile employment opportunities. As a growth management instrument, the Swedish model has an advantage over the German model in that the time period is flexible and must be determined at the time of each plan-approval decision that grants development rights. Indeed, while in Germany the implementation period is generally not used as a growth management instrument, in Sweden there is an increasing (though still small scale) use to incentivize commercial developers in urban redevelopment projects²⁰.

6. Can the international differences be explained?

The findings show major differences among the regulatory takings laws of most of the countries in the "sample". There are also no signs of transfer of legal knowledge between countries, except between Germany and Sweden whose laws do share some distinct concepts. Based on the set of countries analysed here, one cannot point out either trends of divergence or of convergence; each country's law

²⁰ In Sweden too, the implementation period is not viewed as primarily a growth management instrument, but in recent years the time frame is being increasingly used for this purpose. The limited and relatively new use of this tool – mostly vis a vis commercial developers in major urban redevelopment projects - reflects the character of the development process in Sweden, where commercial developers are not yet as important a sector as in many other countries. Imposition of a time limit on private, non-commercial developers is not customary (based on a conversation with Thomas Kalbro, the author of the Swedish chapter, December 2008).

on regulatory takings seems to follow its own track. The only trend of convergence is the greater propensity observed in all the countries in this research, for government authorities to rely on negotiated solutions to solve or mitigate some of the adverse effects of regulations on land values. However, this trend is part of a broader tendency towards more private-public partnerships, and is not specifically related to regulatory takings law.

The only modest trend of legal convergence, discernable with a magnifying glass, is found among European countries and is due to their legal connection with the EHCR and their affiliation EU (though this latter affiliation is not of direct legal relevance). But this convergence is not toward the middle of the "scale" of compensation rights but affects only one side of the scale – the extreme non-compensation doctrine. In other words, the mild convergence in Europe works to some degree to enhance property protection.

The intuitive question to ask is: how can one explain these differences? In the absence of enough prior comparative research on regulatory takings or a relevant theory from which to derive criteria or hypotheses to guide the selection of countries, I relied on common-knowledge variables: affiliation with one of the two major Western legal families, institutional structure (federal or unitary states), geographic proximity and language-cultural affinity. The findings show, that the differences among the countries cannot be "explained" according to any of these variables because there are significant differences within each grouping.

Belonging to the same legal family has not led to a convergence of the law on regulatory takings. Eight of the nine European countries – all but the UK – are civil-law countries. As I have shown, this group of countries exhibits both extremes on the property-rights scale. The modest push towards convergence noted above is due not to the fact that the eight countries are of the civil-law family, but to the impact of the jurisprudence of the ECHR – a supra-national institution that "belongs" to neither family. The same picture holds for the five countries in the common-law group - the UK, the USA, Canada, Australia and Israel²¹. These countries had links with British law at some point back in their legal history. The comparative findings show that there is no "British approach". These five countries also span the two extremes on the scale: Canada on one side (extremely restrictive) and Israel on the other (unworkably excessive). The USA is somewhere in the middle. Whatever was these countries' history, today there are few legal similarities.

Nor will knowing whether a specific country is unitary or federal be of much help in predicting the country's stand on regulatory takings law. The eight unitary states exhibit almost the full range of degrees of compensation rights for regulatory takings: France on the one extreme and the Netherlands and Israel on the other extreme. The five federal jurisdictions also encompass a broad range: Canada takes the most extreme no-compensation stand among the countries analysed, with Australia a close

²¹ Israel is regarded by comparative-law scholars as a mixed system, but with strong attributes of the common law tradition. On the one hand, precedence is a major source of the law and common law is still prominent in a few areas, on the other hand, statutory law is dominant in most fields of law today. See Glendon, Carozza and Picker (2007). p. 948 (about mixed jurisdictions in general) and pp. 976-982 (about Israel). In the field of planning law, the British influence dates back to the British Mandate on Palestine from 1921 to 1948. The planning law is a direct derivative of legislation enacted during that time.

next, whereas Germany is among the group of countries with generous compensation rights (though not as extreme as Israel or the Netherlands), The two remaining federal countries – the USA and Austria – take midway positions.

What about geographic proximity? One would have assumed that knowledge transfer would occur at least among closely located countries, particularly those with similar languages or culture. This assumption too, is not supported by the findings of this study. Germany and Sweden are the only case where there was a transfer of a legal approach (from Germany to Sweden). These pair of countries does have a language affiliation and they are not too distant geographically, but they are the only such pair. There are countries closer to Germany with no signs of knowledge transfer. Most striking are Germany and Austria which are not only geographically close, they also share the same language and many cultural elements; yet their laws on regulatory takings bear little similarity. There are also major differences between other pair of neighbors: Netherlands and Germany, Netherlands and Flemish Belgium (the latter is not covered here, but I can comment that their laws on regulatory takings bear no resemblance²²), Netherlands and France (geographically quite close), the UK and France or the Netherlands across the Channel, Sweden and Finland, or Canada and the USA. Ironically, another couple of countries with strong similarities on takings law are the Netherlands and Israel, despite the absence geographic proximity, language or cultural affinity, shared legal history, of any evidence of knowledge transfer. The similarities are due to pure chance.

Could there be some other explanatory variables outside the legal system, such as population density, urban patterns, or pulse of development? Based on the present set of countries, I do not see evidence for such conjectures, but perhaps further research may prove otherwise.

So, as counterintuitive as this may seem, there is no escape from concluding that, based on the current state of knowledge, there are no apparent explanations based on the usual assumptions about legal families, institutional structure or transfer of knowledge among proximate or culturally-similar countries. This conclusion brings me back to the starting point: The problem at hand may be universally shared by all countries with land use regulations, yet each country has adopted a legal solution largely insulated from the experiences of other countries. The explanation for why a particular legal approach on regulatory takings emerged in a particular country at a particular time is a question for historians or political scientists. This may reflect the views held by an individual who was a decision maker (legislator or judge) at a particular point in time when there was an opportunity to mold the law (whether through legislation or through an important court decision). Alternatively the emergence of a particular legal approach may reflect the influence of an effective political lobby. These questions may be worthy of further research, but its contribution will be more in the realm of political science than law.

²² The Flemish planning law recognizes compensation rights only in rare circumstances, whereas Dutch law grants the broadest compensation rights among all the countries in this book (as explained in Chapter 2). For Flemish Belgium law see (title unofficially translated by me: Articles 84-86 of the Decree of 18 May 1999 *Providing for the Organisation of Town and Country Planning*. Moniteur Belge — 08.06.1999 — Belgisch Staatsblad

More than external factors, the availability of comparative knowledge may in itself contribute to legal changes. Further comparative research may become a factor in the exchange of legal knowledge and thus, may affect the future evolution of regulatory takings law in various parts of the world.

New Horizons in Teaching Law

Michael Reynolds¹

Faculty of Science and Technology, Anglia Ruskin University, Chelmsford, United Kingdom
and Senior Visiting Fellow, Law Department, London School of Economics and Political
Science

Email: michaelpreynolds@btinternet.com

Abstract

This paper attempts to evaluate the dichotomy between the professional law and academic curricula describes drawing upon my student experiences in four leading colleges of London University as well as experience in the teaching of law at undergraduate, professional and post graduate/doctoral level at Anglia and at a leading professional law school in London for City trainees.

My impression is that the age old gap between practice and academic study is still too wide and that academicians must understand more fully the needs of the professions if the professions are to properly benefit from graduates in this discipline and in others. The question troubling economists and politicians in our country today is whether the country can afford to go on as we are. In the context of education this translates to whether we can continue to afford the dichotomy that exists between academia and practice: the greater the divergence, the greater the suspicion on both sides; the greater the convergence, the greater the benefit to the economy and society as a whole. To that latter cause I know from my experience of your members, whether as working colleagues or expert witnesses in many actions I have fought, the RICS like my profession is committed and will always be committed to that end. As Professor Twining suggested the key questions are: what should every lawyer know and what should he be able to do?ⁱ

Some of my views derive from my teaching experiences at Anglia Ruskin University and our rapport with the RICS examiners and practitioners. Their feedback has been especially useful in describing their key concerns and requirements for best practice. Further integration between the office and the classroom is of interest to both practitioners and teachers to the professions. Whilst knowledge is a key facet, teaching analytical skills is also important. Thus, it is no use a law or engineering graduate having taken professional examinations not to be able to write intelligibly and to understand basic commercial and legal interests and rights. It is no use surveying graduates who would like to do claims work not being able to understand the basic principles of the law of damages, and it is no use, as I experienced recently an arbitrator, not understanding fundamental concepts of natural justice. The latter example demonstrates the need for more rigorous CPD by those best qualified and experienced to teach it, always remembering that professional education does not stop on graduation, or at qualification: it is continuing. Business can no more afford to cut the training budget than it can the marketing budget.

These are just some of the concerns I outline but there are many others. In the case of construction law and its application there is a real need for those advising in this area to be properly qualified whether as lawyers or experts. My recent work for Anglia's Faculty of Science and Technology

¹ PhD (LSE), LL.M (Lond.), M.Sc (Lond.), F.C.I.Arb., F.C.I.O.B. Solicitor and Chartered Arbitrator. Senior Lecturer, Part-time Anglia Ruskin University and Visiting Senior Fellow, Law Department, London School of Economics and Political Science.

Department in developing a new MSc course in that subject is geared to that end and has some novel features which the surveying profession will undoubtedly welcome and benefit from. Anglia will therefore become a future centre for innovation and research in this field.

Keywords:

Law, teaching, case method, academic and professional approaches.

Introduction

The real question today is why we teach law and not how we teach law. We teach law to practitioners so that they have sufficient legal knowledge to go about their daily business whether as lawyers, surveyors, mediators, adjudicators, arbitrators or chairmen of tribunals. How that knowledge is delivered varies as between universities. The university teaching tradition has focussed on a deeper and broader education than professional schools. In law this is stark. Whilst in civil law countries lawyers are all graduates that has not always been the case in England. In England apprenticeship was the exclusive mode of legal training. The legal profession jealously guarded its qualifying standards and whilst a degree is now an accepted pathway the profession still qualifies its members. This has caused tensions in the past and maybe it will be accepted that the university is the best place for acquiring knowledge and the professional law school the best place for attaining pre-requisite skills to practice. Other professions do not have this problem and it may be with the onset of Alternative Business Structures (ABS) this problem will dissipate.

In England the gulf between academics and practitioners is vast. It is no wonder that Twiningⁱⁱ has described this earlier educational system as: “a dreary procession of disappointments and missed opportunities.” Many a legal career has also been so described. Up to the 1970s the education of lawyers may be described as apathetic with a general sense of public and professional disinterest apart from keeping a cap on the entrance levels. Because of this law descended to being the poor relation of medicine and engineering in the universities.ⁱⁱⁱ By 1993 law and politics were placed at the bottom of the scale of public funding. I regard this as an incredible scandal given the importance of those subjects in national life.

I recently asked a law professor from South America whether this was the same in his country and he said that in his country every lawyer was a teacher and a practitioner, otherwise they could not call themselves lawyers. I also asked a former member of New York’s Supreme Court and she said that on the whole the law firms in the United States had little affinity with the universities.² This surprised me because of the strong affinity between the Law Schools and the United States Government. After all, Archie Cox, special prosecutor of Watergate was a former Harvard Law Professor and served as

² However, I note that many of the current lecturers at Harvard Law School are practising lawyers or judges.

President Kennedy's Solicitor General. In England, there is little affinity, but then we probably have the best jurists of all who happen to be the members of our new Supreme Court.

Until the 1850s Roman law, Ecclesiastical law and Admiralty law were taught at Oxford. It did not teach commercial law subjects; neither did it teach any subject taught at the Inns of Court which had a monopoly on teaching the Bar. Whilst Chief Justice William Blackstone held the Vinerian Chair at Oxford in 1755 and London University started a course in 1828 with Cambridge following a little later it was the Inns of Court that provided what Graveson described as "the university of the Common Law." (Graveson, 1968, p.1)^{iv} He also described how it was not the tradition for English practitioners to attend university "either to read law, or until the latter part of the nineteenth century to read any other subject." This divergence in England between professional and academic teaching accounts in my view perhaps for the uneasy relationship between academics and lawyers. As the Dean of the Faculty of Law of King's College wrote (Graveson, p.3) :

Academic lawyers too often forget that this academic tradition is alien to the professional tradition of training in the United Kingdom.

Whilst that may have been the case in 1968, that is not true today. Today practitioners and some judges teach construction law. Graveson's description does not extend to those academics that practice, or to the senior judiciary who have excellent relationships with leading academics. This English tradition contrasts with the American tradition. Whilst attorneys and academics have a similar relationship to their English cousins I detect a closer affinity between the American law schools and practitioners' offices. This may have an historical origin in the fact that in 1776 the Anglo-American colonists withdrew from the Inns of Court system to educate their lawyers at University Colleges. These colleges regarded legal training as post graduate study. The American law school curricula, unlike its English counterparts, taught a wider range of subjects so that today in many respects the American student may be better prepared for practice than his English counterpart. This is probably because one of the founders of modern law teaching, Christopher Columbus Langdell, the Dean of the Harvard Law School said that the teaching of law should be a matter of academic training for the practice of law (White, G.E., 1985. p. 26).^v We can learn much from Langdell in terms of this approach. Langdell's students were required to read the case, to ponder the judgement, and to analyse it so that they understood how the judge had arrived at his decision and importantly what the decision, i.e. the principle of law was. Contrast that approach to the reading of text books and the learning of principles from text book writers or learning principles from lecture notes- not quite the same exercise. A reading of cases not only gives the student an understanding of the application of the principle, but of analytical and logical reasoning; and not only in the primary context of the principle, but in a variety of factual matrices. It is the ability to identify and dissect what Langdell called

“fundamental legal doctrine” (Langdell, C. 1871)^{vi} that is essential for students to grasp for use in practice.

Whilst the benefits of the American system may appear greater to an English Common lawyer I may be guilty of wishful thinking since „the grass is always greener on the other side.“ It may look “greener” to me because the students in the United States are more mature and law is taught as a post graduate subject. Thus my comparison is not like for like. However a study of the benefits of a post graduate degree for all lawyers would in my view be a good idea. Just as in my time at Law School in the late 1960s most of my classmates at the College of Law in London were graduates or had other work experience.

One of the challenging propositions for academics to consider is whether the law is led by teachers, or by the judiciary-who make the law. Some jurists say the judges. Others say the teachers who inspired the judges. It may therefore be a bit of both. We may therefore ponder whether Lord Atkin would have declared the “Neighbour” principle in *Donoghue v Stevenson*^{vii} if Justice Cardozo had not decided to implement a theory of negligence enunciated by Supreme Court Justice Oliver Wendell Holmes in the seminal American case on causation, *Palsgraf v. Long Island R. Co.* (1928).^{viii} In that case, Justice Cardozo found the railway company liable when a passenger boarding a train dropped a package which exploded and caused some scales to be dislodged injuring the plaintiff. Cardozo came to the unprecedented view that a tortious act occurs when a risk, possible of apprehension, is taken and results in harm. Was that finding the product of Cardozo’s education, or his innovative judicial skill?

It has been said of Cardozo that (White, G.E. 1985)^{ix}

where his findings were creative, his opinions tended to minimise that creativity; where his articulation or application of principles constituted substantial innovations in the state of legal doctrine, his opinions did not often emphasise these innovations. Innovative judges, in fact, regularly suggested that their decisions preserved in tact the corpus area of law.

In my view judges like Cardozo extended and applied the principle as Langdell had taught it at Harvard. It did not simply “maintain an area of law” but applied the principle of the law to changing social, economic and political eras. Thus, was Cardozo also able to extend negligence in *McPherson v Buick Motor Co* (1916).^x Judges like Cardozo and Atkin, and later Lord Denning were creative and activist. They recognised changing social times and used the law to do justice to the case. Was it their legal education that made them such champions of justice, or their innate sense of justice, or their experiences in practice? It must be a combination of both because without the education they would not have so understood the principles they were applying and without the experience of practice they would find it difficult to do justice to the parties. Each of them also had a natural intelligence that distinguished them from others.

Today’s challenge

The challenge today, however, is not to emphasise the differences between law curricula of professional and academic institutions, or between the professions and the academic fraternity, but to suggest how they must address more effectively a common foe: ignorance, denial of right and obligations, crime, and all the attendant problems we see today in economic decline whether accelerated by greedy speculators who gamble with other people's money, who bet on a financial loss, or fall in the market, or the currency, or by people who live way beyond their station in life and their means. The economic depression of recent times will most probably lead to greater conflict not only within the jurisdiction as without, not only within states, but between states, and hence there will be a great need for experts and lawyers to join common cause in resolving these pressing problems. Hence such initiatives as the disputes panels being established to deal with the transfer pricing problems and taxation (Baistrocchi, E.A. 2006)^{xi} and also the encouraging steps being taken by CEDR (2009)^{xii} promoting settlement in international arbitration disputes.

Bridging the gap

In this paper we consider the ways in which law has been taught, is being taught and how possibly it may otherwise be taught. Law is taught in university colleges and professional law schools and in many professional institutions where law plays a role e.g. the Chartered Institute of Arbitrators. There is considerable difference between the teaching of law for practical purposes and the teaching of law for academic purposes. Are we teaching law to make good lawyers or to teach law in general? The one purpose does not necessarily complement the other. Some students excel at the academic study of law, whilst others find difficulty with the practical aspects, and still others find it hard to adapt to an office environment. That is a problem with teaching law in law faculties and schools. It is less of a problem in my experience of teaching science faculty students at Anglia because their knowledge of law is not necessarily for practical purposes. What we have taught at Anglia mirrors in some ways the case study approach pioneered by Dean Langdell of Harvard Law School over a century ago. Over the last few years our construction law students have been set the practical assessment of analysing construction problems and legal issues involved in the procurement and construction of a fictitious international airport in the Thames estuary. Whilst we did not require them to calculate the quantities of material required for construction, nor advise where the dredged material would be deposited, nor how much space that would take, we did ask them to advise on the formation of the contract by e-mail, the terms applicable, the disputes process to be followed, and an alternative type of contract that would avoid the adversarial relations of conventional standard forms. Many students took this back to their practices and worked with their senior managers in producing some excellent assessments with an average mark in the mid 60s and several first class papers. Whilst we do not expect our students to be able to analyse complex international construction legal problems their ability to grasp the key aspects of the problem and the issues will undoubtedly stand them in good stead in the years to come.

Also whilst the assignment was divided into five parts two were essay type questions and the rest were specific problem issues. What we did here was to amalgamate the law faculty type essay question with the professional law school practical type question, albeit we were teaching science students. Our students responded well to the shorter problem questions. Their reasoning and analysis was good; whereas they had difficulty with the longer essay part questions, quite natural where science students have more of an appetite for analysis than discussion and debate. The variety and dichotomy of the questions made them interesting and challenging and we hope to build on this further this year.

The other experiment we tried at the end of the first semester construction law was a mock arbitration. For that we took the case study and the students played the parts of various witnesses as if the study had been an international arbitration hearing. By playing the roles of the contractors, architects and engineers and project managers the students got a very good feel for the real roles they might one day play.

What we try to teach are the fundamental principles of law. In construction because the syllabus has a wide coverage of the law of contract, construction contract standard forms and construction law principles we teach the key points. We teach them what to look for and we take them through some leading case studies. In teaching subjects such as property law we teach them the general principles, and focuss on the key elements of interest to the surveyor e.g. title to property, easements and covenants. In the latter two areas we take them through some leading judgements and ask them to identify the key facts and the decision or principle and then ask them to apply it to other factual scenarios. We also train them in the use of researching law and using all the latest digital search facilities in the law library. Thus, if the law is made of practical use and interest we find that our students attain a genuine interest in the subject or a particular aspect of the subject. Many of them follow this up in later career by becoming expert witnesses, adjudicators and arbitrators. Sometimes science students do better than law students.

Law examinations and assessments

Law examinations do not test how a student has performed in class, how he or she has prepared, nor more importantly how the candidate will perform after qualifying, or how effective they will be in business or academia. The measure of written examinations is limited. What is the point of teaching advocates to be good writers, to know the law, but not to test oral skills, inter personal skills and forensic skill? What is the point of teaching solicitors if they too lack inter personal or management skills? And what is the point of just teaching English Common Law in a global economy. The recent Great Recession and the prospect of a Depression, worse than the 1930s, demands that lawyers and others widen their skills and their horizons. From the chambers in the Temple to the City Office in Canary Wharf lawyers are realising that the old ways will not do: neither professionally nor economically and that they will have to change to survive. That demands a challenge to legal

education. It must be delivered more efficiently and productively. Courses today are delivered in a variety of formats whether as seminars, lectures, workshops, small group sessions, webinars, video conferences, or teams of students working together on a project. Whilst these formats may keep students awake there is a great need to underpin such activity with the basic learning techniques and skills. Whilst the digital age has given us many huge advantages, not least the internet, it is important to keep these digital aids and modern innovations in perspective. Students still require time to learn and to think. It is no use giving them access to a library of books or a host of websites if they are only to remember the last page of the last book they read or the first article they read on the web link. We need to remember that overshooting the runway of knowledge can be just as bad as landing short of it, or making a hard landing where months of hard work is wasted.

Towards a new Horizon

What we need in the UK today therefore is not less teaching, but more, not just domestic Common Law, but international law; not just English Commercial Law, but transnational trade law, and not just the law of contract or tort, but the specialist areas of law that are so important to industry and the economy. Law has barely touched upon the issues of “cyberspace law;” save in the context of intellectual property rights and transborder intellectual property issues; a whole new horizon, if not, universe of law is emerging to challenge legal scholars and practitioners. Whilst the rights of the private citizen are important to all of us that is not of much value if obligations are not observed and enforced. We have heard in the last decade a great deal of rhetoric about human rights, and whilst these are to be regarded and respected they are of little use without the exercise of the responsibility that goes with the right. For every right there is a corresponding obligation. But today what is needed in legal training for whichever profession is a more productive, practical and dynamic educational rubric that gears students for the challenges of modern day practice. In university all they have to do is learn. In the office they are on their own. Whilst most firms have excellent training provision increasingly they tend to be geared to commercial productivity especially in Busy practitioners do not have time to “train” staff, nor educate them. Practitioners have become jealous of their own education and are not eager to share it with modern graduates. Practitioners take it for granted that graduates know what to do and what is expected of them. The fact is they do not; no more than a person who wins the election can suddenly be transformed from an MP to a Prime Minister, and no more than an experienced counsel can become a High Court judge. None of these transformations occur overnight and some take years of experience.

It was very different when I was an articled clerk and had to learn at the feet of the master, my principal, who happened to be an academic lawyer-a Master of Jurisprudence, and a former student of Cambridge, the Sorbonne and Warsaw as well as being a member of the diplomatic service. He had a great interest in legal training and I was certainly “trained” if that is the right word. Letters were

drafted, then checked, amended or re-drafted, checked again and edited and then approved. Today, time does not always permit such luxury and therefore there is a greater need for those teaching law to bear in mind the practical necessities of those who may practice law or use it in their profession. There is no substitute for a thorough education in this subject, but it must be studied wisely for in the words of Dean Langdell: “if law be not a science, a university will but consult its own dignity in declining to teach it.” (Warren, Z.C. 1886).^{xiii} Many I know oppose Langdell’s approach not least did the great American jurist Oliver Wendell Holmes. But that was in a different age. Today we are in a technological age: the age of digital communication, e-mail and twitter. Law is very much a science, using the means of science to transform the process. But in essence it has always been a science: taking instructions, reviewing and analysing the facts, testing the evidence, applying the legal principle, examining and cross-examining the witnesses, giving an opinion, award or judgement, and often reliant on expert evidence. It could not be more scientific in theory or practice in my experience and that is why it is a pleasure to teach science students.

So far as the judges are concerned, those who teach us the practice of law, I leave you with this thought in the words of that eminent jurist Benjamin Cardozo:

The tide rises and falls, but the sands of error crumble.

Dr Michael Reynolds

Lecturer in Construction Contract Law, Civil Law and Legislation and Property Law, Anglia Ruskin University Faculty of Science and Technology and Visiting Senior Fellow, London School of Economics and Political Science.

ⁱ Twining, W. 1994. *Blackstone's Tower The English Law School*. (The Hamlyn Trust London Stevens & Sons/Sweet & Maxwell.)

ⁱⁱ Twining, W. 1994. p 24.

ⁱⁱⁱ Twining, W. 1994. p.41

^{iv} Graveson, R.H. 1968. Training for the Law-the University Contribution *The Journal of the Society of Public Teachers of Law*.

^v White, G.Edward.,1985. *Tort Law in America: An Intellectual History*

^{vi} Langdell, C. 1871. *A Selection of Cases on the Law of Contracts* v-vii

^{vii} [1932] A.C. 562

^{viii} *Palsgraf v. Long Island R. Co.* 248 N.Y. 339 (1928)

^{ix} White, G.Edward.,1985. *Tort Law in America: An Intellectual History*

^x 217 N.Y. 382 (1916)

^{xi} Baistrocchi, E.A. 2006. The Transfer Pricing Problem: A Global Proposal For Simplification *Tax Lawyer* Vol. 59, No.4 p.943.

^{xii} CEDR. 2009. *Rules for the Facilitation of Settlement in International Arbitration*.

^{xiii} Warren, Z.C. 1908 Dean Langdell’s Address to Harvard Law School Association, November 5, 1886 *History of the Harvard Law School* 374

The proposed directive on industrial emissions: What will change regarding the best available techniques (BAT) – the Finnish perspective

Eeva-Maija Puheloinen¹

¹Faculty of Law,
University of Helsinki,
Finland

Email: Eeva-Maija.Puheloinen@helsinki.fi

Use of the best available techniques (BAT) is a widely recognised principle in environmental law which can be found in many international treaty regimes. The BAT principle is also found in EU law, in particular in the directive concerning integrated pollution prevention and control (IPPC). At the moment, the IPPC is under revision in the European Union – the proposal for a directive on industrial emissions would incorporate seven sectoral directives related to industrial emissions, including the IPPC. The current state of the proposal is that it is under consideration in the European Parliament and the Council. The proposal would clarify and strengthen the concept and the role of the best available techniques (BAT). Accordingly, the BAT would refer especially to documents resulting from the exchange of information organized by the Commission; namely, BAT reference documents (BREFs), and their BAT conclusions. The BAT conclusions would be used, for example, as reference for setting permit conditions, and permit emission limit values should not exceed BAT associated emission levels. The proposed directive would change the role of BAT from a principle to a rule – BAT would lose its dynamic role as something to strive towards and become more of a general rule. It may be that the overall level of environmental protection would improve, but other unwanted consequences may emerge.

Keywords:

Best available techniques (BAT), Industrial emissions, Integrated pollution prevention and control (IPPC)

1 Introduction

Preserving, protecting and improving the quality of the environment is one of the key objectives of the European Union policy on the environment set out in Article 191 of the Treaty on the functioning of the European Union.¹ The policy aims at a high level of environmental protection, its leading principles being the precautionary principle, and the principles that preventive action should be taken, environmental damage should be rectified at source as a priority and the polluter should pay.

¹ Treaty on the functioning of the European Union, consolidated, OJ C 83, 30.3.2010.

One application of the principle of prevention is the use of best available techniques (BAT) which are the most effective techniques to achieve a high level of environmental protection, taking into account the costs and benefits. The principle has been shaped in the 1970s (Kuusiniemi, 2001) and can be found in many international treaty regimes, such as the 1992 UNECE Water Convention² and the 1992 OSPAR Convention,³ where an obligation to prevent harmful releases to the environment is directly related to applying best available techniques. BAT means a shift from the “end-of-pipe” solutions to the regulation of industrial processes with a focus on preventing the harmful discharges in the first place (Sadeleer, 2002, Kuusiniemi, 2001). In addition to the principle of prevention, the BAT can be seen as an application of the at source principle as well as having a connection with the precautionary principle.

The BAT principle is also found in EU law, in particular in the IPPC.⁴ The purpose of the IPPC adopted in 1996 is to prevent or, where this is not practicable, to reduce emissions in the air, water and land from industrial activities.⁵ This goal is to be achieved by applying the best available techniques (BAT). At the moment, the IPPC is under revision in the European Union – in the proposal for a directive on industrial emissions BAT remains a central element.

The largest industrial installations generate a significant proportion of total emissions of major atmospheric pollutants – 83% for sulphur dioxide (SO₂), 34% for oxides of nitrogen (NO_x), 43% for dust and 55% for volatile organic compounds (VOC). They also have other important environmental impacts such as emissions into water and soil, generation of waste and the use of energy. Approximately 52,000 installations across the 27 Member States fall within the scope of IPPC (COM(2007)843 final). In Finland, the number of IPPC-level installations is around 700 (SYKE, 2006). Consequently, concerning environmental protection, the importance of IPPC and its revision is clear.

This paper is a literature review on the proposed directive on industrial emissions and the best available techniques (BAT) – how the proposed directive will change the concept and role of the BAT and what effects the change would possibly have. Also the consequences in Finland are evaluated. This study begins with an analysis of the current legislation and its proposed revision at both EU and domestic level. The legislative evaluation utilises legal dogmatics research methodology, which pursues knowledge of the existing law, but it may also in many cases lead to a change in the law. In this paper I discuss the change in BAT that the acceptance of the directive on industrial emissions would bring, as well as the consequences in practice and the possible effects on environmental protection. I begin by exploring the BAT principle and continue with the shift from IPPC to the directive on industrial emissions. I then introduce the BREFs and their preparation. A more careful analysis of the new BAT introduced by the proposed directive on industrial emissions is then presented. The Finnish perspective as regards to BAT is considered in section 6. Section 7 (Conclusion) summarises the change linking it to a broader discussion.

² Convention on the Protection and use of Transboundary Watercourses and International Lakes, Helsinki 17 March 1992.

³ The Convention for the Protection of the marine Environment of the North-East Atlantic, Paris 22 September 1992.

⁴ Directive 1996/61/EC, codified as Directive 2008/1/EC of the European Parliament and of the Council concerning integrated pollution prevention and control, OJ L 24, 29.1.2008.

⁵ IPPC Art. 1.

2 BAT principle

While best available techniques are a central principle in the field of environmental protection, it is not an absolute. The use of such techniques should not only be technically feasible, but also economically acceptable. Most definitions of BAT include comparison of interest between the costs and the nature of harm being caused to the environment. Still, the cost-benefit analysis should not lead to invalidation of the principle of prevention – the costs must be in balance with the nature and volume of the releases in question in each case. The new technology must be evaluated with the benefit to the environment and the costs in mind, but all the risks of the activity in question must also be taken into account (Sadeleer, 2002).

BAT obligates the operators to reduce emissions as low as technically and economically possible. Applying the BAT principle does not mean committing to a certain technique, process or machinery but it defines a level to which pollution of the environment can be prevented, reduced or limited. All the techniques which can reach this level are considered to be the best available techniques. (Kuusiniemi, 2001).

BAT is mentioned in IPPC as a general principle governing the basic obligations of the operator.⁶ The principle has spread itself from IPPC to the Finnish environmental protection legislation, although the principle has existed already before that in some form, for example, in the Finnish water permit system (Ekroos, 2001). The present national Environmental Protection Act contains a provision concerning the BAT principle similar to the IPPC.⁷

3 From IPPC to directive on industrial emissions

The EU initiated the revision of the IPPC in 2005. Concerning the revision, the Commission thought that BAT was a sound basis for future development of EU legislation on industrial emissions. Nevertheless, the implementation of BAT had been found insufficient and more positive health and environmental effects would not be achieved without further emission reduction from IPPC level installations. Additionally, considerable differences in environmental standards in different Member States had been observed causing distortion in competition. (COM(2007)843 final).

Consequently, the Commission adopted a proposal for a directive on industrial emissions in late 2007 which would incorporate seven existing directives related to industrial emissions into a single legislative instrument.⁸ Beside the IPPC, the proposal would recast the directives on large combustion plants,⁹ waste incineration,¹⁰ solvent emissions,¹¹ and titanium dioxide.¹² At the moment, the proposed directive is under consideration by the European Parliament and the Council. The Council, the Parliament

⁶ Proposed directive on industrial emissions Art. 11.

⁷ Ympäristönsuojelulaki (Environmental Protection Act) (86/2000), Section 4.

⁸ Proposal for a Directive of the European Parliament and of the Council on industrial emissions (integrated pollution prevention and control) (Recast) [COM(2007) 843 final] [SEC(2007) 1679] [SEC(2007) 1682]].

⁹ 2001/80/EC, OJ L 309, 27.11.2001.

¹⁰ 2000/76/EC, OJ L 332, 28.12.2000.

¹¹ 1999/13/EC, OJ L 85, 29.3.1999.

¹² 78/176/EEC, OJ L 054, 25.02.1978, 82/883/EEC, OJ L 378, 31.12.1982 and 92/112/EEC, OJ L 409, 31.12.1992.

and the Commission reached an agreement on the proposal at second reading, and on 7 July 2010 the Parliament adopted the compromise amendment to the proposal. Approval by the Council is expected in the near future. Since the legislative process is still unfinished, I refer here to the outcome of the European Parliament's second reading.¹³

BAT remains a central element in the proposal for a directive on industrial emissions. The proposal would clarify and strengthen both the concept and the role of BAT. What is regarded as BAT at EU level for each industrial sector would be indicated in an information exchange to establish BAT reference documents (BREFs).

4 BREF

BREF stands for best available techniques reference document and is a result from a specific exchange of information. The BREFs are not mentioned in the IPPC, although the exchange of information has been going on since the implementation of the IPPC. IPPC obliges the competent authorities of Member States to follow or be informed of developments in the best available techniques.¹⁴ Exchange of information is organised by obliging the Member States to send the available representative data on the limit values laid down by the specific category of activities and, if appropriate, the best available techniques from which those values are derived.¹⁵

This exchange of information has led to the introduction of best available technique reference documents, or BREFs. These are documents resulting from this exchange of information organized by the Commission and drawn up by the Commission, the competent authorities of Member States and interest groups (the so-called Seville Process). In short, each BREF document generally gives information on a specific industrial or agricultural sector in the EU, techniques and processes used in this sector, current emission and consumption levels, techniques to consider in the determination of BAT, the best available techniques (BAT) and some emerging techniques (EIPPCB, 2010).

In 1997, the European IPPC Bureau (EIPPCB) was established to organise the exchange of information on BAT, associated monitoring and their developments between Member States and industry. EIPPCB forms a technical working group (from 40 to 100 experts) for each BREF to carry out the exchange of information on BAT. The procedure includes a few plenary meetings of the technical working group, sub-group meetings, visits to installations, and submission of draft BREFs for comment. After finalisation, the BREF will be presented to the Directorate General in charge of the Environment within the European Commission at the Consultative Committee (Information Exchange Forum). (EIPPCB, 2010)

EIPPCB had completed the first series of 33 BREFs including six horizontal reference documents (such as energy efficiency) and begun the review of the documents by 2006. At the moment, there are 14 BREFs, which have been formally adopted by the

¹³ Proposal for a Directive of the European Parliament and of the Council on industrial emissions (integrated pollution prevention and control) (recast) - Outcome of the European Parliament's second reading (Strasbourg, 5 to 8 July 2010).

¹⁴ IPPC Art. 11.

¹⁵ IPPC Art. 17.

European Commission, such as the industrial sectors of waste incineration and manufacture of organic fine chemicals. Nineteen documents are being updated, such as those for the ferrous metal processing industry and slaughterhouses and animals by-products industry. The BREFs are hundreds of pages wide documents, for instance, the Reference Document on Best Available Techniques in the Pulp and Paper Industry consist of 509 pages.¹⁶ (EIPPCB, 2010)

In the proposal for the directive on industrial emissions, the BREFs are recognised. BAT reference document is defined as “a document, resulting from the exchange of information organised pursuant to Article 13 (BAT reference documents and exchange of information), drawn up for defined activities and describing, in particular, applied techniques, present emissions and consumption levels, techniques considered for the determination of best available techniques as well as BAT conclusions and any emerging techniques, giving special consideration to the criteria listed in Annex III (Criteria for determining best available techniques).”¹⁷

According to the definition, the BREFs are used in determining the BAT in a particular case. This corresponds to the situation in practice so far. The proposal imposes detailed conditions on BREFs and organizes the exchange of information.¹⁸ The purpose of the exchange of information would be to draw up, review and, where necessary, update the BREFs.¹⁹ A forum composed of representatives of Member States, the industries concerned and non-governmental organisations promoting environmental protection would be established.²⁰

The proposal for a directive on industrial emission changes the role of BREFs as their BAT conclusions become binding. The BAT conclusions are explored below in section 5 (the new BAT).

5 The new BAT

The proposed directive on industrial emission does not bring significant changes to the definition of BAT²¹ compared to the definition in the IPPC.²² Additionally, the criterion for determining best available techniques set out in the proposal²³ is similar to the one in

¹⁶ Available on the Internet at ftp://ftp.jrc.es/pub/eippcb/doc/ppm_bref_1201.pdf, viewed 2/7/2010.

¹⁷ Proposed directive on industrial emissions Art. 3 (10).

¹⁸ Proposed directive on industrial emissions Art. 13.

¹⁹ Proposed directive on industrial emissions Art. 13 (1).

²⁰ Proposed directive on industrial emissions Art. 13 (3).

²¹ Proposed directive on industrial emissions Art. 3 (9): ‘best available techniques’ means the most effective and advanced stage in the development of activities and their methods of operation which indicates the practical suitability of particular techniques for providing the basis for emission limit values and other permit conditions designed to prevent and, where that is not practicable, to reduce emissions and the impact on the environment as a whole:

(a) ‘techniques’ includes both the technology used and the way in which the installation is designed, built, maintained, operated and decommissioned;

(b) ‘available techniques’ means those developed on a scale which allows implementation in the relevant industrial sector, under economically and technically viable conditions, taking into consideration the costs and advantages, whether or not the techniques are used or produced inside the Member State in question, as long as they are reasonably accessible to the operator;

(c) ‘best’ means most effective in achieving a high general level of protection of the environment as a whole.

²² IPPC Art. 2 (12).

²³ Proposed directive on industrial emissions Annex III.

the IPPC.²⁴ In the definition, the main point has been made stronger by leaving “in principle” out of “providing (in principle) the basis for emission limit values” and adding “other permit conditions” to the same connection. The new definition gives some hint of a move towards a less flexible and more powerful BAT.

The big change in the proposal is the introduction of the terms BAT reference document (BREF) and BAT conclusions. The BREFs have been explored above in section 4. The proposal changes the role of BREFs as their BAT conclusions become binding. The BAT conclusions are defined in the proposal for the Directive on industrial emissions as “a document containing the parts of a BAT reference document laying down the conclusions on best available techniques, their description, information to assess their applicability, the emission levels associated with the best available techniques, associated monitoring, associated consumption levels and, where appropriate, relevant site remediation measures”.²⁵

According to the proposal, permit conditions would be set in line with the BAT conclusions, and also stricter conditions are allowed.²⁶ The emission limit values and the equivalent parameters and technical measures would be based on the best available techniques so that emissions would not exceed the emission levels associated with the best available techniques as laid down in the decisions on BAT conclusions.²⁷ Within four years of publication of decisions on BAT conclusions relating to the main activity of an installation, the competent authority should reconsider and, if necessary, update all the permit conditions for the installation concerned.²⁸

If there are no relevant BAT conclusions permit conditions might be set without them when the criteria for determining the best available techniques listed in Annex III are taken into account and the requirements of Article 15 (emission limit values, equivalent parameters and technical measures) are complied with, ensuring a level of environmental protection equivalent to the best available techniques described in the BAT conclusions.²⁹ In a case where none of the relevant BAT conclusions cover the activity or process in question, the BAT is determined by giving special consideration to the criteria listed in Annex III.³⁰ These provisions may be applicable also before the renewal of BREFs and especially their BAT conclusions has been done.

In conclusion, the BAT conclusions introduced by the proposal would, for example, be used as a reference for setting permit conditions, and permit emission limit values should not exceed BAT associated emission levels. Consequently, the starting point is clear – the BAT conclusions will be binding in the future. This means that making of the BAT conclusions will be highlighted.

Decisions on the BAT conclusions will be adopted, according to the proposal for a directive on industrial emissions, in accordance with the regulatory procedure – in other

²⁴ Proposed directive on industrial emissions Annex IV.

²⁵ Proposed directive on industrial emissions Art. 3 (11).

²⁶ Proposed directive on industrial emissions Art. 4 (3) and (4).

²⁷ Proposed directive on industrial emissions Art. 15 (2) and (3).

²⁸ Proposed directive on industrial emissions Art. 21 (3).

²⁹ Proposed directive on industrial emissions Art. 14 (5).

³⁰ Proposed directive on industrial emissions Art. 14 (6).

words with the committee procedure.³¹ The use of the committee procedure was subject of political debate but in the final compromise text the committee procedure remains. The proposal will mean that determining what are BAT conclusions within the BREFs will require more thorough consideration and participation requirements for the technical working groups and the working processes in the discussions of those groups have to be reconsidered (Farmer, 2008).

At the moment, the BREFs are published only in English but their executive summaries are translated into all EU official languages. In the future, the BAT conclusions will be made available in official EU languages.³² This is a must because the BAT conclusions will be legally binding.

Since the current BREFs have not been drawn up with the new BAT in mind, they do not necessarily meet the conditions set for BREFs in the proposed directive on industrial emissions. At least there are no BAT conclusions in the present BREFs. This leads to a transition period during which the current BREFs are updated. This transition time may last longer than the time reserved for the domestic implementation of the directive. During this time the role of the old BREFs may be unclear. They might be treated as the future BREFs if they meet the requirements. Another option is to ignore them. The old BREFs may be treated as subject to the provisions in Articles 14(5-6) which were described above. It may also be that the current BREFs meet the new provisions fully or they can be updated to meet them during the implementation period – some BREFs are already under revision. However, the question of old BREFs under the transition period must be thought through carefully.

6 The Finnish perspective

In Finland, the IPPC has been implemented with Environmental Protection Act³³ and Environmental Protection Decree.³⁴ One special characteristic in the Environmental Protection Act is that it applies not only to the IPPC scale activities but to a great extent also activities on a considerably smaller scale. An environmental permit is required for activities that pose a threat of environmental pollution, and a detailed but not exclusive list of such activities is prescribed in the Decree. Therefore, the operators are currently obliged to have an environmental permit for their activities quite comprehensively. Taking into account the large number of activities at various levels of operation (from small scale to large scale) and the proposed change in BAT, implementing the directive on industrial emissions will undoubtedly bring some changes to the environmental permit system in Finland.

According to Section 28 (1) of the Act, a permit is required for activities that pose a threat of environmental pollution (environmental permit). Permit regulations concerning emission limit values and the prevention and limitation of emissions must be based on the best available technology.³⁵ According to Section 9 (2)(6) of the Decree, the

³¹ Proposed directive on industrial emissions Art. 13 (5) and Art. 75 (2).

³² The final compromise text of the proposal for a directive on industrial emissions, Art. 13(6).

³³ Ympäristönsuojelulaki (Environmental Protection Act) (86/2000).

³⁴ Ympäristönsuojeluasetus (Environmental Protection Decree) (169/2000).

³⁵ Environmental Protection Act, Section 43 (3).

environmental permit application shall include an assessment of the application of the best available technique in the proposed activities. In the process of determining the case specific BAT the EU-level BREFs must be taken into account. Additionally, domestic BAT reports and other possible information may be utilised (Ympäristöhallinto, 2010).

The domestic BAT reports have been drawn up in order to create background information to the BREFs in question as well as produce reports aiming at promotion of environmental protection and harmonizing the environmental permit practice in Finland. Such BAT reports have been published in 20 industry sectors (Ympäristöhallinto, 2010). For instance, a BAT report has been drawn up for installations with a total rated thermal input of less than 50 MW (SYKE, 2008). A similar EU-document is the BREF concerning large combustion plants dealing with combustion installations with a rated thermal input exceeding 50 MW (EIPPCB, 2006).

Because of the probable acceptance of the proposed directive on industrial emissions, the role of BAT will change to a more binding nature, as mentioned above. Since in Finland, the BAT has so far had the same role regardless of whether the activity in question is at IPPC level or on a smaller scale, it remains to be seen what approach will be taken with the national level BAT – will it follow the lead of the BAT of the proposed directive on industrial emissions, remain as it has been before or change in some other direction? It would be probable and reasonable to preserve the flexibility BAT has at the moment in the future concerning small-scale activities.

7 Conclusion

BAT was introduced primarily as a principle. The principle includes a cost-benefit – analysis with a target on achieving a high level of environmental protection. In IPPC, the BAT is characterized as a basis for setting emission limit values, parameters or equivalent technical measures without prescribing the use of a particular technique or technology and taking into consideration the technical characteristics of the installation concerned, its geographical location, and local environmental conditions.³⁶ The BREFs have already made the BAT less dynamic during the IPPC and this development is sure to continue with the implementation of the directive on industrial emissions. The proposed directive would change the role of BAT from a principle to a rule – BAT would lose its dynamic role as something to strive towards and become more of a general rule. Will there be room left for the BAT principle?

There can be seen many consequences on the change of BAT. Firstly, the new BAT forces the operators to the same level, which raises standards for the low-end operators. In conclusion, the new BAT may well lead to a higher level of environmental protection as a whole within the EU which was Commission's original goal. Secondly, the authorities do not have as much discretionary power when setting the environmental permits. This can be a good and a bad thing – there may be less poor permits because of the limited discretion but also it may be that there could be less very good permits since the permit authorities cannot pay attention to the local circumstances as much as before.

³⁶ IPPC Recital 18.

Consequently, that is the third point – the new BAT conclusions do not take into account the local characteristics unlike the previous IPPC BAT because the BREFs are pan-European. This may cause unnecessary economical costs when no addition to the level of environmental protection is gained. This means that the environmental protection may not be as cost-efficient as it has been so far. In addition, this one size fits all thinking applies to the activities as well. BREFs are documents about a specific industrial or agricultural sector but the activities inside that sector may vary greatly. It is impossible to include all the activities and their processes into the BREFs properly. This may lead to applying BAT conclusions and their emission levels into activities which they do not fit so well. According to the proposed directive on industrial emissions, deviation from the BAT conclusions is possible but the conditions seem strict. It remains to be seen how this will be applied in the future.

Additionally, whether the promotion of new techniques will be delayed remains to be seen. The renewing of BREFs is slow – according to the final compromise text the Commission should aim to update the BREFs in maximum of every eight years.³⁷ This would mean that new innovations would be adopted in eight-year leaps, so to say, and this may slow down technical innovations. The eight year leaps may also mean intensive lobbying before accepting new BREFs. Therefore the process of adopting or renewing BREFs must be explicit and transparent.

The many different BATs (the BAT principle, the new BAT, the national level BAT in Finland and the transition period BAT) can be confusing and this must be taken into account in incorporating the directive into Finnish legislation. Implementing the directive on industrial emissions will be a good opportunity to have a broader look at the environmental protection legislation nationally in Finland as well.

8 Acknowledgements

The Academy of Finland, the Finnish Funding Agency for Technology and Innovation (TEKES) and Ekokem are acknowledged for their financial support.

9 References

- COM(2007)843 final. Communication from the Commission to the Council, the European Parliament, the Economic and Social Committee and the Committee of the Regions: Towards an improved policy on industrial emissions. Brussels, 21.12.2007. Available on the Internet at <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=COM:2007:0843:FIN:EN:PDF>, viewed: 12/5/2010.
- EIPPCB (2010). European IPPC Bureau website. <http://eippcb.jrc.es/>, viewed: 13/5/2010.
- EIPPCB (2006). Integrated Pollution Prevention and Control: Reference Document on Best Available Techniques for Large Combustion Plants. July 2006. European Commission. Available on the Internet at <http://eippcb.jrc.es/reference/lcp.html>, viewed: 14/5/2010.

³⁷ Proposed directive on industrial emissions, recital 12.

- Ekroos, M. (2001), Yleiset periaatteet, velvollisuudet ja kiellot, in Kuusiniemi, K., Ekroos, M., Leinonen, J., Marttinen, K., Siitari-Vanne, E. (Eds) *Uusi ympäristönsuojelulainsäädäntö*, pp. 82-85. Helsinki, Edita.
- Farmer, A. (2008), Revising IPPC: Incremental change rather than a radical overhaul of EU industrial emissions policy. *Environmental Law Review*, 10 (2008): pp. 258-282.
- Kuusiniemi, K., Ekroos, A., Kumpula, A., Vihervuori, P. (2001), *Ympäristöoikeus*, WSOY, Helsinki.
- Sadeleer, de N. (2002), Environmental principles: From political slogans to legal rules, pp. 84-86. Oxford University Press; New York.
- SYKE (2008), Suomen ympäristökeskuksen raportteja 8/2008: Parhaan käyttökelpoisen tekniikan huomioonottaminen ympäristölupamenettelyssä. Esimerkkinä alle 50 MW:n polttolaitokset sekä pintakäsittelylaitokset. Available on the Internet at <http://www.ymparisto.fi/download.asp?contentid=82087&lan=fi>, viewed: 14/5/2010.
- SYKE (2006), Suomen ympäristökeskuksen raportteja 14/2006: Ympäristönsuojelulain mukaiset valtion ympäristölupapäätökset 2005. Available on the Internet at <http://www.ymparisto.fi/download.asp?contentid=59423>, viewed: 15/8/2010.
- Ympäristöhallinto (2010), website of Finland's environmental administration. <http://www.ymparisto.fi/default.asp?node=1445&lan=fi> , viewed: 14/5/2010.

Minority Protection Doctrines: From Company Law and Equity to Strata Title

Rachel Leow¹

¹Faculty of Law
National University of Singapore
Singapore

Email: rachel.leow@nus.edu.sg

Abstract:

This paper explores the application of minority protection doctrines in company law and equity to strata title developments in Singapore and New South Wales. As majority rule is used in matters of governance in both companies and strata title developments, there is a corresponding need for minority protection which is particularly evident in strata title because of the social and economic impact of the home. While current Australian case-law suggests that the doctrine of 'fraud on the minority' is applicable to strata title, this paper demonstrates that this view is superficial and conceptually inadequate. Instead, this paper seeks to demonstrate that the doctrine of fraud on a power should be accepted instead as being more conceptually rigorous and practically useful. It concludes that fraud on a power has value in supplementing the remedies available to minority subsidiary proprietors under statute to provide more comprehensive protection for minority subsidiary proprietors.

Keywords:

Company law, Equity law, Minority protection, Property law, Strata title

1 Introduction

The law of strata title is becoming of increasing importance in today's world as existing land is subject to much competition for use by offering another way to maximize land use by partitioning the horizontal division of airspace. While both companies and strata title utilize majority rule for governance, little consideration has been given to the applicability of company law minority protection doctrines to strata title. The sparse current academic work on this area has generally accepted Australian authority supporting the proposition that the company law doctrine of 'fraud on the minority' is applicable (Gray et al, 2007; Bradbrook, et al, 2007; Ilkin, 2007; Alford and Sommer, 2005; Tan et al, 2009), or has accepted 'fraud on a power' without justification despite relying on the same Australian authority (Teo, 2009). These accounts are superficial and fail to grapple with the conceptual nature of the body corporate in strata title law as well as considerations peculiar to strata title law. This essay seeks to address this gap by investigating the doctrines of fraud on the minority and fraud on a power, evaluating Australian case-law and then comparing the applicability of the doctrines in strata title from a conceptual perspective in the context of Singapore and the Australian state of New South Wales. These two jurisdictions have been selected because of the complexity and thoroughness of their strata title laws which were necessitated by high

population densities in these two jurisdictions. Close analysis demonstrates that the Australian authorities are muddled and that ‘fraud on the minority’ is conceptually inappropriate in strata title law. Instead, the equitable doctrine of ‘fraud on a power’ has judicial support and is more theoretically defensible and practically effective and thus should be adopted.

2 A Hypothetical Situation

We start with the following hypothetical scenario as a framework for understanding the role and utility of protection of minority doctrines. A strata title development has a clearly differentiated majority and minority, with the majority holding 80% of the total share value and the minority holding 20%. At a general meeting of the management corporation, a resolution is passed to enter into a contract with a related company¹ to carry out certain services for the strata title development, such as renovation works, though the exercise of voting power by the majority, where all the procedural requirements for the calling of the meeting were complied with. Some questions come to mind: can the minority invalidate the resolution, and if so, under what circumstances? Can the minority obtain rescission of the contract if it has already been entered into? Who can bring the action, and who pays the costs of the litigation? Does it make any difference if the contract price was at fair market value and there is no loss to the minority?

In both Singapore and New South Wales, alternative dispute resolution bodies established under statute are unable to provide a remedy as they may not make orders invalidating proceedings if there was no prejudice or if the same result would be obtained.² This hypothetical scenario will be returned to later to verify the solutions proposed in the course of this paper.

2.1 System of Strata Title Ownership

While the precise nature of strata title ownership is determined by the specific wording of the governing statutes in a particular jurisdiction,³ the underlying idea behind strata title ownership is the subdivision of high-rise buildings such that the flats therein can be owned by different legal persons under the Torrens system of land registration.⁴ Upon the registration of the strata scheme, a management corporation is incorporated to manage and maintain the development and common property after it was developed.⁵ Strata title units will have a share value or unit entitlement which will determine, *inter alia*, the voting rights of the proprietor during special resolutions⁶ and the amount of contributions levied by the management

¹ See s5-6, Singapore Companies Act (Cap. 50, 2006 Rev. Ed.) [“Companies Act”]; Part 1.2, Division 6, Australia Corporations Act 2001.

² Section 103, BMSMA (Singapore); Section 153, SSMA (New South Wales).

³ The main statutes governing strata title are the Land Titles (Strata) Act (Cap. 158, 1999 Rev. Ed.) [“LTSA”] and the Building Maintenance and Strata Management Act (Cap 30C., 2008 Rev. Ed.) [“BMSMA”] in Singapore and the Strata Schemes Management Act 1996 [“SSMA”] in New South Wales.

⁴ Section 2 of the LTSA and section 46(1), Land Titles Act (Cap. 157, 2004 Rev. Ed.) (Singapore), ss6(1) and 7(1) of the New South Wales Strata Schemes (Leasehold Development) Act 1973 (New South Wales)

⁵ Section 10A(1), LTSA and Section 24(1), BMSMA (Singapore); sections 8-9, SSMA (New South Wales)

⁶ Section 30(2)(a), LTSA and Section 62(1)(a) BMSMA (Singapore); section 17-18 of Schedule 2, SSMA (New South Wales)

corporation.⁷ The primary duty of the management corporation is the management of the development,⁸ though in practice, some or all of the management corporation's powers and duties are delegated to a managing agent or manager.⁹ The use and enjoyment of the strata title development is governed by the by-laws of the development.¹⁰ Majority rule plays an important role in the management of the strata title development as it is the primary basis on which resolutions are passed at general meetings, the members of the management council or executive committee elected¹¹ and the by-laws enacted.¹² Under the current statutory framework, it is not possible simply to use the provisions applying to companies to the bodies corporate established by strata title legislation as the provisions of the Singapore Companies Act or Australia Corporations Act are either expressly¹³ or impliedly excluded.¹⁴

Both Singapore and New South Wales have specialized dispute resolution methods for strata title developments, namely the Strata Titles Boards¹⁵ in Singapore and the Strata Schemes Adjudicator and Consumer, Trade and Tenancy Tribunal in New South Wales.¹⁶

3 Rationale for Majority Rule

3.1 Company Law

Under company law, the default method of governance is by majority rule through an exercise of voting power by the members¹⁷ at general meeting. Although the day-to-day management of the company is performed largely by the directors, directors are typically also elected by the members in general meeting. The principle of majority rule is supported largely from a contractarian perspective where minorities are theoretically minorities through an exercise of free will - as minorities choose their extent of risk exposure in the corporate enterprise, they should be content with the corresponding level of clout. Majority rule also theoretically encourages profit maximization as the majority has the greatest incentive to prioritize profit-making as an objective of the company.

However, the contractarian view of corporate membership is inaccurate as minorities may be minorities due to external factors, like a lack of financial means to purchase a larger shareholding, a lack of available shares on the open market for purchase, and restrictions on transfer of shares under the company's memorandum and articles. Also, the majority's interest in profit-maximization is probably for the majority themselves and not the company.

⁷ Section 30(2)(c), LTSA and Section 40(2) BMSMA (Singapore); section 78(2) SSMA (New South Wales)

⁸ Section 29, BMSMA (Singapore); Ch 3, SSMA (New South Wales)

⁹ Section 66-68, BMSMA (Singapore); section 9(b) of the SSMA (New South Wales)

¹⁰ Section 32, BMSMA (Singapore); Ch 2 Part V, SSMA (New South Wales)

¹¹ Section 53(1), BMSMA (Singapore); Section 16, SSMA (New South Wales)

¹² Section 32(3), BMSMA (Singapore); Chapter 2, Part 5, Division 3, SSMA (New South Wales)

¹³ Section 11(2), SSMA (New South Wales)

¹⁴ In Singapore, the exclusion of the Companies Act appears to be implied by section 24(1), BMSMA.

¹⁵ Governed by Part VI of the BMSMA.

¹⁶ Chapter 5, SSMA (New South Wales)

¹⁷ There is a conceptual distinction between membership and holding shares in the company, though in the vast majority of cases a shareholder will also be a member. The words 'member' and 'shareholder' are used interchangeably in this essay to refer to members. A shareholder who is not a member is not entitled to exercise membership rights, including the right to vote (Tan, 2009).

The strongest reason for majority rule is simply one of pragmatism in enhancing administrative efficiency and allowing for timely corporate action to be taken when necessary (Chew, 2007). In situations involving a grouping of persons, differences of opinion will inevitably arise and some mechanism has to be in place to ensure that the company is not paralyzed by disagreements between different individuals (Koh, 2009). Majority rule thus has value as a useful practical business administration mechanism, although its juridical basis is left in some doubt.

3.2 Strata Title

The contractarian perspective is usually less dominant in strata title because it is impractical and unrealistic for the average individual homeowner to purchase multiple units in a development simply to ensure that his rights are protected from abuse by the majority. Furthermore, the contractarian theory is flawed even in the case of a company where the interests are purely commercial. Instead, the focus should be on majority rule as a practical means of ensuring that the management corporation can function smoothly in carrying out its obligations with minimal delay in ensuring the due administration, control and maintenance of the common property (Tan, 2002).

4 Rationales for Minority Protection

Majority rule is an efficient governance mechanism but formulaic application of majority rule will inevitably cause injustice in some situations. The rationale for minority protection is difficult to ascertain, but vague notions of fairness or justice should be avoided for their lack of clarity.¹⁸ While both company law and strata title law share certain grounds justifying minority protection, the home ownership context has a stronger claim for minority protection.

4.1 Company Law

The protection of valuable proprietary rights is an important factor in protection of minority rights. While shareholders do not have any legal or beneficial interest in the property of the company,¹⁹ shares in the company are movable property and confer rights on the shareholder, in particular the right of participation on the terms of the memorandum and articles of association and in accordance with general law and statute (Tan, 2009). Some of these rights associated with membership are proprietary ones and deserve protection in line with the common law's traditional emphasis on the importance of property. In addition, protection of the shareholder's property rights also protects the shareholder's stake in the corporate enterprise. The right to minority protection is relatively weak compared to that of strata title because of a deep-rooted reverence for majority rule as the norm for dealing with intra-corporate affairs in light of the main role of the company as a profit-making enterprise. This is reinforced by the general policy of judicial non-interference with business decisions.²⁰

¹⁸ Eg. Chew, 2007 at 7 ("Where majority rule results in unfairness or injustice, it is submitted that it is natural, for want of a better description, to assert that the minority has higher rights to justice and fair treatment that will take precedence, although the basis of such an assertion often remains conveniently undetermined.")

¹⁹ *Macaura v. Northern Assurance Co Ltd* [1925] AC 619

²⁰ *Vita Health Laboratories Pte Ltd and others v Pang Seng Meng* [2004] 4 SLR(R) 162 at [17]; *ECRC Land Pte Ltd v Ho Wing On Christopher* [2004] 1 SLR(R) 105 at [49]; *Intraco Ltd v Multi-Pak Singapore Pte Ltd* [1994] 3 SLR(R) 1064 at [30]

4.2 Strata Title

In contrast to company law, strata title has greater justification for the intervention of the courts in minority protection. Similarly to company law, the proprietary angle in strata title is one of the strongest reasons for minority protection. The traditional conception of land was that it was the sole and despotic dominion which one man claims and exercises over the external things of the world, in total exclusion of the right of any other individual in the universe. While this conception has been overtaken by the concept of property as a bundle of rights rather than absolute ownership (Gray and Gray, 2009), the subsidiary proprietor is still the sole proprietor of his own lot and as such has the right to use it as he pleases, subject to statute and the by-laws of the estate (Tan, 2002). Subsidiary proprietors also hold the common property as tenants-in-common.²¹ The concept of a man's home as his castle has been eroded statutorily by allowing for termination of the strata scheme by non-unanimous collective sales in Singapore²² and compulsory acquisition of land (Tang, 2007).²³ However, these are limited statutory exceptions and do not warrant withholding minority protection during the continuance of the strata development.

There are also wider socio-political implications to home ownership. An individual's personhood may become intimately bound up with his or her proprietary possessions (Radin, 1982). Homes satisfy social and psychological needs, acts as a physical framework for family life, is a place of self-expression and provides a feeling of security (Fox, 2002). Home ownership is also seen as a social and cultural unit (Fox, 2002), while being property with commercial value. Home ownership thus has a wider social and political aspect than companies which are focused on commercial profit, justifying greater minority protection in strata title.

5 Fraud on the minority

5.1 Origin, Function and Rationale

Fraud on the minority is also known as the common law derivative action²⁴ and is actually a development in equity; the reference to 'common law' distinguishes it from the statutory derivative action.²⁵ It originated under the court's inherent equitable jurisdiction with respect to civil procedure in a piecemeal fashion. As a derivative action, it is pursued by a shareholder of a company but in respect of loss suffered by the company or an infringement on the company's rights rather than damage suffered by the shareholder himself. The shareholder's entitlement to bring a derivative action is therefore derived from the company's right of action; this is merely a procedural device and is not a new substantive cause of action

²¹ Section 13(1), LTSA (Singapore); *Houghton v Immer (No 155) Pty Ltd* [1997] NSWSC 608 (New South Wales)

²² Part VA, LTSA (Singapore).

²³ See Section 5 of the Land Acquisition Act (Cap. 152, 1985 Rev. Ed) (Singapore); Section 41, Lands Acquisition Act 1989 (No. 15 of 1989) (Australia)

²⁴ Despite the common misuse of 'fraud on the minority' to refer interchangeably with the statutory ground of oppression found in Section 216, Singapore Companies Act and Section 232, Australia Corporations Act, this will be avoided in this paper to prevent conceptual confusion. While the common law derivative action is merely a procedural tool, oppression is a substantive cause of action.

²⁵ Section 216A, Singapore Companies Act; Part 2F.1A, Australia Corporations Act

conferring new rights. This was developed as an exception to the proper plaintiff rule in *Foss v Harbottle*²⁶ that that the proper plaintiff in an action in respect of a wrong alleged to be done to a company is *prima facie* the company. The issue of whether the plaintiff-member has *locus standi* to bring the action is to be determined as a preliminary issue prior to the main trial.²⁷ It follows from the nature of the common law derivative action as a procedural device that establishing the ground of fraud on the minority only confers *locus standi* on the member who then has to raise a substantive cause of action as a ground for obtaining relief from the company.

5.2 Application in Company Law

The starting point is that the phrase ‘fraud on the minority’ has elucidated no authoritative exposition and has been described as a term of art,²⁸ although it is clear that ‘fraud’ is used in its wider equitable sense;²⁹ it is not limited to common law fraud requiring an element of dishonesty.³⁰ While it is possible to compile a list of examples when fraud on the minority would arise, these discrete categories do not provide much guidance as principles for novel situations. For example, it has been suggested that appropriation of the company’s money, property or opportunities, the majority obtaining a benefit at the expense of the company or the majority preventing an action being brought are groups of cases where fraud on the minority has been found (Tan, 2009). However, these categories do not appear to be at all similar in formulation at first glance.

The common law derivative action suffers from fundamental conceptual uncertainty, as well as a lack of a clearly defined procedure and inherent uncertainty as a matter of equitable discretion (Tan, 2009). It is unclear whether the focus is on act of wrongdoing instead of the effects of wrongdoing as experienced by the company. The leading Singaporean text supports the latter view by formulating a three-limbed test that the majority must have obtained some sort of benefit; the benefit was obtained at the expense of the company or that some loss or detriment was caused to the company; and that the majority used their controlling power to prevent an action being brought against them in court (Tan, 2009; Chew, 2007). However, case-law in Singapore appears to focus on the act of wrongdoing instead, as seen by the adoption of a test where it is necessary to show wrongdoing, majority control or being in the ‘seat of power’ and an exercise of that control or power to prevent an action being taken by the company.³¹ In conclusion, the common law derivative action “may be treated as moribund (if not actually dead) in Singapore” (Tan, 2009) due to the increasing use of the statutory derivative action and oppression actions. In Australia, the death of the common law derivative action has already occurred when it was statutorily abolished in 2001.³²

²⁶ (1843) 2 Hare 461

²⁷ This is the position in the UK in *Prudential Assurance Co Ltd v Newman Industries Ltd (No 2)* [1982] Ch 204. The *prima facie* approach has also been expressly adopted in Singapore for s. 216A of the Companies Act.

²⁸ *Abdul Rahim bin Aki v Krubong Industrial Park (Melaka) Sdn Bhd* [1995] 3 MLJ 417, cited with approval in *Ting Sing Ning (alias Malcolm Ding) v Ting Chek Swee (alias Ting Chik Sui) and ors* [2008] 1 SLR(R) 197 at [13].

²⁹ *Estmanco (Kilner House) Ltd v Greater London Council* [1982] 1 WLR 2, cited with approval in *Ting Sing Ning (alias Malcolm Ding) v Ting Chek Swee (alias Ting Chik Sui) and ors* [2008] 1 SLR(R) 197 at [13]

³⁰ *Derry v Peek* (1889) 14 App Cas 337

³¹ *Ting Sing Ning (alias Malcolm Ding) v Ting Chek Swee (alias Ting Chik Sui) and ors* [2008] 1 SLR(R) 197 at [13]-[14].

³² Section 236(3), Australia Corporations Act 2001. The remedies now available to a minority shareholder in Australia are an action under the personal rights exception to *Foss v Harbottle* (preserved by Note 3 to s. 236 Corporations Act), a statutory minority shareholders action under s. 232 Corporations Act, a statutory derivative action under Part 2F.1A, Corporations Act and an application for winding up.

6 Fraud on a power

6.1 Origin, Function and Application

The history of the equitable doctrine of fraud on a power is distinguished, originating in the middle of the 18th century (Matthews, 2007). The doctrine was developed in cases relating to the exercise of special powers of appointment, but there is little doubt that other powers now fall within the scope of the doctrine (Thomas, 1998). The doctrine is part of a regulatory web of principles imposed by equity on the exercise of powers (Matthews, 2007) and applies to both fiduciary powers and non-fiduciary powers.

The doctrine of fraud on a power will apply when the power has been exercised for a purpose, or with an intention, beyond the scope of or not justified by the instrument creating the power.³³ Fraud here does not require fraud at common law or any conduct that could be described as dishonest or immoral.³⁴ It is necessary to distinguish between the existence and scope of a particular power from the purpose for which it was exercised as fraud on a power does not capture situations where the exercise of power exceeded the limits of the power granted; such situations are instead captured by the *ultra vires* doctrine (Thomas, 1998).

6.2 Application in Company Law

While company law does not often overtly employ the doctrine of fraud on a power, it can be seen as underpinning various principles which have developed to restrict both shareholder and director action. These principles are better seen as examples illustrating the general doctrine rather than as disparate free-standing doctrines of their own.³⁵ Cases where fraud on a power have been frequently invoked fall into three broad categories: the appropriation of rights, advantages and property belonging to the company; expropriation of rights, advantages and property belonging to other shareholders; and ratification of breach of duties on the part of directors. This list is non-exhaustive and merely illustrative of fraud on a power (Loh, 1996).

One of the clearest examples of such a principle is the proper purposes doctrine under which company directors must exercise their powers for proper purposes. Powers vested in company directors are no different from those which trustees possess and must be exercised only for the purposes for which they were granted (Mok, 2002). They are fiduciary powers conferred by the office of directorship rather than a bare power which is conferred on an individual in a personal capacity (Thomas, 1998). The proper purposes doctrine is now an independent ground of liability for directors from the duty to act *bona fide* in the best interests of the company. The better view is that the proper purposes doctrine embodies a general equitable obligation owed to persons who might suffer detriment from an improper exercise of power rather than a fiduciary duty owed solely to the company (Mok, 2002). In fact, the proper purposes doctrine can be seen as being a particular illustration of the doctrine of fraud on a power.

³³ *Vatcher v Paull* [1915] AC 372 at 378; *Wong v Burt* [2005] 1 NZLR 91 at 12-13.

³⁴ *Vatcher v Paull* [1915] AC 372

³⁵ There is some limited support for this proposition in the authorities. See e.g. Teo, 2009 at p 433, fn 259: "The application of [the doctrine of fraud on a power] to the exercise of powers conferred on a majority of shareholders at a general meeting is well-established."; Loh, 1996 at pp 97-123, which clearly links fraud on a power to limitations on the exercise of voting power by shareholders. Note, however, that for the latter, the discussion of 'fraud on the minority' appears to be used descriptively and interchangeably with fraud on a power, rather than the strict technical meaning of a common law derivative action.

Another area in company law where fraud on a power can be seen is in the alteration of the articles of association which may generally be effected through a special resolution.³⁶ A member's voting rights are property rights and generally may be exercised by the member as he thinks fit in his own interest, as the member owes no duty to the company or to anyone else when he exercises his voting rights. Courts generally will interfere more readily with an alleged abuse of directors' powers than with an alleged abuse of majority shareholder voting power³⁷ because directors' powers are fiduciary while the right to vote is proprietary (Loh, 1996). However, the courts have intervened to limit the power of the special majority to alter the articles of association in certain circumstances, though the exact basis on which they have done so is often unclear, demonstrating that there are still limits placed on a shareholder's rights to vote. The widely accepted position is that the shareholders have to vote '*bona fide* for the benefit of the company as a whole',³⁸ though this has also been criticized for not providing sufficient guidance as to when and on what basis courts will intervene to restrict the rights of the special majority to vote.³⁹ Some attempts to reformulate this test by defining the company to refer to the corporators as a general body⁴⁰ or to replace the test⁴¹ have been criticised (Chew, 2007; Davies, 2008). Notably, in *Gambotto v WCP Ltd*,⁴² the High Court of Australia has rejected the '*bona fide* for the benefit of the company as a whole' test where a distinction was made between cases depending on the existence of an expropriation of valuable proprietary rights. Where the resolution was to amend the articles other than to allow for a power to expropriate proprietary rights, resolutions to alter the articles would be valid unless they were, *inter alia*, beyond any purpose contemplated by the articles. However, a resolution to amend the articles to provide for the power to expropriate could only be allowed if the power was exercisable for a proper purpose and if its exercise would not operate oppressively.⁴³ This area can be better viewed as an example of the doctrine of fraud on a power, with the conventional '*bona fide* for the benefit of the company as a whole' test being the definition of what is a proper purpose. The Australian approach also supports the doctrine of fraud on a power since the exercise of the power to amend the articles will be restrained when it was for an improper purpose, whether or not such exercise was related to the expropriation of proprietary rights.

The preceding discussion demonstrates that the concept of fraud on a power has well-established application to restraining action by the directors and shareholders in general

³⁶ Section 26(1), Singapore Companies Act. An exception is where the articles of association are subject to entrenching provisions; see section 27, Singapore Companies Act.

³⁷ *Ngurli Ltd & Anor v McCann* (1953) 90 CLR 425 at 439.

³⁸ *Allen v Gold Reefs of West Africa Ltd* [1900] 1 Ch 656

³⁹ *Minority Shareholders' Rights and Remedies* at 64-65

⁴⁰ *Greenhalgh v Arderne Cinemas Ltd* [1951] Ch 286 per Lord Evershed (Court of Appeal, UK). Cf. *Peter's American Delicacy Co Ltd v Heath* (1939) 61 CLR 457 (High Court, Australia) per Dixon J: "To say that the shareholders forming the majority must consider the advantage of the company as a whole in relation to such a question seems inappropriate, if not meaningless, and at all events starts an impossible inquiry. The 'company as a whole' is a corporate entity consisting of all the shareholders".

⁴¹ *Greenhalgh v Arderne Cinemas Ltd* [1951] Ch 286 at 291 per Lord Evershed: "That is to say, the case may be taken of an individual hypothetical member and it may be asked whether what is proposed is, in the honest opinion of those who voted in its favour, for that person's benefit. I think that the matter can, in practice, be more accurately and precisely stated by looking at the converse and by saying that a special resolution of this kind would be liable to be impeached if the effect of it were to discriminate between majority shareholders and the minority shareholders, so as to give the former an advantage which the latter were deprived".

⁴² [1995] HCA 12 ["*Gambotto*"].

⁴³ *Gambotto* at [25]-[26] per Mason CJ, Brennan, Deane and Dawson JJ. The court continued to hold that "an expropriation may be justified where it is reasonably apprehended that the continued shareholding of the minority is detrimental to the company, its undertaking or the conduct of its affairs - resulting in detriment to the interests of the existing shareholders generally - and expropriation is a reasonable means of eliminating or mitigating that detriment."

meeting and has outgrown its origins in discretionary trusts. The time is ripe for fraud on a power to be considered as a method for restraining majority power in strata title, similar to the role it plays in company law.

7 The Australian Experience

The Australian experience so far has been instrumental in recognizing the applicability of company law doctrines in strata title. However, the current jurisprudence on the application of minority protection doctrines is unclear and theoretically confused. While the authorities have ostensibly applied the ‘fraud on the minority’ or the ‘expropriation of property’ grounds in vindicating the interests of the minority, it is clear upon close analysis that they were actually applying the ‘fraud on a power’ ground. Thus, there is judicial support for the applicability of the doctrine of ‘fraud on a power’ in strata title.

7.1 Influence of Company Title Home Ownership

Prior to the introduction of strata title legislation in New South Wales, company title was the most common way of accommodating the horizontal division of airspace in Australia.⁴⁴ Under the company title scheme, entitlement to live in a residential home unit building is through share acquisition in a company which owns the building. The purchase of shares in the company would usually give the purchaser only a personal right against the company, though a licence or a lease may also be granted by the company to allow the purchaser the exclusive use and occupation of a specified part of the building and the right to use other parts of the property in common with other shareholders. Company title ownership has significant disadvantages, notably a lack of security of separate title and ownership and the possibility of the shareholder being deprived of the right of occupation of the unit in event of certain breaches if the articles of association are altered.⁴⁵

Company title ownership may be seen as a middle ground between company law and strata title, with the strata title company being run and managed similarly to other commercial companies, but where the interests at stake are those in home ownership. The past use of the corporate structure in home ownership is likely to have influenced courts in Australia to be more willing to apply company law principles in home ownership.

7.2 Confusion: Fraud on the Minority and Fraud on a Power

The Supreme Court of New South Wales Court of Appeal held in the landmark case of *Houghton & Anor v Immer (No. 155) Pty Ltd*⁴⁶ that the doctrine of fraud on the minority was of general application in equity and could apply to strata title ownership, notwithstanding the existence of statutory provision governing strata title ownership generally. *Houghton* concerned the proprietors of a strata title plan which held 80% and 20% of the aggregate unit

⁴⁴ LexisNexis, *Halsbury's Laws of Australia* (at 9 June 2010) 355 Real Property ‘IV Multiple Ownership’, [355-9010] (online); New South Wales Law Reform Commission, *Report 115 (2007) - Disputes in Company Title Home Units*, online: <http://www.lawlink.nsw.gov.au/lawlink/lrc/ll_lrc.nsf/pages/LRC_r115report> at [2.1]. [“Report on Disputes in Company Title Home Units”]

⁴⁵ LexisNexis, *Halsbury's Laws of Australia* (at 9 June 2010) 355 Real Property ‘IV Multiple Ownership’, [355-9010] (online); *Report on Disputes in Company Title Home Units* at [2.5] and [2.8].

⁴⁶ [1997] NSWSC 608 [“*Houghton*”]. Special leave to appeal the decision to the High Court of Australia was denied in *Houghton and Anor v Immer (No. 155) Pty Ltd and Anor* [1998] HCATrans 273.

entitlement respectively. The appellant majority proprietors used their voting power at a series of general meetings of the body corporate (equivalent to the management corporation in Singapore) to pass resolutions authorizing them to construct 2 penthouses in a lot and the roof, enacted a by-law authorizing the construction and allowing the appellant to retain the consequent improvements, and finally authorized the body corporate to transfer its interest in the penthouse lots derived from the use of the roof as common property for \$1. The authorization of the transfer of property for \$1 was held to be a fraud on the minority as the doctrine of fraud on a power would apply to the powers of the subsidiary proprietor (or the shareholder in the company context) exercisable at general meetings. In Singapore, this situation could be caught by s. 36 of the BMSMA which gives dissatisfied subsidiary proprietors redress pertaining to, *inter alia*, the transfer of common property.

In the subsequent case of *Young and Ors v The Owners – Strata Plan 3529 and Ors*,⁴⁷ the plaintiffs owned a lot in the strata plan which conferred a right to two car parking spaces in an adjoining property but not to any residential rights therein. The defendant body corporate foreshadowed the passing of an exclusive use by-law which would permit the use of a swimming pool and common property only to those who held residential rights in the adjoining property, thus excluding the plaintiffs. The court held that if express consent was required to effect such an extinction of the rights of the plaintiffs, the absence of such consent would make the extinction of rights a fraud against the minority, being a fraud on the power to vote at general meetings and the by-law would hence be invalid. In Singapore, written consent by the subsidiary proprietor and an appropriate resolution is also necessary to pass exclusive use by-laws.⁴⁸ There is no statutory recourse in such a situation in Singapore since section 36 of the BMSMA only deals with transfers of common property, transfers of land to become common property, and amalgamation of the common property of two management corporations.

Further confusion was demonstrated in *Radford v The Owners of Miami Apartments, Kings Park Strata Plan 45236*,⁴⁹ which followed *Houghton* and *Young*. The learned judge considered that fraud on the minority operated such that the interests of the minority subsidiary proprietors qualified the power of alteration of the by-laws. He then continued to describe the juridical basis behind fraud on the minority as fraud on a power, which was then correctly recognized to be a general rule relating to the construction of powers. The discussion in *Radford* again shows the lack of distinction between the doctrines of fraud on the minority and fraud on a power and the lack of clarity in conceptual reasoning between the two.

Overall, current Australian case-law is useful to the extent that it shows us the willingness of the courts to intervene to restrain improper exercises of voting power by the majority in general meeting. However, the implementation of company law doctrines has been unsatisfactory. The loose usage of the phrases ‘fraud on the minority’ and ‘fraud on a power’ lead to uncertainty as to the exact legal principles involved. This uncertainty is particularly troubling when established doctrines from one field of law are transplanted into another, as any differences in application may be due either to misunderstanding of the doctrines themselves or to legitimate adaptation necessary to fit the contours of the recipient. As earlier discussed, fraud on the minority refers to a common law derivative action which operates solely as a procedural device to give the minority *locus standi*, while fraud on a power is an

⁴⁷ [2001] NSWSC 1135 [“*Young*”].

⁴⁸ Section 33, BMSMA (Singapore).

⁴⁹ [2007] WASC 250 [“*Radford*”].

equitable substantive cause of action which can be relied on in event of an improper exercise of the power in question. Yet, a plain reading of the cases demonstrates the two doctrines being discussed interchangeably. In *Houghton*, the discussion of the issue of fraud on a power both commenced and concluded that the resolution authorizing the transfer of common property for \$1 was a fraud on the minority. In *Young*, the court held that the doctrine of fraud on the minority was applicable and the use of a by-law to expropriate the minority's property would be a fraud on the power, being from an improper purpose. Given the conceptual differences between the two doctrines and their practical implications, it is submitted that *Houghton*, *Young* and *Radford* are not strong authority for the application of the fraud on the minority ground in the sense of a common law derivative action to strata title developments. On the contrary, the substantive discussion of the doctrines in *Houghton* and *Young* suggest that the courts were in substance applying the doctrine of fraud on a power, since most of the discussion centered on fraud on a power and none of the pre-conditions for fraud on the minority were discussed. In summary, it appears that judicial support in substance swings in favour of the doctrine of fraud on a power.

7.3 Further Confusion: Expropriation of Property

Another alternative ground raised in the cases is the expropriation of property by the majority through a resolution of the company or body corporate. In Australia, the principles relating to the expropriation of property or proprietary rights are found in the High Court of Australia decision of *Gambotto* where it was proposed that the articles of association of the company be altered to allow the shareholder holding 90% or more of the issued shares to compulsorily acquire the shares of the minority owners. The proposed article was held to be invalid.

In *Young*, the New South Wales Supreme Court held that the passing of a by-law which conferred exclusive use of common property on residential lot owners constituted an expropriation of the plaintiffs' proprietary interest as tenant-in-common of the common property under the *Gambotto* principle, hence invalidating the by-law. This finding was in alternative to the ground of fraud on the minority, which has been discussed earlier. *Young* was followed by *Lin and Anor v The Owners – Strata Plan 50276*,⁵⁰ where the body corporate rejected applications to connect improvements in the plaintiffs' shops to the exhaust ventilation system, drains and waste lines, which were common property. The body corporate argued that the exhaust ventilation system was incapable of servicing the plaintiffs' shops because it was overloaded, but this did not justify their rejection of the applications because they had the statutory duty to renew or replace common property.⁵¹ In *Lin*, the court applied the *Gambotto* principle in the present case where the plaintiffs' equitable rights were destroyed by the defendant's refusal to allow the applications. *Young* was also considered in *Radford*, where the learned judge accepted as the law that the doctrine of applying to resolutions authorizing the expropriation of valuable proprietary rights associated with shares appeared to be capable of separate application to equitable interests apart from the discussion of fraud on the minority and fraud on a power. Confusingly, *Radford* later held that a resolution authorizing expropriation of valuable proprietary rights was recognized as a form of fraud on the minority in *Young*.

Once again, the discussion of the *Gambotto* principle is confused. It is often unclear whether the *Gambotto* principle is a separate ground or a subset of a wider doctrine. The confusion

⁵⁰ [2004] NSWSC 88 [“*Lin*”].

⁵¹ Section 62(2), SSMA (New South Wales). The equivalent in Singapore is section 29(1)(b) of the BMSMA.

relating to fraud on the minority has been exacerbated by the further grafting of the *Gambotto* principle onto an already conceptually unstable base. It has been demonstrated earlier that the *Gambotto* principle dealing with expropriation of property rights can be and is better seen as a specific application of the general doctrine of fraud on a power in the company law context. It then follows that the line of cases in strata title law which have relied on the *Gambotto* principle support the doctrine of fraud on a power in controlling the exercise of powers by the majority. Reducing the principles and understanding them as specific applications of a broad theory has the benefit of clarity and consistency in reasoning and should be preferred.

8 The Better Approach: Fraud on a Power

8.1 The Illusory Problem of *Foss v Harbottle* and *Locus Standi*

The most vexed issue is also the preliminary one – who has *locus standi* to bring the action? Yet, the issue of *locus standi* is in fact the very reason why the doctrine of fraud on the minority is unnecessary in strata title law. This is due to conceptual differences between the legal structure of the company and management corporation and their relative strength of separate legal personality.

Both the company and the management corporation are corporate bodies established under statute which have formal separate legal personality status and are regarded as distinct legal individuals. The company has relatively strong separate legal personality status. The main consequences of incorporation are the existence of a body corporate, the capacity to sue and be sued, perpetual succession, power to hold land and liability on the part of the members to contribute to the assets of the company in event of winding up to the extent provided by statute.⁵² The separate legal personality doctrine has been affirmed from the pedigree of *Salomon v A Salomon & Co Ltd*⁵³ to the effect that the company and its members are regarded as legally distinct. There remain some exceptional circumstances where the court has ‘pierced the corporate veil’ to disregard the separate legal personality of the company from its shareholder or directors, though the precise boundaries of when the court can do so are hotly debated. As an artificial juristic person which must act through its constituent organs,⁵⁴ the company’s interests are not clearly aligned with any particular group of interests, except potentially when the company is nearing insolvency.⁵⁵

In contrast, the management corporation has much more limited separate legal personality status, despite being bodies corporate with perpetual succession and legal capacity to sue and sued.⁵⁶ Firstly and most fundamentally, the management corporation is comprised of the

⁵² Section 19(5), Companies Act (Singapore) and section 119, Corporations Act 2001(Australia) which incorporates the common law of corporations except in so far as it has been modified by the Corporations Act: Austin & Ramsay, ‘Ford’s Principles of Corporations Law’, 13th ed, (Australia: LexisNexis Butterworths, 2007) at [4.020]-[4.080].

⁵³ [1897] AC 22. The doctrine has been repeatedly affirmed; see e.g. *Bowman v. Secular Society Ltd* [1917] AC 406; *Macaura v. Northern Assurance Co Ltd* [1925] AC 619; *Re Application by Yee Yut Ee* [1977-1978] SLR(R) 490; [1978] SGHC 32.

⁵⁴ *Sir James Smith’s case* (1691) Carth 217; 90 ER 730; *HL Bolton (Engineering) Co Ltd v TJ Graham & Sons Ltd* [1957] 1 QB 159

⁵⁵ Nearing insolvency, the company’s interests may be identified with creditors’ interests: *Walker v Wimbourne* (1976) 137 CLR 1; *Yukong Line Ltd of Korea v Rendsburg Investments Corp of Liberia (No 2)* [1998] 1 WLR 294. Creditors’ interests may even become dominant: *W & P Piling Pte Ltd v Chew Yin What* [2007] 4 SLR(R) 218.

⁵⁶ *RSP Architects Planners & Engineers v Ocean Front Pte Ltd and anor appeal* [1995] 3 SLR(R) 653; [1995] SGCA 79 at [12] (Singapore); Section 11, SSMA (New South Wales)

subsidiary proprietors of all lots comprised in the strata title plan.⁵⁷ This shows that any separation of the interests of subsidiary proprietors and of the management corporation is illusory despite the presence of the management corporation's formal separate legal personality. Secondly, the strong connection between the interests of the management corporation and those of subsidiary proprietors is reinforced from the 'reverse' representative action under statute.⁵⁸ Such representative actions allow the management corporation to take proceedings on behalf of some or all of the proprietors without requiring the satisfaction of any other requirements such as good faith or that the action would be in the best interests of the corporate body, in contrast to the statutory derivative action under company law.⁵⁹ Also, the management corporation has an unlimited liability structure as the subsidiary proprietors comprising the management corporation guarantee debts of the management corporation lawfully incurred in the course of the exercise of any of its powers or functions or the carrying out of its duties or obligations.⁶⁰ The 'flow through effect' of legal liability from the management corporation to the subsidiary proprietors further demonstrates the lack of an actual separation of legal interests between the two.⁶¹ Lastly, the management corporation has limited powers and functions which are confined by statute to certain restricted circumstances.⁶² All these factors in totality demonstrate that the management corporation's legal personality is relatively weak despite its formal existence and that the interests of the subsidiary proprietors are so intimately related to those of the management corporation's that there is no real distinction between the two.

The proper plaintiff rule under *Foss* and its exceptions envisage a situation where it is possible to clearly differentiate loss between the company and a minority shareholder such that the company should sue for the former. However, it is difficult to draw this line in strata title law. As the sum total of all the subsidiary proprietors, the management corporation would suffer loss whenever a minority proprietor has suffered loss. If the loss between the management corporation and a subsidiary proprietor cannot be meaningfully distinguished, there is no room for operation of the general rule under *Foss* or of the true exceptions to *Foss*. In fact, the *Foss* principle would conflate with the personal injury 'exception', which comprises various causes of action under general law, including fraud on a power. There is hence no room for the operation of fraud on the minority or the doctrine of *Foss* in strata title and the subsidiary proprietor should be able to sue under actions at general law, including fraud on a power. The more conceptually defensible route is to dispense with the procedural tool of fraud on the minority in favour of the substantive action of fraud on a power since the former is superfluous in the context of strata title.

⁵⁷ Section 10A(1), LTSA (Singapore); Section 11(1), SSMA (New South Wales)

⁵⁸ Ss. 85(1)-(2) BMSMA (Singapore); Section 227, SSMA (New South Wales). This is a reverse representative action as the corporate body may represent constituent member(s) in an action. A representative action under company law usually refers to constituent member(s) representing the corporate body. In Singapore, the reverse representative action is often used in actions against developers or architects for defects in the common property; examples of such cases include: *RSP Architects Planners & Engineers v Ocean Front Pte Ltd and anor appeal* [1995] 3 SLR(R) 653; [1995] SGCA 79; *Management Corporation Strata Title Plan No 1938 v Goodview Properties Pte Ltd* [2000] 3 SLR(R) 350; [2000] SGCA 56; *Management Corporation Strata Title Plan No 2297 v Seasons Park Ltd* [2005] 2 SLR(R) 613; [2005] SGCA 16. In New South Wales, actions for structural defects which affect or are likely to affect the support or shelter provided by that part of the building to any other part of the building or its site should be taken under Section 228 of the SSMA instead.

⁵⁹ Section 216A, Companies Act (Singapore); Section 237, Corporations Act (Australia).

⁶⁰ This is clearest in the Singapore provision: Section 44(1), BMSMA. In New South Wales, this appears to be the case pursuant to a court order as to costs in proceedings brought by one or more owners of lots against an owners corporation or vice versa: Section 229, SSMA..

⁶¹ While unlimited companies do exist, but are seldom used in commercial life because limited liability is one of the most attractive features of the corporate structure. Unlimited companies are often incorporated to comply with legislative requirements for specific purposes or functions.

⁶² Section 24(2), BMSMA (Singapore); Section 12, SSMA (New South Wales)

8.2 Application of Fraud on a Power to Strata Title

8.2.1 Supplementary Role in Dispute Resolution

The doctrine of fraud on a power has a supplementary role to the statutory duties and rights provided for in the respective strata title statutes. In Singapore and New South Wales, the rights or remedies which a subsidiary proprietor, mortgagee of a lot, or a management corporation might have under general law in relation to any lot or common property are preserved.⁶³ In *Houghton*, it was held that in light of this provision, the doctrine of fraud on a power is not excluded by the comprehensive provisions of the statute. The same position should be adopted in Singapore as it is clear that the statutory framework seeks to broadly regulate the respective rights and duties in relation to the strata development but not to abrogate additional rights and duties under the common law.

In Singapore, fraud on a power may be raised either before the STB or before civil courts, since the STB does not oust either the jurisdiction of the subordinate courts⁶⁴ or that of the High Court. The position in New South Wales appears similar, as the cases on fraud on the minority and fraud on a power were heard at first instance by the Supreme Court of New South Wales.⁶⁵ Also, an Adjudicator will often have jurisdiction over disputes involving fraud on a power⁶⁶ except, *inter alia*, where the dispute relates to the exercise, or the failure to exercise, a function conferred on an owners corporation which may be exercised only in accordance with a unanimous or special resolution.⁶⁷ Where the Adjudicator does not have jurisdiction over the dispute, the only recourse appears to be to the Supreme Court.

Subsidiary proprietors may prefer proceedings before the STB or the Adjudicator for practical reasons such as lower cost and the greater possibility of peaceful settlement of disputes in line with the spirit of neighbourliness which is surely beneficial since the parties will continue to interact on a regular basis. However, proceedings before the STB or the Adjudicator also have their drawbacks as the STB and Adjudicator have limited jurisdiction.⁶⁸ Even when these bodies have jurisdiction, the remedies available are often less extensive than those at general law under an action of fraud on a power. For example, while STB or the Adjudicator may order the invalidation of any resolution or council election,⁶⁹ no order for invalidation will be made where the failure to comply with the provisions did not prejudicially affect any person and compliance with those provisions would not have affected the result of the election.⁷⁰ A civil action for fraud on a power is thus perfectly compatible with the alternative dispute resolution mechanisms offered by statute in both New South Wales and Singapore as it has a wider scope providing for different remedies

⁶³ Section 123, BMSMA (Singapore); Section 226, SSMA (New South Wales)

⁶⁴ *Teo Kim Hui and Anor v Kwok Wai Hon* [2008] SGHC 232.

⁶⁵ E.g. *Young and Ors v The Owners – Strata Plan 3529 and Ors* [2001] NSWSC 1135; *Lin and Anor v The Owners – Strata Plan 50276* [2004] NSWSC 88; *Houghton & Anor v Immer (No. 155) Pty Ltd* [1997] NSWSC 608

⁶⁶ Section 138(1), SSMA (New South Wales)

⁶⁷ Section 138(3)(c)

⁶⁸ Sections 101-115, BMSMA (Singapore); Chapter 5, Part 4, SSMA (New South Wales)

⁶⁹ Section 103(1), BMSMA (Singapore); Section 153(1), SSMA (New South Wales)

⁷⁰ Section 103(2), BMSMA (Singapore); Section 153(2), SSMA (New South Wales)

8.2.2 *In which situations?*

The clearest situation where fraud on a power could apply is where the majority subsidiary proprietors have misused their voting power, whether to pass resolutions, enact by-laws, make appointments and so on, by exercising the power for a purpose or with an intention beyond the scope of the instrument creating the power. The purpose which the subsidiary proprietors at general meeting should act for should follow the dominant English test for company law, namely whether exercise of the power was *bona fide* for the interests of the management corporation as a whole, as discussed earlier. In strata title, the practical difficulties in company law in determining who the company is in this test are avoided since the interests of the management corporation and the subsidiary proprietors conflate. The Australian approach in *Gambotto* which makes a fundamental distinction based on the expropriation of proprietary rights should be rejected in the strata title context. The majority in *Gambotto* ultimately sought to make a distinction between expropriations aimed at conferring a benefit on the company and those which did not (Davies, 2008). Yet, it is difficult to think of any situation where an expropriation of subsidiary proprietors' proprietary rights will be justified on the ground that the continuing holding of such proprietary rights will be detrimental to the management corporation.⁷¹ A distinction based on expropriation of proprietary rights is unnecessary. Unlike companies where proprietary interests are more limited in scope, proprietary interests are at the core of strata title ownership and most cases involving misuse of voting power will also involve proprietary rights. The better view is that the '*bona fide* for the interests of the management corporation' test should be adopted instead.

8.2.3 *Remedies*

Various remedies are available, depending on the factual scenario and stage at which legal proceedings are instituted. There is a difference in emphasis between remedies under company law and in strata title. In company law, the most frequent remedies are either a forced buy-out of the minority's shares or a winding-up of the company, i.e. a termination of the minority's participation in the enterprise. This is practicable because these are often the only long-term remedies in a situation where the majority and minority are hostile and because corporate participation is essentially for the pursuit of mere financial profit. In contrast, the forced termination of the minority's participation in the strata title development is clearly unpalatable because of the primary aim of home ownership. Minimizing damage to the social fabric by encouraging neighbourliness and restraining improper misuse of power is preferable. Fraud on the power can also pre-emptively deal with misuse of voting power before any damage has occurred by rendering the resolution or by-law void in equity.⁷² As such, it is submitted that fraud on the power can be used to invalidate resolutions passed; to invalidate by-laws; the granting of a mandatory injunction against the passing of a by-law;⁷³ or to obtain rescission of a contract entered into subsequent to a resolution.⁷⁴ In addition, equitable compensation⁷⁵ may be sought. If damages are awarded from the management corporation, the subsidiary proprietors either guarantee the debts of the management

⁷¹ *Gambotto* at [26]

⁷² *Re Cohen* [1911] 1 Ch 37; *Wong v Burt* [2005] 1 NZLR 91. Cf. *Houghton & Anor v Immer (No. 155) Pty Ltd* [1997] NSWSC 608 that such a resolution or by-law would be merely voidable.

⁷³ *Young and Ors v The Owners – Strata Plan 3529 and Ors* [2001] NSWSC 1135

⁷⁴ As suggested in *Houghton & Anor v Immer (No. 155) Pty Ltd* [1997] NSWSC 608, though on the facts of that case, the resolutions had been acted on and the defendants had been able to complete the construction of additional lots and register the lots. It was thus too late to obtain rescission.

⁷⁵ *Houghton & Anor v Immer (No. 155) Pty Ltd* [1997] NSWSC 608

corporation lawfully incurred in the course of the exercise of any of its powers or functions or the carrying out of its duties or obligations⁷⁶, or may be ordered to pay any monies ordered by the courts in actions between subsidiary proprietor(s) and the management corporation.⁷⁷ However, insofar as proceedings by a subsidiary proprietor are successfully brought against the minority, in Singapore, the court has a discretion to specify that any moneys payable are paid by certain lots in certain proportions,⁷⁸ thus avoiding a situation where a subsidiary proprietor would be called to make contributions for damages owed to himself. The position is even clearer in New South Wales, where the management corporation may not levy contributions in respect of such proceedings against the successful subsidiary proprietor.⁷⁹ The doctrine of fraud on a power is thus flexible enough to respond to varied circumstances with a range of appropriate remedies, further demonstrating its superiority over the fraud on the minority ground.

8.2.4 A Hypothetical Situation Revisited

Returning to earlier hypothetical situation posited at the beginning of this paper, fraud on a power can deal with the situation adequately. The minority will be able to invalidate the resolution on proof that when the majority had exercised their voting power in favour of the resolution to grant contracts to a party related to the majority, they had not done so *bona fide* for the interests of the management corporation. It is thus important to examine all the circumstances surrounding the passing of the resolution. Whether the contract price was at fair market value will be an important but not a decisive factor in determining whether the majority had acted *bona fide*. Other factors which should be considered in such a situation include whether the majority has received direct financial gain from their exercise of majority power, whether quotations were obtained from other competitors and considered, and whether the related party who was granted the contract is reputable in carrying out the services contracted for.

In other situations where fraud on a power may apply, different factors should be taken into consideration. For example, where common property was transferred or exclusive use-by laws are passed pursuant to a resolution, the adequacy of compensation and location, type and purpose of the common property in question should be considered. Overall, a sensitive and careful approach should be taken to prevent over-protection of the minority which may hamper the smooth functioning of the management corporation.

9 Conclusion

This essay has attempted to demonstrate the applicability of the doctrine of fraud on a power to strata title developments and its practical and conceptual superiority over the doctrine of fraud on the minority. It has demonstrated that contrary to the general superficial understanding of Australian case-law, a closer analysis reveals conceptual confusion regarding the application of fraud on the minority. A closer examination of both doctrines in light of the unique legal nature of a management corporation yields the conclusion that fraud on a power is the far superior doctrine. While a detailed statutory framework in Singapore

⁷⁶Section 44, BMSMA (Singapore). It is unclear whether this provision would extend generally to *ultra vires* acts as these would obviously not be lawfully incurred in the exercise of the management corporation's powers.

⁷⁷Section 229, SSMA (New South Wales)

⁷⁸Section 87, BMSMA (Singapore)

⁷⁹Section 230(1), SSMA (New South Wales)

and New South Wales is able to provide minority protection in many cases, there are residual situations where fraud on a power has great value in ensuring a comprehensive system of minority protection. It is hoped that this essay has provided a modest starting point by illuminating the conceptual differences between companies and management corporations and the dangers of transplantation of doctrines without careful consideration.

10 References

- Alford, A., Sommer, J. (2005) 'Protection of Minority Owners in a Body Corporate' *Australian Property Law Journal LEXIS*, pp 1-39
- Bradbrook, A., MacCallum, S., Moore, A. (2007) *Australian Real Property Law*, 4th ed Lawbook Co, Sydney
- Chew, M. (2007) *Minority Shareholders' Rights and Remedies*, 2nd ed, LexisNexis, Singapore
- Davies, P. (2008) *Gower and Davies' Principles of Modern Company Law*, 8th ed Sweet & Maxwell, London
- Fox, L. (2002) 'The Meaning of Home' *Journal of Law and Society*, **29**, pp 580-610
- Gray, J., Edgeworth, B., Foster, N., Grattan, S. (2007) *Property Law in New South Wales*, 2nd ed, LexisNexis Butterworths, Australia
- Gray, K., Gray, S.F. (2009) *Elements of Land Law*, 5th ed, Oxford University Press, Oxford
- Ilkin, A. (2007) *NSW Strata and Community Schemes Management and the Law*, 4th ed, Lawbook Co, Sydney
- Koh, P. (2009) *Company Law*, 2nd ed, LexisNexis, Singapore
- Loh, S.C. (1996) *Corporate Powers Controls, Remedies and Decision-Making*, Malayan Law Journal, Kuala Lumpur
- Matthews, P. (2007) 'The Doctrine of Fraud on a Power: Part 1' *Private Client Business*, p 131-142
- Mok, C.Y., (2002) 'Re-thinking Director's Duties: An Analysis of the Proper Purpose Doctrine' *Singapore Academy Law Journal*, **14**, pp 387-417
- Radin, M.J., (1982) 'People and Personhood' *Stanford Law Review*, **34**, pp 957-1015
- Tan, C.H. (gen. ed.), (2009) *Walter Woon on Company Law*, rev. 3rd ed, Sweet & Maxwell, Singapore
- Tan, S.Y. (2002) 'Facets of Communal Living Under the Land Titles (Strata) Act: Common Property, Rights of Subsidiary Proprietors of Individual Lots and the Role of the Management Corporation', *Singapore Journal of Legal Studies*, pp 416-430
- Tan, S.Y., Tang, H.W., Low, K. (2009) *Tan Sook Yee's Principles of Singapore Land Law*, 3rd ed, LexisNexis, Singapore
- Tang, H.W., (2007) 'The Legal Representation of the Singaporean Home and the Influence of the Common Law' *Hong Kong Law Journal*, **37**, pp 81-101
- Teo, K.S., (2009) *Strata Title in Singapore and Malaysia*, 3rd ed, LexisNexis, Singapore
- Thomas, G. (1998) *Thomas on Powers*, Sweet & Maxwell, London

Land Registration and Information Asymmetry Problem in Sub-Saharan Africa Urban Real Estate Markets

Stanislaus Adiaba¹, Felix Hammond¹, David Proverbs¹, Jessica Lamond¹, Colin Booth¹, and Robin Bloch¹

¹School of Engineering and the Built Environment,

University of Wolverhampton, Wolverhampton, WV1 1SB,

United Kingdom

Email: S.Y.Adiaba@wlv.ac.uk; F.Hammond@wlv.ac.uk; D.Proverbs@wlv.ac.uk; J.Lamond@wlv.ac.uk; C.Booth@wlv.ac.uk; R.Bloch@wlv.ac.uk

Abstract

Presently, Sub-Saharan Africa urban real estate markets namely housing, income property and land markets are confronted with deficient information regimes resulting in their underperformance. The enormity of the information problem is known as between 97-99% of the market transactions and relevant information fall outside the radar of formal land registration systems. The sources and causes of the deficient information regime are however, not yet fully understood and the approach to correct the deficiency also remains unknown. The study therefore enquires into the sources and causes of information asymmetry and to measure the impact of the deficient information regime in order to draw attention to the scale of the problem for policy considerations as well as contribute towards bridging the knowledge gap in literature. The study which is ongoing presents its findings on the sources and causes of information asymmetry in Sub-Saharan Africa urban real estate markets based on a thorough literature review. The historical comparison approach is used in tracing the evolution of land registration systems, their trends and influence on information production and use in various land registries. The approach enabled effective comparison of how land registration evolved and why they are relatively successful in Western European countries than in Sub-Saharan Africa. The study finds that in Sub-Saharan Africa, legal, social and technical issues are among the challenges confronting land registries in information production and use. However, legal origins and defective land registration laws are identified as the primary causes and sources respectively of information asymmetry in the real estate markets. The study therefore argues for a need to address the problem and concludes the information problem can be fundamentally addressed through legal reforms that reflect the cultural practices and aspirations of the respective countries and also prescribes the appropriate approach to land registration.

Keywords:

Information asymmetry, Land Information Systems, Land registration, Real estate markets, Sub-Saharan Africa.

1 Introduction

Governments all over the world recognise land information as a public good for which specialised state agencies are responsible for its provision and related services. The responsibilities include production and maintenance of quality land information through land registration systems, facilitating public access and knowledge of the state of titles, charges and encumbrances to real estate property by notice of information (Chatterton *et al.*, 1980; Larsson, 1991; Dale and McLaughlin, 1998; Palmer, 1998; Masser, 1998; Deininger, 2003; Sikor, 2008). These responsibilities are however, not being optimally discharged in most developing countries as compared to their developed counterparts. The effects are huge gaps in information production at the land registries and the problem of information asymmetry among actors involved in real estate transactions.

In Sub-Saharan Africa (SSA), the production and healthy flow of information for formal land delivery processes have not matched the demands of rapid urban growths and real estate markets. From the production side, de Vries (2004) posits that in Sub-Saharan Africa less than 1% of the total land area is administered under any kind of formal land registration system. Toulmin (2008) concurs that presently, the information being produced and certified by various Land Registries for land markets in West Africa region of Sub-Saharan Africa is from only 2-3% of land ownership registered. This implies about 97-98% of required land information is under the control of private actors which brings the legitimacy of the land registration systems into question notwithstanding the effects of the deficient information regime on market transactions, performance and distortions.

Under the present deficient information regime, transactions in Sub-Saharan Africa urban real estate markets are reported fraught with mixed problems of uncertainties, multiple sales, widespread property related disputes, cumbersome procedures for land acquisition, high transaction costs and opportunistic behaviours by information holders (Gough and Yankson, 2000; Kironde, 2000; Hammond, 2006; Hammond and Antwi, 2010). The fundamental causes of the deficient information regime and market underperformance are, however, not yet fully known and how to redress the situation also remains unknown. Nonetheless, various SSA governments are re-engineering their land registration systems in order to provide cheaper, faster, safer and more reliable means of transacting in real estate property.

The study therefore inquires into the sources and causes of information asymmetry in SSA real estate markets with emphasis on urban land markets. It also seeks to measure the impacts of the information problem in order to draw attention to the scale of the problem for policy considerations, and contribute towards bridging the knowledge gap in literature. The study foremost discusses the framework used for thorough literature review. Using the framework, the study from first principles traces the sources and causes of information deficiencies by comparing the formal origins and trends of land information systems in selected Sub-Saharan Africa countries and their Western colonial masters. The comparison is followed by a critical discussion on the legacies of colonial and post-independence laws on land registration and the direction of present land registration reforms. The study further examines their influence on land information production and also, on Sub-Saharan Africa

urban real estate markets. Results on preliminary findings from extant literature reviewed identifying the sources and causes of the information deficient regime in Sub-Saharan Africa then follows. The study concludes based on the literature reviewed with remarks on strategies required to remedy the information asymmetry problem.

2 Framework for Research Review

The review generally covers 21 countries within the six continents of the world using a historical comparison approach to trace how land information systems evolved in the countries reviewed. However, having identified a common trend in the general review of literature, the discussion focuses on SSA countries and their colonial masters for reasons discussed under the heading historical comparison approach.

2.1 The Historical Comparison Approach

The historical comparison approach is mainly adopted to understand the present relations between land information systems and their legal origins and cultures through the past. The past is reviewed with reference to the colonial masters of Sub-Saharan African countries namely Belgium, Britain, Denmark, Germany, The Netherlands, France and Portugal (Feder and Noronha, 1987). The discussion however, focuses on the evolution of land information systems in Britain, France and Germany because of their dominant influence on land tenure and land registration systems in SSA compared to the other colonial masters. Grey literature exists to support the historical comparison approach adopted.

Dale (1997) posits that land registration owns its origin to land tenure and that Africa inherited formal land tenure systems from the colonial masters who fundamentally lacked understanding of indigenous customs and culture. Furthermore, Bertocchi and Canova (2002) elsewhere provide empirical evidence and report that Africa represents an appropriate setting to analyze the impact of colonial rule on growth. Similarly, Beck *et al.* (2003, p. 672) provide empirical results which suggest that “legal origin matters because legal traditions differ in their ability to adjust efficiently to evolving socioeconomic conditions”. Olsson (2009) also shows empirically that there is a strong positive effect of colonial duration on democracy particularly in former British colonies and also, in countries colonised during the imperialist era.

The present is also reviewed with reference to five Sub-Saharan African countries namely Benin, Ghana, Namibia, South Africa and Tanzania with the intent to explain the present relation between the local practices and land information systems, and land registration reforms. The historical comparison approach in studies of land information systems and related fields is not new and has been used by Bogaerts, 1997; Ting *et al.*, 1999; de Soto 2000; Torhonen, 2004; Enemark, 2009. Thus the historical comparison approach is found appropriate in line with the objectives of the study of tracing the sources and causes of information asymmetry in SSA real estate markets.

2.2 Sources of Information for Historical Comparison Approach and Concepts

The review involves information gathered through international land administration systems inventories such as the Cadastral Template (Steudler *et al.*, 2004 and Rajabifard *et al.*, 2007), Inventory of Land Administration Systems in Europe and North America (Manthorpe, 2005) and the Western Humankind – Land Relationship from the tribal period to the Twentieth Century by Ting *et al.* (1999) as guide to the historical comparison. The Cadastral Template provides generic information about cadastral and land administration systems in responding countries (see www.cadastraltemplate.org, Steudler *et al.*, 2004 and Rajabifard *et al.*, 2007). It also reflects the cultural and judicial settings of the respective countries and furthermore provides a worldwide comparison of land registration systems (Enemrk, 2009).

Even though Bandeira *et al.* (2010) comment that the Template does not include a benchmark framework that would allow for obtaining policy implications, the study finds the qualitative and quantitative indicators from countries across five continents provide a very useful descriptive study of the land information systems from the responding countries. North America is not covered by the Template. However, from the Inventory of Land Administration Systems in Europe and North America produced by the HM Land Registry (Manthorpe, 2005), the study draws on the inventory information to cross reference country information provided in the Cadastral Template and also information on North America (See <http://www.unece.org/env/wpla/welcome.html>).

The study further dwells on extensive grey literature on land information systems which was used to corroborate and fill the gaps in the above mentioned sources in the discussions. The study however, observed some key terms could be slippery as they have been variously defined by different writers. Hence those adopted in the study are deemed to encapsulate other definitions, are popular and where necessary modified to accomplish the purpose of the study and also in line with research guide suggestions (Creswell, 2009). The definition of relevant technical terms such as land administration, land information systems, cadastre and cadastral systems are provided.

The UN Economic Commission for Europe (UN-ECE, 1996) Guidelines on Land Administration defines land administration as "the process of determining, recording and disseminating information about tenure, value and use of land when implementing land management policies. It is considered to include land registration, cadastral surveying and mapping, fiscal and multi-purpose cadastres and land information systems." (see Dale and McLaughlin, 1999; Steudler, 2004; Steudler *et al.*, 2004; Bandeira *et al.*, 2010).

Land Information Systems (LIS) are tools for legal, administrative and economic decision making and their operations include procedures and techniques for systematic acquisition and assembly of data; their processing, storage, maintenance, retrieval, analysis and dissemination (FIG-United Nations, 1983; Bedard, 1986; Dale and McLaughlin, 1989; Larsson, 1991). The usefulness of the system depends on how often it is updated, its accuracy, completeness,

accessibility, comprehensiveness, transparency and the extent to which the system has been designed for the benefit of users rather than for the producer of the information (see Dale and McLaughlin, 1989; Hammond, 2006; Hammond and Antwi, 2010). Quintero (2004) broadens the meaning and conceives LIS as the final product of a complete network of three general land related recording systems namely the cadastral survey, land registration system and all relevant public, private, and mixed land information sources. This conception of LIS appears to include unofficial procedures undertaken by individuals or small groups for their private transactions. This meaning falls in tandem with the huge information gap in Sub-Saharan Africa real estate markets the sources and cause which the study investigates.

Bandeira *et al.* (2010) contribute that normally a land information system consists of a registry (juridical cadastre) and a cadastre (fiscal cadastre). The registry serves as the database and compiles the information on legal aspects of properties such as legal ownership, spatial extent, charges and liens. The cadastre on the other hand is the geographical and alphanumerical database and it describes the features, use and spatial coordinates of land parcels. These databases, that is the registry and the cadastre, commonly referred to as cadastral systems, are supposed to provide reliable and current information on land ownership, use, value and spatial location for users.

3 Evolution of Land Recording Systems in Europe

Modern land information systems evolved in response to the need for land records which started in the form of cadastres in Western countries. An insight therefore into the origin of land recording systems in Europe may be helpful in gaining a better understanding of the evolution of land information systems in Sub-Saharan African countries compared to their colonial masters particularly Britain, France and Germany. Larsson (1991) offers two basic historical reasons for formal land records as the need for the state to know all land units for taxation or other fees and dues and secondly, the need for prospective buyers of land to get publicity for their acquisition of land. Dale (1997) espouses that the formal basis of dealing in land is either known as land registration or cadastral system and further explains that land registration owns its origin from land tenure whilst the origin of cadastre tend to lie in land and property taxation. Thus the trend of evolution of formal land recording systems in many Western European countries was for fiscal purposes and later for legal ownership in land (Dale, 1997; Ting *et al.*, 1999).

3.1 Historical Overview of British, French and German Cadastres and Land Information Systems

The great survey of Britain which was known as the Domesday Book and carried out in 1086 during the reign of William the Conqueror was a record of information collected for levying of taxes and has since then been one of the oldest public records in Britain (Plucknett, 2001; Larsson, 1991; Abdulai, *et al.*, 2007). According to Plucknett (2001, p. 12) "The land was described county by county, village by village, the owners and their subtenants were listed and their holdings valued, even the farm stock was recorded, with a view to settling clearly the rights of the crown and the taxable resources of the country". In many instances, the records were accompanied by a concise summary of the customs of a county or city so as to give an insight into the local law in force (Plucknett, 2001 and Larsson, 1991). This suggests

that the survey took cognisance of indigenous laws and practices an approach which the study questions whether it was replicated in Sub-Saharan Africa.

Plucknet (2001) posits further that from the Domesday Book, the idea of an officially compiled documents which was unquestionable appeared to have spread to the rolls of the Exchequer, and then to the rolls of the Courts of law. Consistent with other literature, Plucknett (2001) suggests that the evolution of land records may be regarded as financial in origin, and only later becoming judicial. The Domesday Book, however, as observed by Larsson (1991) was not based on survey maps unlike the French Cadastre which was also an attempt of a nationwide inventory of land ownership under Napoleon I of France and which became the cradle of European Cadastre.

Most European cadastres were modelled or drew inspiration from the French cadastre of Napoleon I which started in 1807. This was purely a fiscal cadastre which was compiled and recorded in large scale maps inventorying individual properties often as consequence of the financial policies of Napoleon I (Dale, 1997; Steudler, 2004). Larsson (1991, p. 25) opines that Napoleon I brought real breakthrough to cadastral records in that Napoleon visualised the linkage between the cadastre and the legal land registry and declared that:

A good cadastre will be the best compliment of my civil law code to achieve systematic order in the area of real estate property. The plans must be so developed and be made exact that they will permit at any time to define and record the boundaries of the land property limits and to prevent the confusion or law suits otherwise arising.

The French cadastre according to extant literature became a model for most European countries as a result of the domineering position of France at that time (Larsson, 1991; Dale, 1997; Cadastral Template, 2003). Most continental European countries however, established systematic land registration systems which were based on local conditions and varied in quality and extent from the French Cadastre. Germany was the first European country to take important steps to introduce registration of title based on cadastral survey which started in the mid Nineteenth Century and by 1900 it was extended to cover the whole of Germany (Larsson 1991; Enemark, 2005). The common feature about the origin of most European land recording systems was that they were originally for fiscal purposes and later became legal. They evolved and developed in accordance with local circumstances and needs unlike in SSA.

4 Evolution of Land Recording Systems in Sub-Saharan Africa

The evolution of formal land recording systems in Sub-Saharan Africa took a reverse turn from their origins in Europe. Land recordings the study finds started from legal to fiscal (levying taxes if any, valuation and planning) purposes in Sub-Saharan Africa unlike in the case of their European colonial masters. The study argues this may contribute as a source to the present information problem. For instance in South Africa, the first cadastral survey in 1658 was for transfer of a piece of land along the Lisbeek River which was released to a servant of the Dutch East India Company. Furthermore in Tanzania the cadastral system was introduced by the German colonial administration in 1893 and later developed by the British and subsequently used for alienation of land to European settlers. In Namibia, at the beginning of the First World War, central survey control networks were established for only European settlers (see Cadastral Template, 2003).

Furthermore, Bruce (1998) opines that the French Civil Code of 1830 was decreed and initially applied in Goree in Dakar, Saint Louis and Rufisque in Senegal which were towns with significant number of French citizens. Later the application of the Code was broadened to permit European settlers to legitimize their land and also paved the way for initiation of land reforms (Bruce, 1998). According to Larsson (1991) the French cadastral system naturally exerted great influence on its colonies in North and Western Africa. Similarly, the English land law influenced the attitudes towards practice regarding land ownership and leaseholds in many former colonies and dominions. Dale (1997), Bruce (1998) and Lavigne-Delville (2000) further concur that both French Civil Code and English Land Law left lasting marks on the legal and institutional culture of Sub-Saharan African countries.

4.1 Legacies of Colonial Land Administration in Sub-Saharan Africa

Glaeser and Sheifer (2002) from an empirical study note that most countries in the world inherited their legal systems from their colonial settlers and masters rather than developed them indigenously. In Sub-Saharan Africa, the colonial masters' subjection of the respective countries to nonlocal legal regimes that departed fundamentally from the dominant customary systems of land tenure and natural resource management left a legacy of dual tenure system namely statutory and customary tenures (Bruce, 1998; Lavigne-Delville, 2000; Le Meur, 2005). Bruce (1998) provides that the Portuguese, Belgians and the French articulated clearly in legislation that "occupied lands" were to be governed in accordance with local custom while all other lands were declared to be state property but in practice this was extended to all untitled lands by the administrators. The French colonial administration initiated land reforms which were based on the Napoleonic Doctrine of 1810 that supplanted traditional land tenure systems with European system, effectively commoditised land and designated the state as the sole custodian of all land in the French colonies in Africa. By 1830 it instituted legal control of land related transactions. In 1899 a land decree was enacted and this created two main categories of land namely public and communal lands (Bruce, 1998).

McAuslan (2003) also, observes that between the 1920s and 1950s there was deliberate suppression and imposition of administrative controls on the customary tenure operations in

SSA. McAuslan (2003) further posits that almost in all Anglophone African countries, the major approach to land tenure was a type of maintenance of dual system and the imposition of tight administrative control. Generally the colonial powers imposed their own laws as the basic laws on the colonies but allowed customary law to operate in certain areas in as much as it did not conflict with the colonial administration (Blocher, 2006).

Dale (1997) also reveals that in most parts of Africa, land registration was introduced to protect the rights of European settlers while the rights of indigenous people were protected by customary law. Bruce (1998) mentions that the French Land Decree of 1899 instituted land registers which became major colonial land reform throughout the whole of colonial Africa. One of the tenure systems was branded by the colonial French administration as 'native' or *indigenes* while the other system governed European residents and the French labelled Western acculturated members of the indigenous population (Njoh, 2004 and 2009). Despite resistance from the indigenous people to the introduction of the dual system of land delivery, the legacy of dual tenure systems from both French and British colonial rule continued into the post-independence era and persisted to the present day (Bruce, 1998; Njoh, 2004).

4.2 Land Administration in Post Independence Sub-Saharan Africa

Most Sub-Saharan African countries upon attainment of independence enacted land legislations and introduced policies that continued with the colonial legacy of land administration and management (Bruce, 1998; Lavigne-Delville, 2000). The legislations and policies were in the form of nationalization, vesting or some other form of these. Nigeria, for instance, revived the Protectorate legacy by proclaiming state ownership of all lands in the country (Bruce, 1998; Lavigne-Delville, 2000). According to Lavigne-Delville (2000) in Francophone countries land was often nationalised and this amounted to abolition of customary systems.

In the past two decades, however, most countries in Sub-Saharan Africa among other developing countries have introduced various land reforms and policies aimed at promoting economic growth, sustainable development and poverty reduction (Toulmin and Quan, 2000; Le Meur 2005) and also ensure that efficient land policies integrate modern statutory law with customary laws that govern many people's day-to-day lives. Bruce and Knox (2009), Sikor and Muller (2009) concur that the most common approach presently to land reform in Sub-Saharan Africa is land registration and land titling with focus on the customary tenure which were hitherto neglected. The study argues that these reforms and policies are without a thorough understanding of the sources and causes of the deficient information regime and underperformance of the land markets and are attempts to address the effects and not the main causes of the information problem. The study therefore discusses the sources and causes of information asymmetry in SSA real estate markets further as follows.

5 Sources of Land Information Asymmetry in Sub-Saharan Africa

The study reviews the relations between laws, registration systems and information asymmetry in SSA and argues that the extent and coverage of land registration systems depend on the kind of laws underpinning the registration systems. Three basic principles common to almost every land recording according to the relevant registration laws are;

Type of registration system;

Form of registration of land ownership; and

Approach to establishment of registration records (Cadastral Template, 2003).

5.1 Laws and Land Registration Systems

Deed and Title Registration Systems are the major types of land registration systems (Larsson, 1991; Palmer, 1998; Zevenbergen, 2002; Karikari *et al.*, 2003; Torhonen, 2004; Enemark, 2009; Enemark 2005a; Rajabifard *et al.*, 2007; Enemark *et al.*, 2009). There are, however, variant versions of both systems. For instance, the Torrens System of Australia which is a variant of the Title System (Enemark, 2005; Enemark *et al.*, 2005). The study finds that the type of registration system in a country may not necessarily offer any advantage over the other. Rather, it is a question of whether the system responds to the needs of local practices. If the type of registration system evolved in response to the needs of a country, from the culture of a country or it was designed within the social, legal history and economic context of a country (Enemark *et al.*, 2005; Dale, 2000) then probably no particular registration system can be considered to offer the best incentive in the absence of an empirical evidence (Deininger and Feder, 2009). The following examples buttress these assertions.

Examples of countries operating the Deeds System with 100% coverage are Belgium, France, the Netherlands and Hong Kong 99.9%. Examples of countries operating Title System with 100% registration are Germany, Denmark, Sweden, Switzerland, India, South Korea and Australia 99% (Cadastral Template, 2003). In contrast, there are countries that operate either one or both systems but have incomplete coverage. Examples of Deeds System with incomplete coverage are South Africa with 70% urban land registered, Namibia with 60%, and Mexico with 61% urban land registered. Examples of countries operating Title System and their coverage of registered urban lands are Malaysia 98%, Japan 72%, Tanzania 10% and Portugal with 17,000,000 parcels out of which 2,055,000 parcels are registered (Cadastral Template, 2003; Manthorpe, 2005).

Palmer (1998) earlier argues that both title and deeds registration systems acting as information systems reduce information asymmetry between parties in transaction. Palmer (1998) argues further that both systems "require similar quality of information in order to function well" consequently it is better to introduce administrative reforms that will improve quality of information rather than to attempt to transform from deeds to title system. Bittner and Frank (2002) argue that the deeds and title systems can be compared as their fundamental principles are similar and both systems are different solutions to a common problem of

determining the legal rights to a parcel of land. Zevenbergen (2002) after a comparative study of four countries including Ghana's land registration system observes that the flaws in the country's Deeds System in terms of technical, organizational and legal aspects should have been corrected rather than introduce Land Title Registration Law of 1986.

Baird and Jackson (1984) and Woodman (1986) concur that legal rules can affect the amount of information available and reduce the uncertainty necessary for contracting parties to bear. Palmer (1998) notes registration systems offer security if only it provides sufficient information and accurate information as well as protection of rights after registration. Palmer (1998) further mentions that in the advanced economies, land registration provide confidence that property rights will be easy to identify, verify, and protect against unlawful actions. Thus results of legal actions will be easy to predict. From a historical point of view, Deininger and Feder (2009) reveal that laws and customs have underpinned the social and economic benefits of land registration systems in developed societies.

5.2 Laws and Registration of Land Ownership

The second cadastral principle which influences information production is what the law prescribes for registration of land ownership. Land registration by law may take the form of either compulsory or optional, or both. Compulsory or optional land registration may depend on a country's social, economic, legal and political environment. In any case, the registration law must be grounded on the culture and customary practices of the country for it to gain acceptance by the citizens and enhance their knowledge of the registration laws (Hayek, 1983; de Soto, 2001). Furthermore, Mitchell *et al.*, (2008), Feder and Nishio (1999) point out that one of the factors that may lead to adverse social impact of land registration is the degree of understanding of the landholders of the processes and benefits of land registration.

5.3 Approach to Establishment of Cadastral Records

The third principle is the approach to land registration which hinges on whether specific laws pave the way for either systematic or sporadic approach to registration, or both. Registration is said to be sporadic when it is performed upon demand by the individual land holder (Feder and Nishio, 1999). Extant literature suggests sporadic approach does not provide an incentive for voluntary land registration and has been a failure in past registration practices in some developed economies. Larsson (1991) notes objections against the sporadic approach to registration and cautions that;

Sporadic approach to registration of titles takes a long time to complete the register;

Registration costs per unit of land are high; and suggests that

Systematic registration provides an overview of all the existing parcels and titles within an area.

Feder and Nishio (1999) concur that the potential for adverse social impacts is larger when registration and titling are sporadic and when the cost incurred by the individual for registration is also high (see Mitchell *et al.*, 2008, p.465). This principle it appears has been

largely ignored in Sub-Saharan African countries from the colonial era till present day. The practice mostly has been optional registration with sporadic approach as shown below.

Table 1. Summary of Cadastral Principles for Reviewed Sub-Saharan African Countries

(Sources: Adapted from Cadastral Template, 2003 and Zevenbergen, 2002)

COUNTRY	REGISTRATION SYSTEM	REGISTRATION BY LAW	APPROACH TO REGISTRATION
Ghana	Deeds	Optional	Sporadic
Namibia	Deeds	Optional	Sporadic
South Africa	Deeds	Optional	Sporadic
Tanzania	Title	Optional	Sporadic

Zevenbergen (2002) comments on the introduction of Ghana's land title registration which is systematic to have progress slowly since 1986 and further says “in practice most first registration has taken the form of more or less sporadic registration”. Further discussions on factors contributing to the causes of land information problems in Sub-Sahara Africa from legal deficiencies are provided.

6 Urban Land Delivery, Registration and Customary Laws Nexus

The study further traces the pattern of formal land recordings evolution in Sub-Saharan Africa to the colonial urbanisation process which was also influenced by laws as possible source of information asymmetry. Cohen (2003) suggests that to understand urban change in Africa requires a consideration of the role of colonialism in social, economic and political history of the various countries. Kironde (2001) notes that the socially segregated land use pattern of the colonial policy and attempts to exclude the Africans from urban land or limit their access to it could not work as the customary tenure ensured a parallel system of access to land. Njoh (2004) and Cohen (2003) extend the argument further that the colonial laws of registration which were skewed initially towards transfer of land to European settlers in Sub-Sahara African countries resulted in the creation of exclusive urban areas. These areas enjoyed a high level of urban infrastructure and services while the indigenous areas were marginally serviced. These developments, the study argues are naturally to be rivalled by the customary tenure in providing access to land for the natives outside the European zoned settlements.

A review of literature indicates the process of parallel land delivery through the customary system gained momentum after independence when many Sub-Saharan African cities began to grow rapidly due to high population growth and high spatial mobility. The high social mobility was influenced by the better access to health and educational facilities, urban bias in terms of trade between primary products and manufactured goods and the availability of newly established public sector jobs (de Soto, 2000; Cohen, 2003). The high social mobility triggered demand for land for urban development. Most of the land delivery was therefore through the informal land delivery mechanism as the registration process faced competition from the informal system and did not gain legitimacy of the indigenes (Kironde, 2000; Sivam, 2002; Deininger and Feder, 2009). Fourie (1998) indicates that informal land delivery destroys the currency of land records to the extent that the records will no longer mirror what is on the ground. According to Fourie (1998) this partly explains why land records in the registries of most countries are not current, contain insufficient information and are fraught with uncertainties.

According to UNCHS (1996) cited in Fourie (1998, p.55) a total of 20-80% of all land delivery in the developing world is carried outside the formal sector and the extent of the informal land delivery varies from country to country. de Vries (2004) posits that in Sub-Saharan Africa less than 1% of the total land area is administered under any kind of formal land registration system. Consistent with earlier estimates, Toulmin (2008) reveals that currently the information being produced and certified by land registries for land markets in West Africa region of Sub-Saharan Africa is from only about 2-3% of land ownership registered. In Burundi it is estimated that less than 1% of land is registered (Toulmin, 2008) and in Kenya, 6% of land available for registration had been registered as at 1999 (Miceli and Kieyah, 2003). The coverage of registration system in Madagascar according to Bruce and Knox (2009) is about one-fifth (6.67%) of the country, while in Ghana the national coverage is about 8% (Hammond, 2006; World Bank, 2005). In Cameroon it is estimated that out of the 1,600,000 parcels nationwide only 100,000 representing 6% are registered and less than 20% of the urban plots demarcated have been titled (McAuslan, 2003).

Critical examination of literature indicates informal land delivery system is often the consequence of legal deficiencies, high transaction costs and bureaucratic bottlenecks. Fourie (1998) argues assertively that in order for the informal land sector to become an asset, the land registration and land information systems of a country must be transformed. The study argues that laws of the respective countries must facilitate the transformation. Sanders (1987) suggests laws must depart fundamentally from the colonial and post-independence laws which fundamentally failed to transform let alone eradicate indigenous customary laws. Concerning which laws must be transformed McAuslan (1989) argue there are both good and bad laws with the latter serving as an obstacle to efficiency.

Determining which law is good or bad, the study argues that the basis must be laws which are dynamic as society changes. Seidman (1975) observes that society is in constant flux but little a government changes its laws. Thus society changes while laws remain static. Sanders (1987) opines both colonial and post independence laws are ineffective and not in

consonance with what Africans consider to be their laws. Hendrix (1996) argues that the imported laws and registration systems have failed mainly because;

The system is incapable of addressing fractures of estates;

Inadequate in addressing all forms of property ownership which should legally be eligible for registration.

Lavigne-Delville (2000) concur that legislation, legal principles and conception of law are still profoundly alien to the customary principles and land holding practices even though according to Toulmin and Quan (2000) efforts are being made to learn from the mistakes of the past. De Soto (2000) however, challenges the status quo that creating a system that integrates both formal and informal systems is not about drafting laws and regulations that look only good on paper. The challenge is rather about designing legal norms that are rooted in the people's beliefs, are legitimate and more likely to be obeyed and enforced. Stuart (1981) argues that the underperformance in some markets can be restored through legal intervention. The study argues the deficient information regime in Sub-Saharan African urban real estate markets has legal roots and it requires legal intervention based on local practices to address the problem and bridge the information gaps.

7 Conclusion and Remarks

The study concludes that the primary source of the huge information gap in Sub-Saharan urban real estate markets is the insufficient information at the land registries due to the predominance of the informal sector in land delivery process. Thus legal origins of land information systems and deficient land registration laws are identified as the primary causes of information gap in Sub-Saharan urban real estate markets. The study also notes that laws which are based on local practices provide incentives for registration rather than the type of land registration system.

The study therefore suggests that the information problem can be fundamentally addressed through legal reforms that reflect the cultural practices and the aspirations of the respective countries. This approach the study finds will ensure spontaneous legitimacy of the registration systems and also serve as an incentive in attracting transactions in the informal sector to the formal sector. Best practice from the study indicates that mapping the entire country should be the first step towards tackling the deficient information problem. Also, the Sporadic Approach to registration must gradually be phased out as piecemeal gathering of data does not provide comprehensive information in the land registry. The systematic approach to registration is suggested as a long term solution. It is envisaged a cost-benefit analysis of the current deficient information regime compared to an improved information system that produces quality information taking into account local practices and circumstances would inform government policy direction in bridging the information gap.

8 References

- Abdulai, R.T., Ndekugri, I.E., Olomolaiye, P.O. and Proverbs, D.G. (2007), 'Land registration and security of land tenure', *International Development Planning Review*, **29** (4), pp 475-502.
- Baird, D., Jackson. T. (1984), 'Information, Uncertainty, and the Transfer of Property', *Journal of Legal Studies*, **13**, pp 299-320.
- Bandeira, P., Sumpsi, J. M., & Falconi, C. (2010), 'Evaluating land administration systems: a comparative method with an application of Peru and Honduras', *Land Use Policy*, **27**, pp. 351-363.
- Barnes, G. (2003), 'Lessons learned: an evaluation of land administration initiatives in Latin America over the past two decades', *Land Use Policy*, **20**, pp 367-374.
- Beck, T., Demirguc-Kunt, A. And Levine, R. (2003), 'Law and finance: why does legal origin matter', *Journal of comparative Economics*, **31**, pp 653-675.
- Bedard, Y. (Winter, 1986), 'A Study Of The Nature Of Data Using A Communication-Based Concept Framework Of Land Information Systems', *The Canadian Surveyor*, **40**, pp 449-460.
- Bertocchi, G. And Canova, F. (2002), 'Did colonization matter for growth? An empirical exploration into the historical causes of africa's underdevelopment', *European Economic Review*, **46**, pp 1851-1871.
- Blocher, J. (2006), 'Building on Custom: Land Tenure Policy and Economic Development in Ghana' *Yale Human Rights and Development Law Journal*, **9**, pp 166 – 202.
- Bogaerts, T. (1997), 'A Comparative Overview of the Evolution of Land Information Systems In Central Europe', *Computers, Environment and Urban Systems*, **21** (2), pp 109-131.
- Bruce, J.W. (Dec.1998), 'Country Profile of Land Tenure', Africa Land Tenure Centre Research Paper No. 130.
- Bruce, J. W. and Knox, A (2009), 'Structures and Stratagems: Making Decentralization of Authority over Land in Africa Cost-Effective', *World Development*, **38** (8), pp 1360 -1369.
- Cadastral Template (2003), 'A Worldwide Comparison of Cadastral Systems. Established under UN mandate by Resolution 4 of the 16th UNRCC-AP in Okinawa Japan in July, 2003. www.cadastraltemplate.org, viewed:23/10/09
- Chatterton, W. A. (1980), 'Land Information Systems for the Twenty-First Century. Real Property', *Probate and Trust Journal*, **15**, pp 890-900.
- Creswell, J. W. (2009) '*Research Design. Qualitative, Quantitative, and Mixed Methods Approaches*', Third Edition. University of Nebraska-Lincoln. Sage Publications, USA.
- Dale, P. (1997), 'Land Tenure Issues in Economic Development', *Urban Studies*, **34** (10), pp. 1621-1633.

Dale, P. and McLaughlin, J. D. (1989), *Land Information Management. An Introduction with Special Reference to Cadastral Problems in Third World Countries*. Clarendon Press. Oxford.

Dale, P. and McLaughlin, J. (1999), *Land Administration*. Oxford University Press Inc. Oxford, New York.

de Soto, H. (2000), *The Mystery of Capital: Why Capitalism Triumphs in the West and Fails Everywhere Else*, Bantam Press, Black Swan Edition, London.

Deininger, K. (2003), *Land Policies for Growth and Poverty Reduction*. World Bank and Oxford University Press, Washington DC.

Deininger, k. And Feder, G (2009), 'Land Registration, Governance, and Development: Evidence and Implications for Policy', *The World Bank Research Observer*, **24** (2), pp 233-266.

Enemark, S. (2005), 'A Cadastral Tale', *Week on Geomatics 2005*, Bogota, Colombia, 8 – 13 August, 2005.

Enemark, S. (2009), 'Land Administration Systems', *Map World Forum*, Hyderabad, India, 10-13 February, 2009.

Enemark, S., Williamson, I. and Wallace, J. (2005), 'Building Modern Land Administration Systems in Developed Economies', *Spatial Science*, **50** (2), pp.51-68.

Farvacque, C. And McAuslan, P. (1992), 'Reforming Urban Land Policies and Institution in Developing Countries', *Urban Management Program Policy Paper* No. 5

Feder, G. and Feeny, D. (1991), 'Land Tenure and Property Rights: Theory and Implications for Development Policy', *World Bank Economic Review*, **15**, (1), pp 135-153

Feder, G. and Noronha, R. (1987), 'Land Rights Systems and Agricultural Development in Sub-Saharan Africa', *World Bank Research Observer*, **2** (2), pp143-169.

Fourie, C. (1998), 'The role of local land administrators: an African perspective', *Land Use Policy*, **25** (1), pp 55-66.

Glaeser, E. and Sheifer, A. (2002), 'Legal Origins', *Quarterly Journal of Economics*, **117** (4), pp 1193-1230.

Gough, K.V. and Yankson, P.W.K. (2000), 'Land markets in African cities: The case of pre-urban Accra, Ghana', *Urban Studies* **37** (13), pp 2485-2500.

Hammond, F. N. (2006), 'The Social Cost of Real Estate Market Information Gaps in Ghana', *Urban Policy Research*, **24**, (3) pp 391-408.

Hammond, F. N. and Antwi, Y. A. (2010), *Economic Analysis of Sub-Saharan Africa Real Estate Policies*, Palgrave Macmillan, Basingstoke and New York.

Hayek, F. A. (1983), *Knowledge, Evolution and Society*, The Adams Smith Institute, UK.

Jacoby H. G. And Minten, B. (2009), 'Is land Titling in Sub-Saharan Africa Cost-Effective? Evidence from Madagascar', *The World Bank Economic Review*, **21** (3), pp 461-485.

Karikari, I., Stillwell, J. and Carver, S. (2003), 'Land Administration and GIS: The Case of Ghana', *Progress in Development Studies* **3** (3), pp 223-242

Kironde, J. M. L. (2000), 'Understanding land markets in African urban areas: the case of Dar el Salaam, Tanzania', *Habitat International* **24** (2), pp 151-165.

Lavigue-Delville, P. (2000), 'Harmonizing formal and customary land rights in French-speaking West Africa', In: Toulmin C. and Quan J. (eds), *Evolving land rights, policy and tenure in Africa*, DFID/IIED/NRI, London, pp. 97-121

Larsson, G. (1991), *Land Registration and Cadastral Systems: Tools for land information management*, Longman Scientific and Technical. Longman Group UK Ltd, England.

Le Meur, P. (2005), 'Land Tenure Information and Land Policy in Sub-Saharan Africa. Synthesis Report (Benin, South Africa), FAO normative work on land tenure data and policy-making, September 22-23, 2005, FAO, Rome.

Manthorpe, J. (2005), 'Inventory of Land Administration Systems in Europe and North America'. HM Land Registry, London. <http://www.unece.org/env/wpla/welcome.html> viewed 12-11-09

McAuslan, P. (2003), *Bringing the law back: essays in land, law and development*, Ashgate Publishing Ltd. Aldershot, England.

Mitchell, D., Clarke, M. and Baxter, J. (2008), 'Evaluating land administration projects in developing countries', *Land Use Policy* **5**, pp 464-473.

Njoh, A. J. (Oct. 2004), 'The experience and legacy of French colonial urban planning in sub-Saharan Africa', *Planning Perspective*, **19**, pp 435-454.

Njoh, A. J. (July, 2009), 'Urban planning as a tool of power and social control in colonial Africa'. *Planning Perspective*, **24** (3), pp 301-317.

Olsson, O. (2009), 'On the democratic legacy of colonialism', *Journal of Comparative Economics* **37**, pp 534-551

Plucknett, T.F.T. (2007), *A concise history of the Common Law*, Fifth Edition, the Lawbook Exchange Ltd. Union, New Jersey.

Quintero, J. R., (2004), Land Information Systems. International Encyclopedia of the Social & Behavioral Sciences, pp. 1-6.

Rajabifard, A., Williamson, I., Steudler, D., Binns, A. and King, M. (2007), 'Assessing the worldwide comparison of cadastral systems', *Land Use Policy*, **24**, pp 275-288.

Sanders, A. J. G. M. (1987), 'Towards A People's Philosophy Of Law', *Journal of African Law*, **31**, pp 37-43.

Steudler, D., Williamson, I. P., Rajabifard, A. and Enemark, S. (2004), 'The Cadastral Template Project', *FIG Working Week 2004*, Athens , Greece, May 22-27, 2004.

Ting, I ., Williamson, I.P., Grant, D. and Parker, J.R. (1999), 'Understanding the Evolution of Land Administration Systems in Some Common Law Countries'. Survey Review.

Toulmin, C. (2008), 'Securing land and property rights in sub-Saharan Africa: The role of local institutions', *Land Use Policy* **26**, pp 10-19.

Toulmin, C. And Quan, J. (2000), 'Evolving Land Rights, Policy and Tenure in Sub-Saharan Africa', In: Toulmin C. and Quan J. (eds), *Evolving land rights, policy and tenure in Africa*, DFID/IIED/NRI, London, pp. 1-30.

Torhonen, M. (2004), 'Sustainable land tenure and land registration in developing countries, including a historical comparison with an industrialised country', *Computers, Environment and Urban Systems*, **28** (5), pp 545-586.

Woodman, G. R. (1987), 'Land Registration Without Prejudice: The Ghana Land Title Registration Law, 1986', *Journal of African Law*, 31, pp 119-135.

Zevenbergen, J. (2002), *Systems of Land Registration-Aspects and Effects*, PhD Thesis Published by Nertherlands Geodetic Commission, delft, The Nertherlands

Conceptualising the Singapore Real Estate Investment Trust

Seah Xiande Aaron*

*LLB (NUS)

Email: aaronseah@yahoo.com

ABSTRACT

The Singapore Real Estate Investment Trust (“S-REIT”) is an investment vehicle structured as a trust to hold, trade and manage real property or property mortgages for the benefit of investors called unitholders. A unique feature of the S-REIT is that unitholders are expressly excluded as the equitable owners of the underlying S-REIT assets. Two prominent commentators have taken the view that this renders the S-REIT an invalid trust because the exclusion of equitable ownership offends the beneficiary principle, which is a fundamental trust law rule. However, they suggested that the S-REIT is possibly valid by being construed as a Re Denley non-charitable purpose trust. This paper argues that the Re Denley trust is not a non-charitable purpose trust. Instead, the Re Denley trust is a trust for human beneficiaries. Accordingly, it will be shown that the S-REIT is a Re Denley trust for human beneficiaries, thereby not offending the beneficiary principle. In support of this view, the jurisprudential bases of the beneficiary principle and alleged need to vest equitable interest in the beneficiaries shall be investigated. The further argument advanced in this paper is that if the S-REIT is considered invalid because equitable ownership does not vest in the unitholders, the beneficiary principle should be re-conceptualised as being premised on the enforcer principle. This means that the S-REIT should be valid if there are enforcers who can hold the trustees accountable to the obligations undertaken.

I. INTRODUCTION

In a world where financial innovations abound, it is unsurprising to find financiers and lawyers collaborating to come up with more creative ways to raise funds and improve investment returns. One such financial ingenuity that proved to be successful is that of securitisation (Burns 2006; Finch 2008, pp.759–761; but see Hovanesian and Conlin (2002), who reported that securitisation has somehow lost its grandeur). Securitisation is a process where viable assets with the potential to generate future cash flows are hived off to an investment vehicle (Borrows and Smith 2002, pp.1–4). The investment vehicle is in turn managed for the benefit of a multitude of investors, who will subsequently be paid returns from the future cash flows generated. An example of this investment vehicle is the Real Estate Investment Trust (“REIT”), which essentially owns and manages commercial property and other real-estate related assets such as mortgages. As its name suggests, the REIT is structured as a trust which interest is divided into units. Investors subscribe to the units to be entitled to the profits generated by the assets held under the trust. In addition, investors are also entitled to the residual REIT assets on dissolution of the REIT.

The success of the REIT has been nothing short of phenomenal. Ever since the REIT began its existence in the United States in 1960 (Campbell, Samuels and Yuasa 2006, p.219; Chan, Erickson and Wang 2003, p.3), the use of the REIT has proliferated in many jurisdictions. This was due in no small part to the publicity generated by Kimco Realty REIT’s Initial Public Offering in the US in 1991 (Schaefer 2005, p.5) and the fact that most other assets in the United States have been securitised (Hovanesian and Conlin 2002). Furthermore, owing to the inherent flexibility of the REIT as a property investment vehicle to

create a robust property market, jurisdictions such as the United Kingdom (HM Treasury 2005, pp.5–6) have readily embraced the use of the REIT. As of December 2009, the global market capitalisation of REITs worldwide has been valued at a whopping \$568 billion and is set to grow in 2010 (Zieminski 2010). In Singapore, the total market capitalisation of the REITs listed on the Singapore Exchange Ltd. (“SGX”) was approximately S\$26 billion as of 30 October 2009 (Lee and Foo 2010, p.40).

The REIT was introduced in Singapore in 1999 when the Monetary Authority of Singapore issued the first REIT regulations (Tan 2005, p.109). However, the taxation levied on the Singapore REIT (“S-REIT”) discouraged the setting up of the REIT in Singapore as an investment vehicle. It was only in 2001 when the Inland Revenue Authority of Singapore granted a case-by-case ruling to exempt S-REITs from taxation did the REIT become attractive as an alternative investment vehicle. SingMall Property Trust was granted such tax exemption but its listing was withdrawn because of lack of investor interest (there was a 20 percent under-subscription of the 530 million units offered; Tan 2005, p.111). In 2002, the first S-REIT, CapitalMall Trust, was listed on the SGX. As of 8 April 2010, a total of 20 S-REITs is listed on the SGX (SGX Real Estate Investment Trusts 2010).

A unique feature of the S-REIT is that most S-REIT trust deeds expressly provide that the equitable interest in the underlying S-REIT assets does not vest in the unitholders (Ascendas REIT Prospectus, p.146; Suntec REIT Prospectus, p.151; K-REIT Introductory Document, p.120; Saizen REIT Trust Deed; Koh 2006, p.182). The equitable interest does not vest for taxation reasons. When equitable interest is excluded, stamp duty need not be

paid when unitholders transfer their units to another person.¹ It is noted that other common law jurisdictions such as Australia (Deveson *et al* 2006), Canada (Armstrong and Ulmer 2006, p.41), the United States (Campbell, Samuels and Yuasa 2006, p. 222), Hong Kong (Leung and Smith 2006) and Malaysia (Gooi and Teong 2006) do not have a practice of excluding beneficial interest. In Australia (Deveson *et al* 2006, p. 23), Hong Kong² and Malaysia,³ the most likely reason for doing so is that stamp duty is required to be paid whenever units under the REIT are traded, so there is no need to exclude beneficial interest to avoid payment of stamp duty.

However, this special trait of excluding equitable interest has proven to be the S-REIT's Achilles' heel, for two prominent commentators, Professor Hans Tjio and Mrs. Lee Suet Fern, have argued that the S-REIT offends one of the cardinal requirements of trust law, namely that of the beneficiary principle. According to the beneficiary principle, a trust must have ascertainable beneficiaries for it to be valid.⁴ While the S-REIT technically has beneficiaries in the form of unitholders (see also Sin (1997, p.69) who has taken the view that the S-REIT's closest cousin, the unit trust, has "certainty of objects because the objects are the same persons who have transferred the subject matter [i.e. subscription monies] to the trustee (or transferees of such persons)"), thereby satisfying the beneficiary principle, Tjio and Lee were of the view that the beneficiary principle is complied with only when the unitholders own the underlying S-REIT assets (Tjio and Lee 2006, p.50). To support their view, the learned authors relied on Michael Hwang JC's (as he then was) remark in *Lee Chuen Li v*

¹ See Stamp Duty Act (Cap 312, 2000 Rev Ed Sing) s 4 read with First Schedule Article No. 3(a).

² The Stamp Duty Ordinance (Cap 117) s 19 (Hong Kong);

³ The Stamp Act 1949 (Malaysia) read with First Schedule Item 31.

⁴ *Leahy v AG for New South Wales* [1959] AC 457 (PC) 478–479.

*Singapore Island Country Club*⁵ that “[l]ike nature, the law abhors a vacuum of ownership”. Since Hwang JC’s statement was made in the context of ownership in an unincorporated association,⁶ where assets are usually held on trust for the members, it appears that Tjio and Lee were suggesting that there has always to be two owners over S-REIT assets, namely the legal and equitable owners. In this case, given that the trustee has the legal interest in the trust property, the unitholders should correspondingly have an equitable interest in the trust property. Accordingly, if equitable interest does not vest in the unitholders, the beneficiary principle is offended. As a result, the learned authors have taken the position that the S-REIT is a trust for the purpose of investment in commercial real property which is *prima facie* invalid for being a non-charitable purpose trust.

Although Tjio and Lee considered the S-REIT to be a non-charitable purpose trust, they argued that the S-REIT should be upheld on the basis that it is a *Re Denley*⁷ (“*Re Denley*”) trust (Tjio and Lee 2006, pp.52–53; see also Tjio 1993, p.114). However, the learned authors rationalised the *Re Denley* trust to be an exceptional non-charitable purpose trust because they regarded the employees in *Re Denley* not as “traditional beneficiaries” but as “indirect beneficiaries with a negative *locus standi* to see that the trust was properly administered” (Tjio and Lee 2006, p.53). While Tjio and Lee are correct in conceptualising the S-REIT as a *Re Denley* trust, this author respectfully disagrees that *Re Denley* is an exceptional non-charitable purpose trust. Instead, it is suggested that the *Re Denley* trust should be conceptualised as a trust for human beneficiaries.

⁵ [1992] 2 SLR(R) 266, para 48.

⁶ *Ibid*, where His Honour further described the club members as “co-owners”. Since an unincorporated association has no separate legal personality, any property has to be held on trust for the benefit of the club members.

⁷ [1969] 1 Ch 373.

The first argument advanced in this paper is that it is erroneous to suggest that equitable interest has to vest in the beneficiaries for the beneficiary principle to be satisfied. An investigation into the underlying principle of the beneficiary principle in Part III shows that the beneficiary principle was essentially formulated to aid the court in the enforcement of trusts. Further, an examination of the basis for the vesting of equitable interest in the beneficiaries demonstrates that this is done in order to give the beneficiaries *locus standi* to enforce the trust. However, as case law illustrates, the enforcement of a trust is possible even when equitable interest does not vest in the beneficiaries. It appears that so long as there are identifiable beneficiaries, a trust which does not vest equitable interest in the beneficiaries will nevertheless satisfy the beneficiary principle. Since the employees in *Re Denley* are the identifiable beneficiaries, the *Re Denley* trust is actually consistent with the beneficiary principle. Accordingly, the *Re Denley* trust is a trust for human beneficiaries.

The second argument put forth in this paper is that if the S-REIT is not considered to be a *Re Denley* trust but an invalid non-charitable purpose trust, this author suggests that the beneficiary principle should be re-conceptualised as being premised on the enforcer principle to validate the S-REIT. This means that if the unitholders are able to hold the trustees accountable to their management of the S-REIT assets, the S-REIT should be a valid non-charitable purpose trust (see also Tjio and Lee (2006), p.55 who appeared to agree with this view but at the same time expressed apprehension about adopting a pure enforcer principle). This author will, in Part IV, show that this proposition is supported by the various authorities and presents a better view of the relationship between a trustee and beneficiary.

Before the discussion proper, this author shall describe the detailed structure of a typical S-REIT in Part II to facilitate the understanding of the arguments in Parts III and IV.

II. STRUCTURE OF A TYPICAL S-REIT

An S-REIT is a trust settled in Singapore for the commercial purpose of holding, trading and managing real property or property mortgages for the benefit of beneficiaries who are called unitholders. The interest in the S-REIT is divided into equal and undivided transferrable units. Individuals will then be invited either through prospectuses or advertisements to subscribe to the units.

If the units are *not* tradable in a stock exchange, the S-REIT is called a private S-REIT. Private S-REITs are usually open-ended investments in that trustees may have an obligation to sell or redeem units. Should private S-REITs have a value of more than S\$5 million⁸ or more than 50 investors,⁹ it will be subject to the whole of the Securities and Futures Act¹⁰ and Code on Collective Investment Scheme (MAS 2002; see further Tjio 2004, paras.3.16–3.21). If the units are tradable in a stock exchange, the S-REIT is called a public S-REIT. Public S-REITs are closed-ended investments because the trustees are not required to sell or redeem any unit. Any “redemption” is done by the unitholder selling his or her units on the stock exchange. Since public S-REITs are traded on the Singapore Stock Exchange Ltd. (“SGX”), the listing requirements (see SGX Listing Manual) set forth by the SGX, in addition to those prescribed by the Securities and Futures Act¹¹ and Code on Collective Investment Scheme (MAS 2002; see further Tjio 2004, paras.3.16–3.21), have also to be satisfied.

⁸ Securities and Futures Act (Cap 289, 2006 Rev Ed Sing) ss 272A and 302B.

⁹ *Ibid* ss 272B and 302C.

¹⁰ *Ibid*.

¹¹ (n 8).

The figure below shows a typical S-REIT.

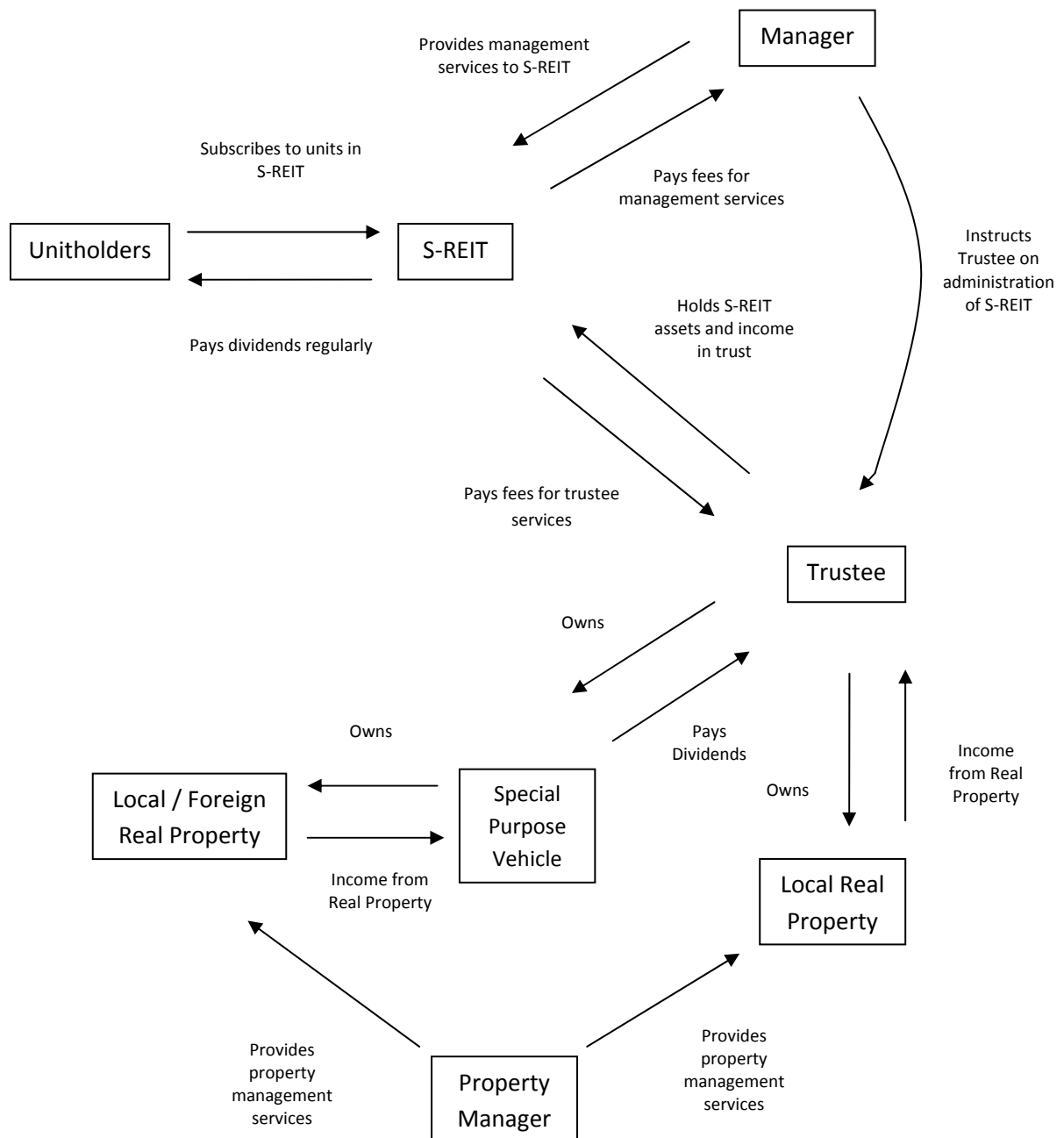


Figure 1. Graphical illustration of a typical S-REIT

The S-REIT is formed by a declaration of trust over S-REIT assets for the benefit of the unitholders (Ascendas REIT Prospectus, p.146; Suntec REIT Prospectus, p.151; K-REIT

Introductory Document, p.120; S aizen REIT Trust Deed). These assets include the subscription monies put forth by the unitholders. The declaration of trust is evidenced by a trust deed, which contains other provisions governing the rights and obligations of the unitholders, trustee and managers.

(1) Rights and Obligations of the Manager

The manager is responsible for providing management services to the S-REIT. It has to devise investment strategies and directions to maximise returns from the S-REIT assets. After the investment strategies and directions have been formulated, the manager will instruct the trustee to act according to the strategies and directions (see also Sin (1997, p.167) who also argued that the trustee and manager owe fiduciary duties to each other). The manager is usually remunerated in respect of the management services rendered.

(2) Rights and Obligations of the Trustee

The trustee, which is usually a Singapore corporate entity, is responsible for the proper administration of the S-REIT. The trustee is obliged to use the money pooled by the unitholders to purchase real estate investments as instructed by the manager or as set out in the trust deed or legislation¹² (see also MAS 2002) or both. If the trustee is investing in local real property as joint owner, the trustee has to do so through a special purpose vehicle (i.e., the trustee has to own a special purpose vehicle which in turn owns the piece of real property) (MAS 2002, Appendix 2 (Property Funds), para.6.4). For taxation reasons, the trustee is also advised to use a special purpose vehicle to invest in overseas properties (Koh 2006, p.181). The trustee is also obliged to receive income from the real estate investments and make

¹² Securities and Futures Act (n 8); Trustees Act (Cap 337, 2005 Rev Ed Sing).

dividend payments accordingly. Further, the trustee is usually entitled to receive fees in respect of the trustee services rendered.

(3) *Rights and Obligations of the Property Manager*

A property manager is hired to provide property management services. These services generally relate to the upkeep of the real properties held under the S-REIT. The property manager will be remunerated accordingly for the services provided.

(4) *Rights and Obligations of Unitholders*

A trust deed typically provides that the trustee holds the S-REIT assets on trust for the benefit of unitholders. However, the unitholders have no “equitable interest” in the underlying S-REIT assets (Koh 2006, p.182; Ascendas REIT Prospectus, p.146; Suntec REIT Prospectus, p.151; K-REIT Introductory Document, p.120; Saizen REIT Trust Deed). Furthermore, unitholders have no right to demand that the trustee transfer the ownership in the S-REIT assets in their joint names as typically permitted under the rule in *Saunders v Vautier*¹³ (see also Matthews 2006). Unitholders are also not entitled to interfere with the management of the S-REIT. However, unitholders have the right to ensure the due administration of the trust and receive income distributions, audited accounts, annual reports and a share of the residual S-REIT assets after the S-REIT has been wound up (K-REIT Introductory Document, p.IV-26; Koh 2006, p.182). Should the trustee breach the terms of the trust deed, unitholders are entitled only to damages (as opposed to specific performance or injunctive relief) (Saizen REIT Trust Deed). Another important feature of the S-REIT is that unitholders are liable towards the S-REIT only to the extent of the amount paid or payable for a unit of the S-REIT (Ascendas REIT Prospectus, p.151; Suntec REIT

¹³ (1841) 4 Beav. 115.

Prospectus, p.II-25; K-REIT Introductory Document, p.IV-26; Saizen REIT Trust Deed; Lee and Foo 2010, pp. 61–64), rendering the unit holders' position being similar to that of shareholders in a limited liability company (Lee and Foo 2010, pp. 61–64).

III. THE BENEFICIARY PRINCIPLE AND THE S-REIT

(1) *Does Equitable Interest need to vest in the Beneficiaries to comply with the Beneficiary Principle?*

Tjio and Lee argued that the beneficiary principle is offended when the beneficiaries do not collectively own the underlying trust property (Tjio and Lee 2006, p.50). This means that there has to be legal and equitable owners respectively over trust property, such a view being implicit from the learned authors' citation of Hwang JC's view in *Lee Chuen Li*¹⁴ that "[l]ike nature, the law abhors a vacuum of ownership". *Lee Chuen Li* was a case involving the right to claim relief for forfeiture in a club membership. The claimants in that case argued that such a right subsisted in their favour because they had a proprietary interest in the assets of the club. Hwang JC appeared to agree with the claimants when he remarked that the club members held the club assets as "co-owners".¹⁵ Since the club is an unincorporated association with no separate legal entity, any club asset has to be held under a trust.¹⁶ Accordingly, by being "co-owners", the club members must have been the beneficiaries under the trust over the club assets.

¹⁴ (n 5).

¹⁵ *Ibid.*

¹⁶ *Chen Cheng v Central Christian Church* [1995] 3 SLR(R) 806, para 31.

The view of Tjio and Lee also appears to be supported by the rule in *Saunders v Vautier*,¹⁷ where the beneficiaries who are of the age of majority and have an indefeasible interest in the trust property may ask for a transfer of the legal title to the trust property from the trustee to themselves. Hence, *Saunders v Vautier*¹⁸ demonstrates that beneficiaries technically own the underlying trust property, for the trustee is obliged to effect the transfer when requested by these beneficiaries (see also Matthews 2006; Matthews 1996, p.13–15).

However, this author respectfully disagrees with Tjio and Lee, for an investigation into the basis of the beneficiary principle reveals that the beneficiary principle was developed to allow a court to address a breach of trust through the help of the beneficiary. Further, a scrutiny of the basis of vesting an equitable interest shows that equitable interest is vested to give *locus standi* to the beneficiaries to enforce the trust. However, the vesting of an equitable interest is only but one way of giving *locus standi* to the beneficiaries to enforce the trust. There are many instances where the beneficiaries are entitled to enforce the trust even in the absence of a vested equitable interest. These instances can be broadly grouped into the discretionary trust and *Re Denley* trust categories. Thus, as these two categories of trust show, if a court is able to identify beneficiaries to enforce the trust, the trust will be upheld. There is no need for the beneficiaries to “own” the trust property.

¹⁷ (n 13).

¹⁸ *Ibid.*

(a) Basis of Beneficiary Principle

The beneficiary principle was formally introduced into trust law by an *obiter dictum* of Lord Parker in *Bowman v Secular Society Ltd*¹⁹ (but see Baxendale-Walker (1999, pp.145–147) who argued that the beneficiary principle was never binding but persuasive authority) where His Lordship remarked: “[a] trust to be valid must be *for the benefit of individuals ...*”. (emphasis mine)

The beneficiary principle was reiterated in *Re Wood*,²⁰ where Harman J held the trust therein to be invalid for want of beneficiaries. A few years later, in *Re Astor’s Settlement Trusts*,²¹ Roxburgh J considered the authorities in detail including *Re Wood*²² and Lord Parker’s *obiter dictum* in *Bowman*²³ and held that a trust to maintain “good relations between nations” and preserve “the independence of newspapers” was invalid for offending the beneficiary principle. The beneficiary principle has since been applied to invalidate trusts such as one for the research and development of a “British alphabet”²⁴ and one for providing “some useful memorial to myself”.²⁵

What then is the rationale for having the beneficiary principle? The rationale appears to be the courts’ fear of trustees behaving as though they are the absolute owners of trust property. This concern was expressed by Sir Grant MR in *Morice v Bishop of Durham*:²⁶

¹⁹ [1917] AC 406, 441.

²⁰ [1949] Ch 498.

²¹ [1952] 1 All ER 1067.

²² (n 20).

²³ (n 19).

²⁴ *Re Shaw* [1957] 1 WLR 729.

²⁵ *Re Endacott* [1960] Ch 232.

²⁶ (1805) 10 Ves Jr 522.

“There can be no trust over the exercise of which this court will not assume a control, for an uncontrollable power of disposition would be ownership and not trust. ... There must be somebody in whose favour the court can decree performance.”²⁷

The aversion can be further seen through Roxburgh J’s description in *Re Astor’s Settlement Trusts*²⁸ of the distribution of large accumulated funds without proper control over the trustees as being against public policy (but see Hayton, Matthews and Mitchell 2007, para.8.144, where a power could be construed instead, the basis being “*verba ita sunt intelligenda ut res magis valeat quam pereat*” (i.e., an interpretation which gives effect is preferred to one which makes void). However, this appears to be possible only where the intention is to create a charitable trust.²⁹)

Such fear is real because ownership at common law allows the owner to exclude others from enjoyment of the owned asset (Bridge 2002, p.12; McFarlane 2008, pp.132–133). Thus, it would be highly inequitable for trustees to behave as though the trust property belonged to them when they are in fact subjected to a trust.

In order to ensure that trustees do not behave as absolute owners, a court will require the assistance of a person to petition the court to enforce the trust. A trust for purposes is unlikely to assist the court in this regard because there is technically no person who will ensure that the trustee perform his or her obligations. A court will therefore find that no trust has been created in such circumstances (i.e., the trust is invalid). Professor Paul Matthews (1996, p.1) brought this point out forcefully when he satirically wrote that “animals and trees could not

²⁷ See also *Re Davidson* [1908-1910] All ER Rep 140, 144 (Sir Cozens-Hardy MR).

²⁸ (n 21).

²⁹ *Inland Revenue Commissioners v McMullen* [1981] AC 1 (HL) (Lords Hailsham LC, Diplock, Salmon and Keith).

[complain to the Chancellor who is the keeper of the King's conscience]" if a trust settled for their upkeep (which from a legal perspective is a trust for a purpose since animals and trees have no legal personality) was being misapplied (Matthews 1996, p.1).³⁰ By parity of reason, the trust in *Re Astor's Settlement Trusts*³¹ has to be invalid because the "nation" or the "newspaper" could not possibly petition a court if the trustee misapplied the trust property therein. Similarly, the trusts in *Re Shaw* and *Re Endacott* are also invalid because it would be impossible for the "British Alphabet" and deceased settlor in each respective case to appear in court to accuse the trustees of breach of their fiduciary duties! Therefore, in the light of the development of the beneficiary principle, it has become customary for trust law to regard a pure purpose trust as being void (Hayton, Matthews and Mitchell 2007, para.8.144; Mowbray *et al* 2008, para.4-38).

(b) Basis of Vesting Equitable Interest in the Beneficiaries

The recognition of a beneficiary as having an equitable interest in trust property is closely linked to the enforcement requirement (see also Thomas and Hudson 2010, para.6.13). As noted by Roxburgh J in *Re Astor's Settlement Trusts*,³² "a trustee would not be expected to be subject to an equitable obligation unless there was somebody who could enforce a correlative equitable right."

If a trustee is not the absolute owner of the trust property, there has to be a regime to ensure that the trustee does not behave otherwise. Since the common law regards the trustee

³⁰ See also *Re Davidson* (n 27) 144 (Sir Cozens-Hardy MR); *Re Alsagoff Trusts* [1956] 1 MLJ 244, 245 (Murray-Aynsley CJ); *Re Shaw* (n 24) 745 (Harman J.); *Leahy* (n 4) 479.

³¹ (n 21).

³² *Ibid.*

as the owner of the trust property, thereby allowing the trustee to deal freely with the trust property, equity intervenes by recognising the beneficiary as being equitably entitled to the trust property.³³ In order to give effect to such recognition, equity creates an equitable interest over the trust property and vests it in the beneficiary (Gray 1994, p.163).³⁴ As a result, if the trustees breach their fiduciary duties, by virtue of the equitable interest, the beneficiary has *locus standi* to appear before a court to enforce the trust.

Another way of analysing the need for equity to vest an equitable interest is the fact that, unlike the civil law, gratuitous undertakings are generally not enforced at common law (Chitty 2008, para.3-002).³⁵ This means that when a trustee agrees to hold trust property for the benefit of another, the common law will not enforce this obligation unless the settlor or beneficiary has provided some consideration (Chitty 2008, paras.3-001–3-003).³⁶ The need for consideration is a real stumbling block if a settlor declares himself a trustee or if a settlor contracts with a trustee to hold trust property gratuitously for the benefit of another. The common law will ignore the claims of both the settlor and beneficiary. However, even if consideration is provided, the beneficiary has to surmount the privity of contract doctrine³⁷ to have *locus standi* to sue the trustee if consideration had been provided by the settlor.³⁸ This conundrum leads to an unsatisfactory result because the trustee in these specific situations need not account to anyone at common law. Since “equity will not suffer a wrong without a

³³ *Baker v Archer-Shee* [1927] AC 844 (HL); see also *Saunders* (n 11), where a beneficiary who is of the age of majority and who has an absolute and indefeasible equitable interest in the trust property can call for the transfer of the legal title to the trust property; *Foskett v McKeown* [2001] 1 AC 102, where the equitable interest entitled the beneficiaries to claim the whole of a windfall benefit arising from the trust property.

³⁴ See *Commissioner of Stamp Duties (Qld) v Livingston* [1965] AC 694, 712, where Lord Radcliffe remarked that equity “calls into existence and protects equitable rights and interest in property only where their recognition has been found to be required in order to give effect to its doctrines”.

³⁵ This author notes that a possible exception is a contract executed by way of deed.

³⁶ *Re Hudson* (1885) 65 LJ Ch 811; *Re Cory* (1912) 29 TLR 18; *Williams v Roffey Bros & Nicholls (Contractors) Ltd* [1991] 1 QB 1, 19.

³⁷ Contrast the current position under Contracts (Rights of Third Parties) Act (Cap 53B, 2002 Rev Ed Sing); Contracts (Rights of Third Parties) Act 1999 (c 31) (UK).

³⁸ The settlor also faces the same issue if the beneficiary is the one who provided the consideration.

remedy” (Meagher, Heydon and Leeming 2002, paras.3-010–3-030), it is imperative that equity intervene and give *locus standi* to the beneficiaries by vesting an equitable interest in them.

On a more general note, since a beneficiary’s equitable interest in the trust property is binding in equity against every person in the world except the *bona fide* purchaser for value without notice³⁹ and the beneficiary’s successor in title,⁴⁰ the beneficiary is considered to have a proprietary interest in the trust property (Hayton, Matthews and Mitchell 2007, paras.2.12 and 2.15; Mowbray *et al* 2008, para.1-06). The proprietary nature of the beneficiary’s equitable interest can also be seen from the fact that an assignment of an equitable interest will protect the assignee from the insolvency of the assignor-beneficiary and that a trust ceases to exist when the trust property is lost without fault of the trustees (Penner 2008, pp.27–29).⁴¹ This is in contrast to a purely personal right which in the former case disentitles the assignee from claiming an interest in the trust but must prove in the assignor-beneficiary’s insolvency and which in the latter case entitles a beneficiary to sue the trustee despite the trust property being lost. Furthermore, where the beneficiaries are *sui juris* and have an absolute and indefeasible interest in the trust, they are entitled to terminate the trust and demand the transfer of the trust property to themselves.⁴²

It is therefore apposite for Lord Browne-Wilkinson in *Westdeutsche Landesbank Girozentrale v Islington London Borough Council*⁴³ to acknowledge that: “[o]nce a trust is

³⁹ *Pilcher v Rawlins* (1872) 7 Ch App 259, 268–269 (Lord James); *Westdeutsche Landesbank Girozentrale v Islington London Borough Council* [1996] AC 669, 705 (Lord Brown-Wilkinson).

⁴⁰ *Wilkes v Spooner* [1911] 2 KB 473.

⁴¹ *Westdeutsche* (n 39) (Lord Brown-Wilkinson).

⁴² *Saunders* (n 13).

⁴³ (n 39). See also *Abdul Hameed Sitti Kadija v De Saram* [1946] AC 208 (PC) 217; *Gosling v Gosling* (1859) Johns 265; *Re Nelson* [1928] Ch 920.

established, as from the date of its establishment *the beneficiary has, in equity, a proprietary interest in the trust property*” (emphasis mine). Arguably, the beneficiaries are v indicating their “proprietary interest” by enforcing the trust.⁴⁴

Prima facie, from the above discussion, it appears that if the equitable interest in trust property is not vested in the beneficiaries, the beneficiaries will have no right to enforce a trust (Matthews 2002, p.223). This would offend the beneficiary principle whose object, as discussed above, is to ensure that a court can effectively enforce a trust.

(c) Categories of Non-charitable Trusts which are valid even when Equitable Interest does not vest in Beneficiaries

However, it is not always true that equitable interest in a trust property has to vest in the beneficiaries so as to give rise to a right to enforce the trust or a proprietary interest in the trust (Thomas and Hudson 2010, paras.6.13–6.20 and 6.40–6.48). Lord Radcliffe implicitly acknowledged that equitable interest need not always vest in the beneficiaries in *Commissioner of Stamp Duties (Qld) v Livingston*⁴⁵ when he said:

“This dilemma [that there must be a separation of legal and equitable estates] is founded on a fallacy, for it assumes mistakenly that *for all purposes and at every moment of time* the law requires the separate existence of two different kinds of estate or interest in property, the legal and the equitable. There is no need to make this assumption.”
(emphasis mine)

⁴⁴ See e.g., *O'Rourke v Darbishire* [1920] AC 581, 626–627, where Lord Wrenbury remarked that the beneficiary's equitable proprietary right entitled the beneficiary there to have access to the trust documents.

⁴⁵ (n 34).

Professor Kevin Gray (1994, p. 163) also made a similar observation when he wrote that:

“An equitable right of property finds its origin not as a pre-existing component of some larger interest which is then hewn free as a block of equitable entitlement; instead it represents the result of a doctrinally-driven movement which impresses new rights upon the pre-existing estate under the mandate of the controlled conscience of equity.”

This means that the view of Tjio and Lee that equitable interest has to vest in the unitholders is simply a generalisation of the mechanics of trust law. This generalisation has to be particularised according to the context. In this regard, Professor Patrick Parkinson (2002) has astutely observed, *inter alia*, that there are non-charitable trusts which are valid despite the equitable interest not vesting in the beneficiaries. These trusts are namely the (1) discretionary trust and (2) *Re Denley* trust.

(i) Discretionary Trust

In a discretionary trust, the trustees hold trust property with a mandatory duty to choose from among the class of beneficiaries those who are to receive an interest in the trust property.⁴⁶ Prior to any allocation of the equitable interest, equitable interest in the trust property does not vest in anyone (contrast this with *Everton* (1982) who argued that a quasi-proprietary right vests in the beneficiary).⁴⁷ Accordingly, Hwang JC's statement in *Lee Chuen*

⁴⁶ *Mettoy Pension Trustees Ltd v Evans* [1991] 2 All E.R. 513 (Warner J).

⁴⁷ *Gartside v IRC* [1968] AC 553; *Vestey v IRC* [1980] AC 1148 (HL); *Re Brookes' Settlement Trusts* [1939] Ch 993, 997–998 (Farwell J).

*Li*⁴⁸ that “[l]ike nature, the law abhors a vacuum of ownership” cannot be regarded as applicable in all contexts, especially in the context of the discretionary trust which is a well-recognised form of the trust.

Although equitable interest does not vest, this does not prevent the persons in the class of beneficiaries from petitioning the court to enforce the trust. The right to seek the court’s assistance to supervise the administration of a discretionary trust does not depend on the vesting of any equitable interest.⁴⁹ Thus, in *Re the Trusts of the Abbott Fund*,⁵⁰ Stirling J held in *obiter* that although the two lady beneficiaries did not have a fixed interest in the trust fund, they could nevertheless apply to court to have the funds administered accordingly. Further, in *Gartside v IRC*,⁵¹ Lord Reid held that members of the class of beneficiaries have the right to be considered a potential recipient of the trust property.⁵² Lord Wilberforce also added that the members have the right to demand that the trustees exercise their discretion “fairly”, “reasonably” or “properly”.⁵³

(ii) *Re Denley* Trust

In *Re Denley*, Goff J upheld a trust over “land [to be] used as and for the purpose of a recreation or sports ground primarily for the benefit of the employees of the company and secondarily for the benefit of such other person or persons (if any) as the trustees may allow

⁴⁸ (n 5).

⁴⁹ *Schmidt v Rosewood Trust Ltd* [2003] 3 All ER 76 (PC) para 51. See also *Hartigan Nominees Pty Ltd v Rydge* (1992) 29 NSWLR 405; *AG for Ontario v Stavro* (1994) 119 DLR 750.

⁵⁰ [1900] 2 Ch 326.

⁵¹ (n 47).

⁵² *Ibid*, 606.

⁵³ *Ibid*, 618; see also *ibid*, 606 (Lord Reid), where His Lordship remarked that members can demand the trustee to act “in bona fide”.

to use". In the case, although the trust was worded as promoting a purpose, Goff J held that the beneficiary principle was not offended (*Mowbray et al* 2008, para.4-39).⁵⁴ His Honour construed the trust as one for the *factual* benefit of the employees, and this factual benefit entitled the employees to bring an action to enforce the trust. This proposition was approved and applied by Megarry J in *Re Northern Developments (Holdings) Ltd* (see also Pawlowski and Summers 2007, pp.443–445).⁵⁵ (but see *Re Grant's Will Trusts*,⁵⁶ where Vinelott J preferred to consider *Re Denley* as a discretionary trust; Matthews, on the other hand, suggested that *Re Denley* is a fixed trust of which "the intention of the settlor or testator was to benefit some person or persons, and hence the *purpose* is subordinate to the *benefit*". Under this analysis, the equitable interest appears to be vested in the employees, something which no court has agreed with (*Mowbray et al* 2008, para.4-40). Hence, it is unclear whether Matthew's view on this point is the preferable one.)

Hence, the *Re Denley* trust is another example of a trust where the beneficiaries (who are in the case the employees) could still enforce the trust without the equitable interest in the trust property vesting in them.

(2) *The S-REIT being a Re Denley Trust*

This author agrees with Tjio and Lee that the S-REIT closely resembles a *Re Denley* trust and could possibly be valid on that basis. The only difference in opinion lies with the fact that Tjio and Lee viewed the *Re Denley* trust as an exceptional non-charitable purpose trust while this author considers the same to be a trust for human beneficiaries. Tjio and Lee adopted

⁵⁴ *Re Denley* (n 7) 383–384.

⁵⁵ unreported, 6 October 1978 Ch. D. See also *Carreras Ltd v Freeman Mathews Ltd* [1985] Ch 207.

⁵⁶ [1980] 1 WLR 360, 370.

such a stance because the employees therein were not “traditional beneficiaries”, having merely a “negative *locus standi*” to enforce the trust (Tjio and Lee 2006, pp.52–53). Implicit in this reasoning is that the employees have no equitable interest vested in them, so the employees were not “traditional beneficiaries” for the purposes of the beneficiary principle. This probably led Tjio and Lee to conclude that the *Re Denley* trust is an exceptional non-charitable purpose trust.

This author respectfully disagrees with Tjio and Lee that the *Re Denley* trust is an exceptional non-charitable purpose trust. As seen in Part II I(1) above, the purpose of having the beneficiary principle is to enable the beneficiary to bring to the court’s attention of any breach of trust. Equity then vests an equitable interest in the beneficiary to give him or her *locus standi* to enforce the trust. However, as was discussed, enforcement of the trust can be effected even though the beneficiary is not vested with an equitable interest. This is seen in the case of the discretionary trust and *Re Denley* trust in Part III(1)(c) above. Further, this view is supported by Lord Walker in *Schmidt v Rosewood Trust Ltd*,⁵⁷ where His Lordship remarked:

“The right to seek the court’s intervention does not depend on entitlement to a fixed and transmissible beneficial interest. The object of a discretion (including a mere power) may also be entitled to protection from a court of equity”

It appears therefore that so long as there are *identifiable* beneficiaries, a court will allow those beneficiaries to enforce the trust (Pawlowski and Summers 2007, pp.444–445;

⁵⁷ (n 49).

Baxendale-Walker 1999, p.225).⁵⁸ In *Re Denley*, Goff J identified the employees as being the beneficiaries of the trust, entitling them to enforce the trust (but see Matthews (1996, p.14) who preferred to construe *Re Denley* as a trust which although is expressed as a purpose, “the intention of the settlor or testator was to benefit some person or persons, and hence the purpose is subordinate to the benefit” [emphasis inherent from source]). *Re Denley* thus opens up the possibility of a court making a finding of fact where necessary to identify beneficiaries in cases where it is unclear who the beneficiaries are. In the case of the discretionary trust, the identifiable beneficiaries are the objects who form the class of beneficiaries. As such, there is strictly no need for the vesting of an equitable interest to enforce a trust.

Since the unitholders are the identifiable beneficiaries (for the trust deed stipulates that the S-REIT assets are to be held on trust for the benefit of the unitholders), the unitholders should have the right to enforce the S-REIT (see also Trukhtanov 2007).⁵⁹ Accordingly, the S-REIT is a *Re Denley* trust for human beneficiaries, thereby not offending the beneficiary principle.

IV. IF THE S-REIT IS A NON-CHARITABLE PURPOSE TRUST, WILL IT NEVERTHELESS BE VALID?

While the S-REIT arguably does not offend the beneficiary principle from the analysis in Part III, a conservative court may hold a contrary view. Like *Tjio and Lee*, a court may be averse to finding that the beneficiary principle has been satisfied when equitable interest in a

⁵⁸ See also *Re Bowes* [1896] 1 Ch 507.

⁵⁹ See also *Citibank NA v QVT Financial LP* [2006] EWHC 3215 (Ch), where the English Court of Appeal decided to uphold the terms of the trust despite there being a possible breach of the irreducible core of trust obligations.

trust does not vest in the beneficiaries. This is especially so in the light of Lord Browne-Wilkinson's *obiter dictum* in *Westdeutsche Landesbank Girozentrale v Islington London Borough Council*,⁶⁰ where His Lordship remarked that when a trust has been established, the beneficiary has an equitable proprietary interest in the trust. Moreover, in *Baker v Archer-Shee*,⁶¹ the House of Lords by a bare majority held that equitable interest has to vest in a legatee-beneficiary after the administration of a deceased's estate. An application of this view would probably mean that the S-REIT will be regarded as a non-charitable purpose trust.

In this case, a better way to validate the S-REIT is to re-conceptualise the beneficiary principle as being premised on the enforcer principle. Tjio and Lee appeared to support this approach although they had some reservations about the adoption of a pure enforcer principle (Tjio and Lee 2006, pp.55 and 57). When the beneficiary principle is conceived as the enforcer principle, the question to ask is not whether there is a beneficiary with an equitable interest in the S-REIT assets. Instead, the question is whether there is an enforcer who has the right to hold the trustee accountable to the obligations undertaken (Thomas 2006, pp. 218–221). This author shall (1) proffer arguments to support the enforcer principle and (2) explain how the enforcer principle is applicable to validate the S-REIT in the ensuing paragraphs.

(1) *Arguments for the Enforcer Principle*

The enforcer principle was first articulated by Professor David Hayton (2001) in his article: *Developing the Obligation Characteristic of the Trust* (see also Hayton, Matthews and Mitchell 2007, paras.8.157–8.167; Hillard 2003). According to Hayton, a trust should be

⁶⁰ (n 39), where His Lordship remarked that a trust creates a legal and equitable title. See also *Abdul Hameed* (n 43).

⁶¹ (n 33).

valid on the basis that an enforcer has been identified to a trust (Leigh 1955, p.137). This author agrees with this view for two reasons. First, the right of the beneficiaries against the trustee is essentially *in personam* and not *in rem* in nature. Secondly, the enforcer principle does not conflict with the beneficiary principle.

(a) Beneficiary's right against the Trustee being *in personam* in nature

As discussed in Part III above, the main concern of courts is that trustees are not held accountable to the obligations they undertook and behave as though they are the absolute owner of the trust property. In order to prevent trustees from misbehaving, there is a need for someone to petition the court to enforce the trust. From the analysis of the beneficiary principle in Part III above, it can be seen that this duty has been conveniently vested in the beneficiaries. This has led to some commentators arguing that the beneficiary's right against the trustee is *in rem* in nature (Scott 1917; Martin 2009, p.19). Their main justification for the *in rem* right is syllogistic: since an equitable interest corresponds to a legal estate and the beneficiary is entitled to the equitable interest, the substance of the transaction gives the beneficiary an *in rem* right in the estate (Martin 2009, p.19). Thus, when a beneficiary enforces a trust, the beneficiary is arguably vindicating the *in rem* right it has against the trust property.

However, this author submits that the *in rem* analysis of the beneficiary's right against the trustee is an incorrect view, for this leads to the assumption that there has always to be two owners over the same trust property (see also Stone 1917; Hayton 2001; Parkinson 2002, Part II; Duckworth 1998, p.23; but Pearce and Stevens (2006, p.111) and Thomas and Hudson (2010, paras.1.41 and 7.01) argued that a trust is both proprietary and obligatory;

while Martin (2009, p.22) suggested that the beneficial interest in a trust is *sui generis* in nature). As noted by Professor Frederic Maitland (1936, p.17), a beneficiary can not be the owner of the trust property when the trustee is already the owner:

“An examiner will sometimes be told that whereas the common law said that the trustee was owner of land, equity said that the *cestui que trust* was the owner. Well here in all conscience there seems to be conflict enough. Think what this would mean were it really true. There are two courts of co-ordinate jurisdiction – one says that A is owner, [while] the other says that B is the owner of Blackacre. That means civil war and anarchy. ... Equity did not say that the *cestui que trust* was the owner of the land, it said that the trustee was owner of the land, but added that he was bound to hold the land for the benefit of the *cestui que trust*.”

Thus, the right which the beneficiary has against the trustee is essentially *in personam* in nature. If an owner undertakes to hold the property for the benefit of another, though he is the owner at law, equity will act *in personam*⁶² (see also Maitland 1984, p. 258) against him or her to perform according to the trust obligations undertaken. As such, the beneficiary acquires an *in personam* right against the trustee. Should one insist on classifying the beneficiary's right as an *in rem* right, the beneficiary has, at the very most, only a *negative in rem* right which entitles the beneficiary primarily to exclude third parties from the enjoyment of the trust property and secondarily to recover the trust property from non-*bona fide* purchasers for value without notice of the trust (Nolan, 2006). However, this negative *in rem* right does not necessarily require the equitable interest to vest in the beneficiary. The reason is that this negative *in rem* right is derived from the positive *in rem* right that the trustee and

⁶² *Penn v Lord Baltimore* (1750) 1 Ves Sen 447 1134–1135 (Lord Hardwicke).

his or her transferees possess over the trust property. The net result is that a beneficiary is still acting *in personam* against the trustee and the subsequent transferees.

Hence, while Isaacs J in *Hoystead v Federal Commissioner of Taxation*⁶³ (see also Pearce and Stevens 2006, p.97) acknowledged that “Equity regards the cestui que trust as the true owner of the property itself,” His Honour qualified that view by remarking that: “[b]ut it is, nevertheless true that equity acts only *in personam* and the rights it recognizes and enforces are rights *in personam* and not rights *in rem*.”⁶⁴ Further, as Gray (1994, p. 163) has observed: “[e]quitable rights of property ... derive from conscientious obligations to deal with an asset or resource in a certain way.”

That equity fundamentally wields an *in personam* power was also astutely observed by Gibson J in *Carreras Rothmans Ltd v Freeman Mathews Treasure Ltd*⁶⁵:

“[E]quity fastens on the conscience of the person who receives from another property transferred for a specific purpose only and not therefore for the recipient’s own purposes, so that such person will not be permitted to treat the property as his own or use it for other than the stated purpose.”⁶⁶

This necessarily follows that the separation of the legal and equitable estate in a trust is merely a mechanism to aid a court’s analysis with regard to *locus standi* to enforce the trust

⁶³ (1920) 27 CLR 400, 422.

⁶⁴ *Ibid*.

⁶⁵ (n 55) 222.

⁶⁶ See also *Westdeutsche* (n 39), where Lord Browne-Wilkinson also remarked that equity affects the conscience of the owner of the legal interest.

and possibly the tracing of trust property.⁶⁷ This was alluded to by Lord Radcliffe in *Commissioner of Stamp Duties (Qld) v Livingston*,⁶⁸ where His Lordship stated that:

“When the whole right of property is in a person [e.g. a trustee] ... there is no need to distinguish between the legal and equitable interest in that property, any more than there is for the property of a full beneficial owner. *What matters is that the court will control the executor in the use of his rights over assets that come to him in that capacity; but it will do it by the enforcement of remedies which do not involve the admission or recognition of equitable rights of property in those assets. Equity in fact calls into existence and protects equitable rights and interests in property only where their recognition has been found to be required in order to give effect to its doctrines.*”
(emphasis mine)

Although *Baker v Archer-Shee*⁶⁹ stands for the proposition that equitable interest has to vest in the beneficiary, thereby creating some *in rem* right in favour of the beneficiary, the case can be confined to its facts, for the case involved the analysis of the beneficiary's right in the specific context of taxation. Thus, for taxation purposes, the beneficiary was considered to have an equitable interest in the trust fund (Martin 2009, pp.21–22). This rule, however, cannot extend to all species of trusts for there are trusts such as the discretionary trust and *Re Denley* trust which are valid despite no equitable interest being vested in the beneficiaries. It must also be noted that Lords Sumner and Blanesburgh who dissented in the *Baker v Archer-Shee*⁷⁰ adopted an *in personam* approach rather than a proprietary approach

⁶⁷ See *Re Diplock* [1948] Ch 465; *Foskett* (n 33).

⁶⁸ (n 34).

⁶⁹ (n 33).

⁷⁰ (n 33).

towards the analysis of the beneficiary's interest. That one has to analyse a trust in its specific context but without losing sight of the basic premise of trust law was acknowledged by Lord Browne-Wilkinson in *Target Holdings v Redfern*:⁷¹

“[I]n my judgment it is in any event wrong to lift wholesale the detailed rules developed in the context of traditional trusts and then seek to apply them to trusts of quite a different kind. In the modern world the trust has become a valuable device in commercial and financial dealings. *The fundamental principles of equity apply as much to such trusts as they do to the traditional trusts in relation to which those principles were originally formulated. But in my judgment it is important, if the trust is not to be rendered commercially useless, to distinguish between the basic principles of trust law and those specialist rules developed in relation to traditional trusts which are applicable only to such trusts and the rationale of which has no application to trusts of quite a different kind.*” (emphasis mine)

Therefore, as can be seen from the above discussion, the right acquired by a beneficiary against the trustee is essentially *in personam* and not *in rem* in nature. The most fundamental principle still is that courts will enforce obligations undertaken by trustees to prevent them from dealing with trust property as though they are the absolute owners, and this is achieved by the active enforcement by the beneficiaries through the exercise of their *in personam* right against the trustee. There is strictly no involvement of *in rem* rights between the beneficiaries and trustee *per se*.

⁷¹ [1996] AC 421, 435.

Hence, the true test of the validity of a trust is whether there is someone who possesses an *in personam* right against the trustee and is thereby able to bring to the courts' attention of any breach of trust (i.e., an enforcer).⁷² An application of this test to the differing categories of trusts shows that this is the case: typical non-charitable trusts are valid because the beneficiary has the right to demand that the trustee perform faithfully the duties undertaken by petitioning the court to enforce the trust; charitable trusts are also valid because the Attorney-General can hold the trustee accountable to the trust obligations undertaken by bringing to the court's attention of any breach of trust;⁷³ the discretionary and *Re Denley* trusts are valid because there are beneficiaries who have the right to enforce the respective trusts should the trustee breach his or her fiduciary duties. That the presence of an enforcer is sufficient to validate a trust can be further gleaned from *Re Thompson*,⁷⁴ where Clausen J upheld a non-charitable purpose trust for the "promotion and furthering of fox-hunting" despite such trusts being typically invalid for want of beneficiaries. The *Re Thompson*⁷⁵ trust was upheld simply because the residual legatees had the right to enforce the trust should the trustee there not perform his obligations (but see Pawlowski and Summer (2007, p.446) who argued that exceptional non-charitable purpose trusts may be "policy driven exceptions validated because ordinary people expect them to be valid).

(b) The Enforcer Principle does not conflict with the Beneficiary Principle

Further, since the beneficiary principle is one of the sacrosanct principles of trust law, the enforcer principle should be acceptable as an analytical tool for validating trusts if it does not

⁷² See *Morice* (n 26) (Sir Grant MR).

⁷³ *Leahy* (n 4) 479.

⁷⁴ [1934] Ch 342.

⁷⁵ *Ibid.*

conflict with the beneficiary principle. As seen in Part III above, the purpose of the beneficiary principle is to ensure that the beneficiaries can notify the courts of any breach of trust. The enforcer principle performs the same function albeit the enforcer can be someone other than the beneficiary. Hence, the enforcer principle is the overarching explanation of the beneficiary principle. In this case, this author submits that there is no conflict between the two principles, so the enforcer principle should be an acceptable tool to analyse the S-REIT.

However, this author notes that a potential weakness of the enforcer principle is that disinterested third parties can be designated as the enforcer. Unlike beneficiaries, disinterested third-party enforcers may not perform the enforcer role well, for their interests are unlikely to be jeopardised if they should perform their role defectively. Thus, to allow a disinterested third-party to perform its role defectively is an affront to the fundamental principle that courts enforce obligations undertaken by trustees to prevent them from dealing with the trust property as though they are the absolute owners. As Parkinson (2002, p.680) has observed, to ensure proper enforcement, the enforcer has to be one who is interested in enforcing the trust, for a disinterested enforcer will unlikely complain to the court of any breach of trust. This was probably the unarticulated consideration of the courts in the twentieth century post *Morice v Bishop of Durham*⁷⁶ when they decided that a non-charitable trust, unlike a charitable trust which is enforceable by the Attorney-General, has to be for the benefit of persons. It may then be imperative for equity to validate a trust on the basis of the enforcer principle only when the enforcer is pecuniarily, legally (Tan 2009, p.485–489)⁷⁷ or morally interested in the trust (but see Tjio and Lee (2006, pp. 55–56) who

⁷⁶ (n 26).

⁷⁷ See Special Trusts (Alternative Regime) Law, 1997 (Cayman Islands) s 8(2).

advocated the use of a negative equitable property right or negative *locus standi* instead of a mere enforcer).

In any case, the S-REIT does not suffer from this potential weakness because the unitholders are interested parties in enforcing the S-REIT. Since unitholders usually hold on to the units to receive returns on their investment, it would be foolhardy for unitholders not to be interested in enforcing the S-REIT. A breach of trust will almost certainly affect the rate of return for the unitholders, so it is absolutely in the interest of the unitholders to enforce the S-REIT.

(2) *The Enforcer Principle and the S-REIT*

As seen from the discussion above, the beneficiary's right against a trustee is essentially *in personam* in nature, and an enforcer is all that is required to validate a trust. It necessarily follows that should the S-REIT be considered as a non-charitable purpose trust, it is nevertheless valid because the unitholders take on the role of the enforcer. It is sufficient that the trustee expressly owes an obligation to hold the S-REIT assets for the benefit of the unitholders, and the unitholders are correspondingly conferred a power by the trust deed to enforce the terms of the S-REIT against the trustees. As a result, there is no need for equitable interest (and consequently proprietary interest) to vest in the unitholders.

V. CONCLUSION

In this paper, this author has argued that despite the non-vesting of the equitable interest in the unitholders, the S-REIT does not offend the beneficiary principle. The beneficiary

principle was developed for the simple reason of ensuring that there is someone who has *locus standi* to bring to the court's attention of any breach of trust. Thus, the courts have conveniently given *locus standi* to the beneficiaries by vesting an equitable interest in the beneficiaries. However, in the light of the validity of non-charitable trusts such as the discretionary trust and the *Re Denley* trust, there appears to be no need for an equitable interest to vest in order to give a beneficiary *locus standi* to enforce a trust.

While this author agrees with Tjio and Lee that the S-REIT is essentially a *Re Denley* trust and could be validated as such, this author prefers to view *Re Denley* as a trust for human beneficiaries and not an exceptional non-charitable purpose trust. This can be gleaned from the fact that the discretionary trust and *Re Denley* trust are upheld even when equitable interest does not vest in the beneficiaries. It appears from these cases that trusts which do not vest equitable interest in the beneficiaries will be upheld if there are identifiable beneficiaries who can enforce the trust. It therefore follows that the beneficiary principle will be satisfied even when equitable interest in the S-REIT does not vest in the unitholders because the unitholders are the identifiable beneficiaries who have the requisite *locus standi* to enforce the S-REIT.

However, assuming the beneficiary principle is not satisfied unless equitable interest vests in the unitholders, thereby rendering the S-REIT an invalid non-charitable purpose trust, the beneficiary principle should be re-conceptualised as being based on the enforcer principle. Under the enforcer principle, a trust is valid if there is an enforcer who has the right to bring to the courts' attention of any breach of trust. In this way, the S-REIT can be seen as a valid trust, for the unitholders have a right provided by the trust deed to enforce the terms of the S-REIT against the trustees.

VI. Acknowledgements

The author would like to thank Associate Professor Tang Hang Wu from the National University of Singapore and an anonymous referee for their invaluable comments on the earlier version of this paper. The author would also like to thank Mr. Andrew Chan from Allen & Gledhill LLP for his pertinent insights on the Singapore Real Estate Investment Trust and Saizen Real Estate Investment Trust for readily providing the relevant copies of their trust deed for the purposes of this paper. All responsibility for deficiencies remains solely with the author.

VII. References

- Armstrong and Ulmer (2006), "Canada" in Booth, R. and Boyle, C., eds. *Real Estate Investment Trusts: A Global Analysis*, Globe Business Publishing, London, p. 41.
- Ascendas REIT Prospectus, <http://areit.listedcompany.com/misc/maintextfinal.pdf>, viewed: 01/06/2010, pp.146 and 151.
- Baxendale-Walker, P. (1999), *Purpose Trusts*, Butterworths, London, pp. 145–147 and 225.
- Borrows, J. and Smith, H.(2002), "Introduction" in Borrows, J. gen. ed. *Current Issues in Securitisation*, Sweet & Maxwell, London, pp.1–4.
- Bridge, M. (2002), *Personal Property Law*, 3rd ed., Oxford University Press, Oxford, p.12.
- Burns, T., "Developing a Franchise: Could Securitisation be a Serious Funding Option for Franchisors in the United Kingdom" [2006] JBL 656.
- Campbell, P.J., Samuels, J.B. and Yuasa, M. (2006), "United States" in Booth, R. and Boyle, C., eds. *Real Estate Investment Trusts: A Global Analysis*, Globe Business Publishing, London, p.219.
- Chan, S.H., Erickson, J. and Wang, K. (2003), *Real Estate Investment Trusts: Structure, Performance, and Investment Opportunities*, Oxford University Press, New York, p.3.
- Chitty, J. (2008), *Chitty on Contracts*, Vol. 1, 30th ed. (Beale ed.), Sweet & Maxwell, London, paras.3-001–3-003.
- Deveson et al (2006), "Australia" in Booth, R. and Boyle, C., eds. *Real Estate Investment Trusts: A Global Analysis*, Globe Business Publishing, London.
- Duckworth, A. (1998), "The Colony Strikes Back", 12 TLI 16, 23.
- Everton, A.R., "Towards a Concept of Quasi-Property" [1982] Conv 118.
- Finch, V. (2008), "Corporate Rescue in a World of Debt" 8 JBL 756, 759–761.
- Gooi and Teong (2006), "Malaysia" in Booth, R. and Boyle, C., eds. *Real Estate Investment Trusts: A Global Analysis*, Globe Business Publishing, London.
- Gray, K. "Equitable Property" [1994] CLP 157, 163.
- HM Treasury (2005), *UK Real Estate Investment Trusts: A Discussion Paper*, <http://www.hm-treasury.gov.uk/d/Bud05Reits.pdf>, viewed: 01/06/2010, pp.5–6.
- Hayton, D. (2001), "Developing the Obligation Characteristic of the Trust", 117 LQR 96.
- Hayton, D., Matthews, P., and Mitchell, C. (2007) *Underhill and Hayton, Law Relating to Trusts and Trustees*, 17th ed., LexisNexis Butterworths, London, paras.2.12, 2.15, 8.144, 8.157–8.167.
- Hillard, J. (2003), "On the Irreducible Core Content of Trusteeship – A Reply to Professors Matthews and Parkinson", 17 TLI 144.

- Hovanesian, M.D. and Conlin, M. (2002), "Wall Street's Broken Spirits", *Business Week*, http://www.businessweek.com/magazine/content/02_35/b3797088.htm, viewed: 01/06/2010.
- Koh, J. (2006) "Singapore" in Booth, R. and Boyle, C., eds. *Real estate Investment Trusts: A Global Analysis*, Globe Business Pub, London, pp.181 and 182.
- K-REIT Introductory Document, http://www.kreitasia.com/pdf/KREITAsia_IntroDoc.pdf, viewed: 01/06/2010, pp.120 and IV-26.
- Lee, S.F. and Foo, L.E. (2010), "Real Estate Investment Trusts in Singapore Recent Legal and Regulatory Developments and the Case for Corporatisation", 22 SAcLJ 36, 40 and 61–64.
- Leigh, L. (1955), "Trusts of Imperfect Obligation", 18 MLR 120, 137.
- Leung and Smith (2006), "Hong Kong" in Booth, R. and Boyle, C., eds. *Real Estate Investment Trusts: A Global Analysis*, Globe Business Publishing, London.
- MAS (2002), Appendix 2 (Property Funds), *Code of Collective Investment Schemes*, http://www.mas.gov.sg/resource/news_room/press_releases/2005/CIS_Code_20_Oct_2005.pdf, viewed: 01/06/2010, para.6.4.
- MAS (2002), *Code of Collective Investment Schemes*, http://www.mas.gov.sg/resource/news_room/press_releases/2005/CIS_Code_20_Oct_2005.pdf, viewed: 01/06/2010.
- Maitland, F.W. (1936), *Equity; A Course of Lectures*, Cambridge [Eng.] University Press, p.17.
- Maitland, F.W. (1984), *Equity, also the Forms of Action at Common Law, Two Courses of Lectures* (Chaytor A.H. and Whittaker W.J. eds.) F.B. Rothman, Littleton, Colo, p.258.
- Matthews, P. (1996), "The New Trusts: Obligations Without Rights?" in Oakley, A.J., ed. *Trends in Contemporary Trust Law*, Clarendon Press, Oxford, pp.1 and 13–15.
- Matthews, P. (2002), "From Obligations to Property, and Back Again?" in Hayton, D., ed. *Extending the Boundaries of Trusts and Similar Ring-fenced Funds*, Kluwer Law International, London, p.223
- Matthews, P. (2006), "The Comparative Importance of the Rule in *Saunders v. Vautier*" 122 L.Q.R. 266
- Martin, J. (2009), *Hanbury and Martin: Modern Equity*, 18th ed., Sweet & Maxwell, London, pp.19 and 21–22.
- McFarlane, B. (2008), *The Structure of Property Law*, Hart, Oxford, pp.132 – 133.
- Meagher, R.P., Heydon, J.D. and Leeming, M.J. (2002), *Meagher, Gummow and Lehane's Equity: Doctrines and Remedies*, 4th ed., Butterworths, Chatswood, NSW, paras.3-010–3-030.
- Mowbray, J. et al (2008), *Lewin on Trusts*, 18th ed., Sweet & Maxwell, London, paras.1-06 and 4-38–4-40.
- Nolan, R. (2006), "Equitable Property", 122 LQR 232.
- Parkinson, P. (2002), "Reconceptualising the Express Trust", 61(3) CLJ 657, 680 and Part II.
- Pawlowski, M. and Summers, J. (2007), "Private Purpose Trusts – A Reform Proposal", Conv 440, pp.443–446.
- Pearce, R.A. and Stevens, J. (2006), *The Law of Trusts and Equitable Obligations*, 4th ed., Oxford University Press, Oxford, pp.97 and 111.
- Penner, J.E. (2008), *The Law of Trusts*, 6th ed., Oxford University Press, New York, pp.27–29.
- Saizen REIT Trust Deed, on file with author.

- Schaefer (2005), "Understanding REITs: Fundamental Concepts and Structures" in Wright D., ed. *REITS in Asia from Concept to Completion*, Asia Law & Practice, Hong Kong, p.5.
- Scott, A.W. (1917), "Nature of Rights of the Cestui Que Trust", 17 Col LR 269.
- SGX Real Estate Investment Trusts (2010), http://www.sgx.com/wps/portal/marketplace/mp-en/prices_indices_statistics/securities/reits, viewed: 01/06/2010.
- SGX Listing Manual (2010); <http://info.sgx.com/weblist.nsf/listing+manual?openview>, viewed: 01/06/2010.
- Sin, K.F. (1997), *The Legal Nature of Unit Trust*, Oxford University Press, New York, pp.69 and 167.
- Stone, H.F. (1917), "Nature of the Rights of the Cestui Que Trust", 17 Col LR 467.
- Suntec REIT Prospectus, [http://info.sgx.com/webipo.nsf/IPOAttachmentsName/SUNTEC%20REAL%20ESTATE%20INV%20TRUST/\\$File/suntec_final\(C\).pdf](http://info.sgx.com/webipo.nsf/IPOAttachmentsName/SUNTEC%20REAL%20ESTATE%20INV%20TRUST/$File/suntec_final(C).pdf), viewed: 01/06/2010, pp.151 and p.II-25.
- Tan (2005), "Singapore: Evolution and Future Development of the REIT Market" in Wright D., ed. *REITS in Asia from Concept to Completion*, Asia Law & Practice, Hong Kong, pp.109 and 111.
- Tan, T. (2009), "The Irreducible Core Content of Modern Trust Law" 15(6) T&T 477, 485–489.
- Tjio, H. (1993), "Contract, Property and Unincorporated Associations" SAcLJ 112, 114.
- Tjio, H. (2004), *Principles and Practice of Securities Regulation in Singapore* (Singapore: LexisNexis, 2004), paras.3.16 –3.21.
- Tjio, H. and Lee, S.F. (2006), "Developments in Securities Law and Practice" in Teo, K.S., gen. ed., *Developments in Singapore Law between 2001 and 2005*, Singapore Academy of Law, Singapore, pp. 50–57.
- Thomas, G. (2006), "Purpose Trusts" in Glasson, J. and Thomas, G., eds. *The International Trust*, 2nd ed., Jordans, Bristol, pp.218 –221
- Thomas, G. and Hudson, A. (2010), *The Law of Trusts*, 2nd ed., Oxford University Press, Oxford, paras.1.41, 6.13–6.20, 6.40–6.48 and 7.01.
- Trukhtanov, A. (2007), "The Irreducible Core of Trust Obligations" 123 LQR 342.
- Zieminski, N. (2010), "Global REIT Values to Grow in 2010 – Ernst & Young", *Reuters*, <http://www.reuters.com/article/idUKN1014767820100310>, viewed: 01/06/2010.

The modern position of the statutory authority defence in the UK

Francis Moor

School of the Built Environment,
University of Salford, Salford, M5 4WT,
United Kingdom
Email: f.moor@pgr.salford.ac.uk

Abstract:

This paper aims to evaluate the impact of the new s.158 Planning Act 2008 on the existing doctrine of statutory authority.

The Planning Act 2008 has introduced many changes pertaining to the planning and development of infrastructure projects in the UK. This paper will briefly overview those changes, before focusing on S.158 of The Act, which awards developers statutory immunity from an action in private nuisance. The paper will address the availability of certain actions, in both common law and statutory nuisance, before assessing the impact of any immunity or statutory authority defence on such actions. It will then be important to identify how a history of widening the statutory authority defence has affected its use in practice, and to consider whether s.158 could potentially extend the application of such a defence, before moving on to clarify the existing exceptions to the statutory authority rule. These exceptions have been created by the judiciary's antipathy towards overt infringement on property owners' rights. Therefore, in awarding the defence they will take into account a variety of considerations such as the specificity of the location, the inevitability of the nuisance, the effect on statutory provisions, and the bearing that negligence has on the defence.

Although s.158 of the Planning Act 2008 appears to broaden the opportunity for an exemption defence to be used, many restrictions remain on its application. There is no such thing as 'blanket immunity', and developers may still be held responsible for their un-neighbourly conduct. This is something that developers, and their advisers, will have to be aware of before seeking to rely on the statutory immunity defence.

Keywords: Statutory authority, Nuisance, Infrastructure planning, Tort.

1: Introduction

The Planning Act 2008 was introduced with the intention of streamlining and improving procedures in the United Kingdom for the planning and development of major projects such as power stations, airports and highways. The prescribed measures were a governmental response to the delays and costs incurred in the 7-year planning process for Terminal 5 at Heathrow Airport. The Act creates a range of measures, which have been introduced to ensure that the planning process for these projects, now referred to as Nationally Significant Infrastructure Projects (NSIPs), is conducted in a swifter, more transparent and clearly prescribed manner. In addition to these provisions, which create new policy statements, alter the consultation procedures, make changes to public participation opportunities, and create a new planning application process, the Government proffered statutory authority on such projects under s.158 of the Act. This section, which came into effect on March 1st 2010, awards developers of NSIPs a defence against actions in private or public nuisance. The application of such a defence could have serious implications for neighbours of infrastructure developments, who may suffer adversely as a result of the build, or the operation of the actual development. Although the use of the defence in such situations is not novel or original, this paper argues that by extending the scope of such protection, in the way set out in the Planning Act 2008, the legislature can be seen to significantly depart from the original ideologies of the statutory authority defence, which purported to defend public officials who may have committed a nuisance in the performance of their duties.

This paper will firstly direct its attention to a brief overview of the proposed procedural changes for the new NSIP planning process and then proceed to outline the legal principles of common law and statutory nuisance. The paper will then evaluate the governmental move from a 'blanket immunity' approach for towards nuisance actions to one of 'statutory authority' for developers of NSIPs. It will then be important to review the history of the statutory authority defence, in order to assess what such an authorisation will mean, in effect, for such developers. This paper will then demonstrate how the defence has been conferred in the Planning Act 2008, and consider whether the new provisions extend the scope of its application, before moving on to consider the inherent restrictions to the defence which have been carved out to protect private property owners. It is apparent from over two centuries of case law, that certain restrictions apply to the use of the statutory authority defence; this case law will be relevant in determining how any immunity will be approached in practice.

2: Making the system work

Nationally Significant Infrastructure Projects are, for the first time, set to have their own dedicated planning system, which aspires to achieve the Labour Government's desired outcome of "*a single, unified development consent regime*" (Communities & Local Government etc. 2007, 78). The impetus for this policy and procedural change was the lengthy planning application process and inquiry stage for Heathrow's Terminal 5. The inquiry, which sat for 525 days over a 4 year period, was the longest of its kind in the UK, at a cost of £17 million to the Government, and created an awareness for the need to create a more streamlined and efficient system (Department for Transport 2005). In order to achieve this aim, the Government introduced new procedures and institutions which would organise and administrate the process. The new framework strives to achieve a *trias politica*, or separation of powers, between the legislature (ministerial drafting of National Policy Statements), executive (consideration of each application by the relevant planning authority), and judiciary (appeal available through judicial review). This separation of duties and roles should ideally act as a filter system, ensuring that an accountable body is making policy decisions that are then applied by an impartial and independent body in considering applications, with both stages reviewable through the legal system.

In determining whether certain projects are necessary, and where they should be situated, Government ministers are to produce new National Policy Statements (NPSs). These aim to simplify the system by avoiding repeat enquiries; for example "*... if ministers have said once that they consider nuclear power to be safe and necessary, then it should not be necessary for this to be tested every time an application is made*" (Bassford 2007, 150). Developers must be aware of the NPS for that type of development before making any application, and must also have conducted any consultation procedures prescribed by the Act. It was proposed that applications would then be considered by a newly appointed Infrastructure Planning Commission (IPC), who would determine whether the applications they received were in accordance with the NPS on such matters and/or locations. However, the new coalition Government has committed to abolishing the IPC (as was evidenced in both parties' election manifestos) and will instead legislate for the consideration of applications by a 'Major Infrastructure Planning Unit' in the Planning Inspectorate,¹ thereby countering the anti-democratic argument of opponents to the IPC.

¹ To be implemented by Decentralisation and Localisation Bill, proposed in Queen's Speech, May 25th 2010.

The need to streamline the planning system for major developments, and increase efficiency, transparency and productivity for such projects, is unquestionable, but at times the recent reforms have led to vehement opposition. Commentators such as Ellis (2008), Thompson (2009), and Bassford (2007), have negatively critiqued the new measures in relation to; their impact on public consultation, the removal of certain hearing processes, and the aforementioned lack of democratic accountability of the unelected IPC. One aspect of the recent measures which has received little treatment however, is the abrogation of neighbours' rights to an action in private or public nuisance, with the Government using the pre-existing defence of statutory authority. This paper will first address the meaning of common law and statutory nuisance before exploring the notions of immunity or a statutory authority defence.

3: What Constitutes a Nuisance?

Nuisance has been described by Blackstone (1768) as:

... anything that works hurt, inconvenience or damage. And nuisances are of two kinds: public or common nuisances, which affect the public, and are an annoyance to the King's subjects; ... and private nuisances, which ... may be defined, anything done to the hurt or annoyance of the lands, tenements or hereditaments of another. (Blackstone, 1768, p.216)

Such a definition has been acknowledged as excessively broad however, in relation to both public and private nuisances.

Private nuisance actions have been restricted to apply in situations where there has been unlawful interference with land, which the claimant has an interest in, and where it can be proven that the defendant is legally responsible for the unlawful interference (Elvin & Karas 2002). The common law doctrine of private nuisance protects landowners from interferences with either their land, or their enjoyment of it, and has been referred to as the realm of the "*stinking privy, the urban hog-sty, the fouled or diverted stream, the polluting chimney*" (Brenner, 1973, p. 403). The three main forms of private nuisance, identified in *Hunter v. Canary Wharf*², are:

- encroachment on a neighbour's land
- physical injury to a neighbour's land

² [1997] 2 All E.R. 426.

- substantial interference with a neighbour's quiet enjoyment of his land

Nuisance actions seek to address potential conflicts in land usage, and balance these competing interests, in view of what is 'reasonable'³ in the circumstances. It can be argued that this judiciary measure has been historically successful in balancing such competing interests in regards to the nature of the locality, the duration and frequency of the defendant's conduct, the utility of the defendant's conduct, any abnormal sensitivity of the claimant, and malice on the part of the defendant (Giliker & Beckwith, 2008). The direction given is for the judiciary to consider these matters with a pragmatic approach, rather than pandering to "*the elegant or dainty modes and habits of living, but according to the plain and sober and simple notions among the English people*"⁴. This paper will later direct its attention to the way in which the legislature has circumvented the capacity for such judicial intervention in certain land usage disputes.

In consideration of public nuisance it is appropriate to begin with another broad definition, as used by academics such as McLaren (1972), Spencer (1989), and Coyle & Morrow (2004), to illustrate the ambitious ideals of the principle. Archbold's Criminal Pleading (2010) states:

A person is guilty of a public nuisance..., who (a) does an act not warranted by law, or (b) omits to discharge a legal duty, if the effect of the act or omission is to endanger the life, health, property, morals, or comfort of the public, or to obstruct the public in the exercise or enjoyment of rights common to all Her Majesty's subjects (Archbold, 2010).

In reality however, public nuisance was most effective in providing actions for the protection of health and safety for the public at large in the centuries before legislative measures such as the Public Health Act 1936, the Highways Act 1980 and, most recently, environmental legislation such as the Environmental Protection Act 1990. As Spencer (1989) noted in his seminal work on public nuisance: "*Over the last hundred years ... virtually the entire area traditionally the province of public nuisance prosecutions has been comprehensively covered by statute*" (Spencer, 1989, p.76). Although the doctrine of public nuisance appears increasingly redundant in the modern criminal and tort landscape the judiciary have recently upheld its existence, in the *Goldstein*⁵ case, with the proviso that statutorily defined offences be preferred in most instances.

³ Described in *Sedleigh-Denfield v. O'Callaghan* [1940] A.C. 880 at 903.

⁴ *Walter v. Selfe* (1851) 64 E.R. 849 quoted in Kidner, 1998 at 276.

⁵ *R. v Goldstein* [2005] UKHL 63.

Statutory nuisance has evolved as a legislative way of prohibiting certain activities which could be deemed harmful to the health, well-being, or property interests of the public and has been formed and shaped by the harms identified in the common law doctrines of public and private nuisance. Statutory nuisance thus “*reworks common law nuisance, taking some of its more practical features for remedying interference with property and separating them from the more onerous procedural and standing requirements*” (Coyle & Morrow, 2010, p.126). Powers are given to local authorities⁶ and to certain individuals⁷, under statutory provisions, to enable them to request the abatement of the nuisance; however it has been noted that courts tend to favour claimants who have experienced interference which affects personal comfort (Coyle & Morrow 2010). Thus the application of statutory nuisance may be more restrictive than that of the common law approaches and claimants could still benefit from the availability of a common law action, which may not be possible where defendants have been awarded statutory authority. The statutory authority defence can be seen to infringe on certain rights in action for property owners; this paper will explore the impact of the defence on such rights in a subsequent chapter.

The Government have however, been impelled to arrive at the current provision, which grants the defence of statutory authority to developers of NSIPs. The original proposal, which strongly mirrored the focus on streamlining and efficiency, was a complete or ‘blanket’ immunity from all forms of action in private, public or statutory nuisance for such developers. This paper will now examine the reasons for the change in position; from a clause which proposed total immunity, to a section which awards the defence of statutory authority.

4: The opposition to ‘blanket immunity’

On the 25th of June 2008, M.P. Clive Betts tabled a clause for insertion in the Planning Bill, removing the right to an action in nuisance, stating that;

No proceedings, whether criminal or civil, in nuisance and no civil proceedings in respect of the escape of things from land, other than proceedings for breach of statutory duty, may be brought in relation to development, works or operations authorised by an order granting development consent. (Betts, 2008 Hansard HC Col.371)

⁶ S.80 Environmental Protection Act 1990.

⁷ ‘Persons aggrieved’ under s.82 Environmental Protection Act 1990.

The proposed aim of the measure was to mimic the existing exemptions for nuisance actions afforded under other legislative measures, such as the provisions set out in s.122 of the Railways Act 1993. However the ambiguity of phrasing in the Planning Act led to confusion for practitioners, and even the legislature, with Hazel Blears envisaging a scenario which would result in the removal of any responsibility for the developers in the creation of statutory nuisances (Blears, 2008). Emotive headlines such as, “*Hundreds of thousands will not be able to complain about noisy building sites*” (Hope, 2008) and “*The Planning Bill: a bulldozer through democracy*”, (Cavendish, 2008) heightened concerns and generated the requisite publicity for an anti-governmental backlash.

In October 2008, what has been referred to as an ‘*unprecedented alliance*’ (Miner, 2008), consisting of professionals, council leaders, lawyers, and environmental organisations, joined forces to sanction for the removal of clause 151 from the Planning Bill. The alliance consisted of Environmental Protection UK (EPUK) and the Local Authorities Co-ordinators of Regulatory Services (LACoRS), as well as CPRE (Campaign to Protect Rural England), Friends of the Earth (FoE), and the Local Government Association (LGA). CPRE and FoE, as well as being part of the alliance against clause 151, were leading members of the ‘better planning coalition’ which also included organisations such as the NSPB, Plantlife International, The National Trust and Civic Trust. These parties strengthened the position against the proposed planning reforms with the support of over 5 million members.

Concerns highlighted by the leaders of the opposition, EPUK and LACoRS, included the following;

1. The clause did not appear to recognise the authority of case law in dealing with nuisances successfully in the past.
2. There was no incentive for developers to mitigate effects on local communities.
3. The clause created a lack of clarity as to the scope and application of new environmental protection legislation which should protect from nuisances.
4. Communities would have no voice once consent was given, and no authority to turn to if there was any interference with their enjoyment of their land.
5. Would the nuisance immunity be retrospective in order to cover pre-existing infrastructures?
6. Did the clause abrogate rights enshrined in the ECHR and HRA, which entitle citizens to respect for their home and peaceful enjoyment of their possessions? (EPUK, 2008)

These concerns and requests for clarity were echoed by a number of environmental campaign groups and organisations, with the disquiet spreading and increasing in volume. The Chartered Institute for

Environmental Health (CIEH) added its support to the protest against clause 151, asking for justification from the Government for the application of statutory immunity in such a broad and sweeping fashion;

... it appears to offer protection to activities which currently do not attract (and hitherto have not thought to need) statutory immunity or even the benefit of the older common law doctrine and unlike previous applications, its scope is not limited by a statutory purpose; rather it applies to a long list of projects by virtue of the vehicle chosen to permit their development (Price, 2008).

The CIEH viewed such broad application as ‘tantamount to a licence to pollute’ (Price 2008), echoing the opinion that credit ought to be given to the success of the existing system in balancing competing interests when considering nuisance actions;

Both statutory nuisance controls and the common law doctrine of statutory authority offer some protections to the sort of operations in your list, the first in the form of the “best practicable means” defence, the second through a similar concept of “reasonable care”, and we should be grateful to know why you think it now imperative to over-ride that balance... (Price, 2008).

The Environmental Law Foundation (ELF), after consulting counsel, joined forces with the alliance in declaring the clause unlawful in its removal of the right to redress. A briefing, prepared in October by ELF, found that the proposed clause could;

restrict access to justice for disadvantaged communities affected by nuisance from major power stations, airports, or roads, as their local council could no longer bring such claims; severely restrict rights to claim compensation against nuisance; and allow for operators of major infrastructure projects to claim immunity on grounds of ‘statutory authority’ against prosecution, even when it can be proved they have been negligent (ELF, 2008).

The Bill entered the Committee stage, for debate in the House of Lords, on October 6th. Baroness Andrews began the discussion on clause 151, on October 16th, summing up the intention of the original insertion, which was to replicate the existing statutory authority exemptions for transport schemes. She then went on to acknowledge that the actual clause, as written, had far broader scope and application than that which was originally intended, agreeing with Lord Reay that the proposed clause “*tips the balance too far towards the promoter*”(Andrews, 2008, Hansard HC, Col.881). Reference was made to the fact that the measure would result in a situation where all nuisance proceedings would be barred, even those brought only ‘in relation to’ authorised developments. This

was seen as too great an advance and the Baroness suggested amendments to the clause, which, although still awarding immunity, would provide more appropriate redress for claimants in the form of compensation. On October 20th 2008, the contentious clause 151 was officially removed, and was replaced by this provision;

Nuisance: statutory authority

(1) This subsection confers statutory authority for—

- (a) carrying out development for which consent is granted by an order granting development consent;
- (b) doing anything else authorised by an order granting development consent.

(2) Statutory authority under subsection (1) is conferred only for the purpose of providing a defence in civil or criminal proceedings for nuisance.

(3) Subsections (1) and (2) are subject to any contrary provision made in any particular case by an order granting development consent. (Andrews, 2008, Hansard HC, Col.989).

This amended clause went on to become s.158 of the Planning Act 2008. It seems that, although the original broad-stroke application of immunity has been somewhat restricted in scope by the reassurances given in the House of Lords, the actual phrasing has changed little. In fact, the insertion of the pre-existing term ‘statutory authority’ lends further credence to a move which could result in the abrogation of rights for neighbours and communities, who may have to suffer long-lasting and unforeseen consequences of major developments. This paper will now explore the potential impact of the statutory authority defence on such actions, by reviewing the historical origins of the doctrine and the restrictions carved into its application by the judiciary in over two centuries of case law.

5: The widening of the statutory authority defence

This paper will now address the potential widening of the statutory authority defence in three key ways; through the historical reshaping of the doctrine, the move to implying authority for activities not specifically sanctioned by the legislature, and the augmentation of the list of projects, and their size limits, that has occurred in the Planning Act 2008.

5.1 Inception of the doctrine

The defence of statutory authority has evolved as a governmental response to the harmful consequences of activities condoned or mandated by the State. The legal principle behind it has been acknowledged as; *'cuicumque aliquis quid concedit, concedere videtur et id sine quo res ipsa esse*

*non potuit*⁸ (he who grants something is deemed to also grant that without which the grant would be worthless). In authorizing the activity, or allowing that something be built in a particular location, the legislature are seen to sanction any infringements which might have previously been considered to be unacceptable or unlawful (Brubaker 1995). Initially however, such an authorization was deemed to apply only to public officials, in the performance of a public duty (Linden 1996). This form of statutory authorization is evident in one of the first cases to consider the existence of the defence, *Entick v Carrington*⁹, which related to the unsolicited entry of government messengers onto private land. Although the judgment in *Entick*¹⁰ failed to find such authority, it did acknowledge that Parliament could authorise such interferences with property rights by clear prescription; “*if the Legislature be of that opinion they will make it lawful*”¹¹. This eighteenth-century judgment may seem trivial, especially in the modern context of larger scale industrial nuisances and pollution cases, but it indicates the origins of greater state-infringement on private property rights; something which was to continue over the following century.

The turnpike roads and harbours of the early 18th Century have been identified by Bassford (2007) as the first major infrastructure projects which required immunity from suits in nuisance, in order to install the requisite obstructions or methods of payment. However, infrastructure soon progressed from what was essentially “*a mean and muddy thoroughfare or a simple sea wall*” (Bassford 2007, 135). The dawning industrial era led to the development of major infrastructure provisions such as canals and railways, which would facilitate the burgeoning mills and factories. The most intrusive of these projects were the railway lines, with over 400 being authorized for construction between 1844 and 1846, with the accompanying legislation to ensure their completion (Brubaker 1995): “*At the very height of the reign of the “Railway King” George Hudson there were a vast number of railway bills before Parliament annually.*” (Bassford 2007, 136) These Bills and their resultant provisions included a host of measures which could interfere with personal property rights;

Land could be acquired, surveys carried out, trains run, canals built, navigation interfered with, charges levied, other conflicting legislation could be overridden, roads closed or diverted, criminal offences could be created and action taken if offences were committed. Until the 1940s, planning permission was not, of course, required. An Act was-- and in many cases remains --a panacea for every promoter's ill. (Bassford 2007, 136)

⁸ *Hammersmith & City Railway Co v Brand* (1865-66) L.R. 1 Q.B. 130, Lord Chelmsford at 202.

⁹ [1765] 95 E.R. 807.

¹⁰ *Ibid.*

¹¹ *Ibid.* at 819.

Such large scale interference with landowners' property rights resulted in an increase in actions for nuisances caused; including claims for beer ruined by vibrations from the passing trains¹², fright caused to horses by "*the great roaring, snorting, appalling monster vomiting smoke and fire in all directions*" (first steam locomotives)¹³, and increased premiums for a cotton mill adjacent to a sparking railway line.¹⁴ It is in this abundance of resultant case law that the general principles surrounding statutory authority have been formulated and shaped.

5.2 Widening of the application

The existence of statutory authority provided ample opportunity for developers of these new railway schemes to escape liability for causing any nuisance to citizens with interests in proximate land. This arose as a result of a widening in the statutory exemption considered in *Entick*¹⁵, which had specifically referred to explicit parliamentary authority for the actions of public officials, who were performing a public duty. The courts appeared more willing over the subsequent years to consider the defence applicable in any situations where individuals were seen to be acting under sanction of the legislature, including private commercial enterprises that incorporated an amount of public benefit into their activities (Linden 1966). It seems that the proliferation of railway-related case law in the mid-1800s may have been responsible for interweaving the realms of common law and statutory interpretation, thus incorporating the statutory authority immunity into such actions (Wilde & Smith, 2010).

This widening continued however, with a move to judicially implied authority, as demonstrated in *Hammersmith Railway v Brand*¹⁶. In such instances the application of the defence was widened beyond express statutory phrasing or insertion, and deemed to also occur by 'necessary implication'. This case appears to establish a low threshold for the interpretation of the statutory authority exemption. Such 'back door' authorisation has been highly contentious, with judges having to interpret statutory authority in situations where the effective use or purpose of certain permitted projects would have been impossible otherwise (Pontin 2002).

¹² *The London and North-Western Railway Company v Bradley* (1851) 42 E.R. 290.

¹³ *Rex v. Pease*, quoted in Wilde & Smith pg. 16.

¹⁴ *Stockport, Timperley and Altringham Railway Company* (1864) 33 L.J. 251.

¹⁵ (1765) 19 St. Tr. 1030.

¹⁶ (1869) LR 4 AC HL.

Cases such as *Manchester Corporation v Farnworth*¹⁷ and *Allen v Gulf Oil Refining*¹⁸ demonstrate the implication of such an authority. In these cases the relevant Acts were silent or ambiguous as to certain operational procedures pertaining to a power station and an oil refinery respectively. The judges in both instances overruled the private rights of neighbours, who had brought reasonable actions in nuisance, by the necessary implication of statutory authority. This form of equivocal judicial interpretation has led to some criticism;

A substantial minority of judges have argued that casting courts into the role of ‘soothsayer’ of some potentially mythical Parliamentary intent in this way is an invitation to arrive at radically divergent and unpredictable outcomes. Denning colourfully likened this task to ‘a guesswork puzzle’. (Pontin 2002, pg.351)

It also led to conflicting and, in certain instances, anomalous results as the courts were not inclined to favour the defendants in every case, as evident in *Rapier v London Tramways*¹⁹, a case which pertained to the authorised operation of horse-drawn trams. The statutory authorisation for the tramlines was silent on the matter of where, or how, the requisite horses might be housed but the defendant had built stables that could accommodate approximately 400 horses, which were a severe disturbance to the local residents. The court in *Rapier*²⁰ however, did not find sufficient authorisation in the actual statute for such a nuisance and did not feel compelled to imply it in that instance. They determined that the stables, although necessary for the effective operation of the tram, should have been conducted and managed in a way that did not cause a nuisance to the neighbouring community.

Parliament has, until now at least, been quite reluctant to show a clear and explicit intention to remove citizen’s private rights, and express statutory authority has appeared in very few legislative provisions. Parliament has, it seems, tried instead to abrogate such rights in a more oblique manner due to the political sensitivity of the issue (Pontin 2002). Where immunity has been awarded it has generally only been successfully received when juxtaposed with some form of statutory remedy, as in s.44 of the Land Compensation Act 1973 and, in the current instance, s.152(3) Planning Act 2008, which states that; “*A person by whom or on whose behalf any authorised works are carried out must pay compensation to any person whose land is injuriously affected by the carrying out of the works*”. This mirrors the common law position on the removal of private rights as, where there was a

¹⁷ [1930] AC 171.

¹⁸ [1981] AC 1001.

¹⁹ [1893] 2 Ch. 588.

²⁰ Ibid.

statutory compensation scheme available, it acted as an indication to the courts that any nuisance caused was an inevitable result of the action.²¹

5.3 Application under s.158 Planning Act 2008

Opponents to the original clause 151 continue to raise concerns that there will still be a high possibility that developers will escape liability for nuisances under the new provisions, referring to it as ‘*de facto immunity from nuisance*’ (EPUK, 2008). This is particularly relevant in consideration of the list of projects with such exemption. This list, which appears in s.14 of the Planning Act 2008, contains a considerable amount of structures and developments which could be treated as ‘Nationally Significant Infrastructure Projects’:

- (a) the construction or extension of a generating station;
- (b) the installation of an electric line above ground;
- (c) development relating to underground gas storage facilities;
- (d) the construction or alteration of an LNG facility;
- (e) the construction or alteration of a gas reception facility;
- (f) the construction of a pipe-line by a gas transporter;
- (g) the construction of a pipe-line other than by a gas transporter;
- (h) highway-related development;
- (i) airport-related development;
- (j) the construction or alteration of harbour facilities;
- (k) the construction or alteration of a railway;
- (l) the construction or alteration of a rail freight interchange;
- (m) the construction or alteration of a dam or reservoir;
- (n) development relating to the transfer of water resources;
- (o) the construction or alteration of a waste water treatment plant;
- (p) the construction or alteration of a hazardous waste facility.

Although statutory authority is an established defence against private action, it has previously been used in a more measured and restricted way in relation to very specific projects such as railways and utilities. This large new list of projects appears to extend the application of the defence to a wider variety of developments and their ancillary projects.

Concern has also been raised as the projects in the list have no stringent size limits, which could lead developers to ‘think big’, submitting applications for projects larger than strictly necessary in order

²¹ See *Marriage v East Norfolk Rivers Catchment Board* [1950] 1 K.B. 284.

to benefit from the statutory authority defence (Bassford 2007). This point was noted in the House of Lords committee stage of the Planning Bill:

... the threshold for onshore wind projects is set at 50 megawatts, which means 15 megawatts of annual production of electricity generated and suggests that they are only one-hundredth the size of a normal nuclear power station. They cannot be considered to be big projects. (Lord Reay 2008, Col. 886)

It would seem that the proposed application of the statutory authority defence for infrastructure projects is a far remove from the immunity propounded in *Entick*²², of protection for public officials performing activities of benefit to the public concerned. What appears to be emerging as more common is the protection of private commercial companies, which are engaged in works sanctioned by the State, and an acceptance of any resulting infringements caused by their activities. This, it could be argued, is a substantial stretch of the public benefit argument as, while it can be acknowledged that “*salus populi suprema lex*”²³ (the welfare of the people is the supreme law), it is not always clear who the real beneficiaries are in such circumstances. Mc Laren (1983) states that the law strongly reflects changes of perception within society and what importance they have, emphasising that this is particularly relevant ‘*in times of dramatic social change and economic growth*’ (and, it could be argued, recession). In the current economic climate it may be that the prospect of employment growth, investment in national infrastructure and improved transport links, which encourage trade and tourism, are the incentives being offered to the new generation of disenfranchised neighbours. We are not facing a new industrial revolution, in fact quite the reverse, with industry and manufacturing on the decline in the UK year-on-year but the same justifications are being given for infringements on private property interests (Marvin & Guy 2001). It will therefore fall to the judiciary to continue to apply the restrictions carved out in common law to the statutory authority defence in order to protect private property owners’ rights.

6: The scope and application of statutory authority

The range of activities excused under the statutory authority defence has been exceedingly broad and, at times obscure, covering nuisances such as; cattle noise from a railway yard,²⁴ the emission of unpleasant smells from sewage works,²⁵ and the operation of a smallpox hospital in a residential

²² (1765) 19 St. Tr. 1030.

²³ *British Cast Plate v Meredith* (1792) 100 E.R. 1306 at 1309.

²⁴ *Brighton and South Coast Railway v Truman* (1886) L.R. 11 App. Cas. 45.

²⁵ *Hounslow LBC v Thames Water Utilities Ltd* [2003] EWHC 1197 (Admin).

area.²⁶ However, the actual application of the defence, and its relationship to statutory nuisance, has been restricted to allow for more fair and reasonable results for claimants. These restrictions have been applied through the years in a number of proceedings, shaping the scope of statutory authority in relation to; the proposed location of the nuisance, the inevitability of the nuisance, the effect on statutory provisions and the bearing that negligence has on the defence.

6.1: The importance of location specificity

Obviously in s.158 of the Planning Act 2008, Parliament have made their intention to restrict private rights apparent, meaning that in future developments there should be an amount of legal certainty. However the NPS process in itself could affect the application of the exemption, as courts have been reluctant to proffer the authority where there is a wide discretion available in the choice of sites for the proposed development (Buckley 1981). Unless National Policy Statements were location-specific in every instance, this aspect of implied authority could become problematic. This proved to be the case in *Metropolitan Asylum District Managers v Hill (No.2)*²⁷; although the legislature had empowered the Poor Law Board to designate an asylum in the London area, the statutory framework did not explicitly specify the building or operation of the hospital. The courts could not determine whether the operation of the hospital would cause any nuisance irrespective of its location in London and, therefore, did not find sufficient explicit authority to abrogate common law rights. Lord Watson clarified the scope of the immunity:

I do not think that the Legislature can be held to have sanctioned that which is a nuisance at common law, except in the case where it has authorized a certain use of a specific building in a specified position, which cannot be so used without occasioning nuisance, or in the case where the particular plan or locality not being prescribed, it has imperatively directed that a building shall be provided within a certain area and so used, it being an obvious or established fact that nuisance must be the result.²⁸

The clarity of expression in the 2008 Act should make it easier for judges to identify that such statutory authority exists, however the actual application in each situation could depend on the wording of the National Policy Statements and the amount of consideration given to the location of the proposed development.

²⁶ *A.G. v. Corporation of Nottingham* [1904] 1 Ch. 673.

²⁷ (1881) L.R. 6 App. Cas. 193.

²⁸ *Metropolitan Asylum District Managers v Hill (No.2)* (1881) L.R. 6 App. Cas. 193 at 212.

6.2: Is the nuisance caused an ‘inevitable result’ of the development?

One of these fundamental rules, established in *Manchester v Farnworth*²⁹, is that the nuisance caused will only be excused if it is an inevitable result of the defendant acting within the scope of their Parliamentary authority (Penner 2000). In that case it was also viewed that inevitability would apply where the defendant had tried the best system of nuisance prevention available at the time, allowing for up-to-date knowledge of such systems and taking into account the cost implications.³⁰ The defence in *Manchester*³¹ failed because the damage caused was not an inevitable result of the building of the power station in question; it should instead have been built with taller chimneys, which would have dispersed the emissions more efficiently (Wilde 2008). This position was upheld more recently in the *Allen*³² case, with Lord Wilberforce describing the procedure for courts when considering the defence in some detail:

The respondent alleges a nuisance, by smell, noise, vibration, etc. The facts regarding these matters are for her to prove. It is then for the appellants to show, if they can, that it was impossible to construct and operate a refinery upon the site, conforming with Parliament's intention, without creating the nuisance alleged, or at least a nuisance.³³

This obviously has some bearing on the treatment of future cases relating to Nationally Significant Infrastructure Projects and developers will need to ensure that they do not cause or exacerbate any avoidable nuisances. As Viscount Sumner stated in the *Manchester*³⁴ case: “*The legislature, in authorising a very big concern, must not be taken to be chary of requiring that much care is to be taken*”³⁵. The concerns voiced by opponents to the Act that, under the new provisions, developers would have no need to mitigate their effects on local communities may thus be addressed. The defence will only be available as long as any nuisance they cause is an inevitable result of their actions and they have used the most modern means available in order to try to abate that nuisance. In effect, this replicates the existing defence of ‘best practicable means’, which is available under s.80(7) of the Environmental Protection Act 1990, for defendants who face prosecution under statutory nuisance provisions (Samuels 2004).

²⁹ [1930] AC 171.

³⁰ *Manchester Corporation v Farnworth* [1930] AC 171 at 183.

³¹ *Ibid.*

³² *Allen v Gulf Oil Refining Ltd* [1981] AC 1001.

³³ *Ibid.* at 1013.

³⁴ *Manchester Corporation v Farnworth* [1930] AC 171

³⁵ *Ibid.* at 201.

6.3: Effect on statutory nuisance

Opponents to the initial Planning Bill raised concerns that a total immunity approach could mean that developers would not be held responsible for any breaches of statutory nuisance legislation and would elude potential enforcement proceedings. However the actual wording of the Planning Act appears to remove such concerns, mirroring as it does the previous provisions in the Railways Act 1993. The defence as set out in s.122(3) of the Railways Act 1993 provides exemption; “(i) in any proceedings, whether civil or criminal, in nuisance; or (ii) in any civil proceedings”, whereas statutory authority in s.158(2) of the Planning Act 2008; “is conferred only for the purpose of providing a defence in civil or criminal proceedings for nuisance”.

Immunity from statutory nuisance proceedings under s.122(3) was considered in *Camden LBC v London Underground Ltd*³⁶, with the court determining that statutory authority could only act as a defence to common law nuisances, and would not create a defence in actions prescribed by environmental protection legislation. This means therefore, that immunity will be afforded to any interference with land, as protected by common law nuisance, but will not remove liability for actions which may be prejudicial to health.

Although the reference to statutory nuisance in *Camden*³⁷ was *obiter*, the guidance remains that statutory authority will not be considered where activities have resulted in effects that are injurious to the health of others;

...our provisional view is that the immunity conferred by s.122 applies to common law nuisance but not to all statutory nuisances under the 1990 Act. All the categories of statutory nuisance listed in s.79(1) are defined by reference to the matter complained of being “prejudicial to health” or “nuisance”. In the absence of any contrary expressed intention, nuisance must mean common law nuisance, which is a tort committed against land. There is nothing in section 122 to suggest that the statutory defence was being conferred in relation to all statutory nuisances, which include injury to health which might not be related to land.³⁸

Such a position would appear to uphold the sanctity of environmental protection legislation, which should protect neighbours from certain nuisances, and reinforces the reassurances voiced by Baroness Andrews in the House of Lords Committee stage debate:

³⁶ [2000] Env. L.R. 369

³⁷ Ibid.

³⁸ Ibid. LJ Rose at 379.

The default position set out in our amendment does not confer a defence against any of the statutory nuisances set out in Section 79(1) of the Environmental Protection Act 1990, in so far as they involve matters outside what would constitute nuisance in common law. (Baroness Andrews 2008, Col. 882)

6.4: A negligent act

There is a long case history clearly demonstrating that there will be no defence available where a nuisance has been caused due to the tortfeasor's negligence. The court in *Manchester v Farnworth*³⁹ reaffirmed this position, citing cases such as; *Rex v. Pease*⁴⁰; *Vaughan v. Taff Vale Railway Co.*⁴¹; *Hammersmith Railway v Brand*⁴² and *Geddis v. Bann Reservoir*⁴³: "*If the Legislature has authorized the doing of a particular thing in a particular place it authorizes the doing of it, even though it causes a nuisance, provided that it is not done negligently.*"⁴⁴

The definition of the term 'negligence', as it applies to statutory authority exemption, was propounded in the *Allen*⁴⁵ case, with Lord Wilberforce stating that the word be used; "... *in a special sense so as to require the undertaker, as a condition of obtaining immunity from action, to carry out the work and conduct the operation with all reasonable regard and care for the interests of other persons.*"⁴⁶ More recently, negligence has been considered in relation to statutory authority immunity in the *Marcic*⁴⁷ case, with the House of Lords again upholding the need for operators to perform their functions within the scope of the authority proffered upon them, in a way that is not negligent or too remote. This approach is in accordance with both previous case law and European jurisprudence, as evidenced in *Kane v New Forest*⁴⁸:

Nor on the facts of this case are the defendants immune from a claim in negligence because they were exercising a statutory function under planning legislation... I reject Mr Snells's submission that a planning authority has blanket immunity from claims for negligence whatever the facts. That is simply not consonant with recent

³⁹ [1930] AC 171.

⁴⁰ (1832) 4 B. & Ad. 30.

⁴¹ (1860) 5 H. & N. 679.

⁴² (1869) LR 4 AC HL.

⁴³ (1878) 3 App. Cas. 430.

⁴⁴ Viscount Dunedin, *Manchester Corporation v Farnworth* [1930] AC 171 at 174.

⁴⁵ *Allen v Gulf Oil Refining Ltd* [1981] AC 1001.

⁴⁶ *Ibid.* at 1011.

⁴⁷ *Marcic v Thames Water Utilities Ltd* [2003] UKHL 66.

⁴⁸ [2001] 3 All E.R. 914.

developments of the law both in this jurisdiction and in Strasbourg see for example *Barrett v Enfield London Borough Council* [2001] 2 AC 550 and *Osman v United Kingdom* (1998) 29 EHRR 245.⁴⁹

All of the restrictions outlined above could have a big impact on developers or operators of major infrastructure projects, who will not have a “*carte blanche*”⁵⁰ to create or exacerbate nuisances. Instead they will have to show that any nuisance caused is an inevitable result of the permitted development, as well as having to perform their duties in consideration of the relevant statutory provisions, whilst ensuring that all activities are conducted in a manner that cannot be classified as ‘negligent’. Additionally any NPS pertaining to such developments may have to be very location-specific, in order for the legislative authorisation to be explicit. These exceptions have been carved out in a demonstration of judicial hostility towards the notion of any infringement on private rights, sanctioned or otherwise (Linden 1966). It would seem that many of the opponents to the Planning Act 2008 could find solace in the safeguards these qualifications provide for neighbours of new infrastructure developments.

7: Conclusion

This paper demonstrates that the planning system for major infrastructure projects has been radically reshaped. Among a host of more contentious measures such as the curtailing of consultation, the establishment of an unelected planning commission, and the removal of certain hearing rights, the insertion of the existent ‘statutory authority’ defence appears unremarkable. This paper argues however, that the provisions set out in s. 158 of the Planning Act 2008, while embracing the terminology of the existing legal doctrine of statutory authority, may extend the scope and significance of such a principle and could tilt the balance of conflicting property rights in the developer’s favour.

The previous explicit authorization, in clause 151 of the Planning Bill, to infringe private rights and award complete immunity from actions in private or public nuisance for developers of NSIPs, indicated a governmental disregard for private property owners. This paper has demonstrated the progression from this ‘blanket immunity’ to the statutory authority defence proffered in s.158 of the Planning Act 2008. It has also been shown that such a defence will be awarded restrictively in the

⁴⁹ Ibid. at 921.

⁵⁰ *Midwood v. Manchester Corporation* [1905] 2 K.B. 597 at 606.

court rooms; the judiciary will retain a key role in determining both the scope of statutory authority in such instances and whether the defence should be applied at all. In producing National Planning Statements, government ministers will need to ensure that they are presenting statements that pertain to specific locations wherever possible, or the defence may not be upheld. Additionally, developers will have to heed precedents spanning two centuries pertaining to the application of the statutory authority defence and ensure that they behave in a non-negligent fashion, that any nuisance caused is an inevitable consequence of their sanctioned actions, and abide by the relevant statutory nuisance provisions which pertain to their activities.

Acknowledgements

I would like to thank my PhD supervisor, Mr. Paul Chynoweth, for his unwavering support, encouragement and advice over the past year; this has been invaluable. I would also like to give thanks to the academic and support staff at the University of Salford, whose enthusiastic and friendly attitude makes every problem seem surmountable.

References

- *Archbold's Criminal Pleading, Evidence and Practice*, 2010, London: Sweet & Maxwell.
- Baroness Andrews, 2008, Hansard HL, Col.881, 882 & 989 (16th October 2008). Available from: <http://www.publications.parliament.uk/pa/ld200708/ldhansrd/text/81016-0010.htm> [accessed 22nd April 2010].
- Bassford H., 2007 'Major infrastructure projects: the answer is the Planning White paper. Now, what was the question?', *Journal of Planning & Environment Law*, Dec Supp (Planning: Still Sexy at Sixty?), 134-158.
- Betts C., Hansard HC, Col.371, (25th June 2008). Available from: <http://www.publications.parliament.uk/pa/cm200708/cmhansrd/cm080625/debtext/80625-0017.htm> [accessed 16th April 2010].
- Blackstone's Commentaries, Volume III (1st Edition, 1768).
- Blears H., Hansard HC, Col.342, (25th June 2008) Available from: <http://www.publications.parliament.uk/pa/cm200708/cmhansrd/cm080625/debtext/80625-0012.htm> [accessed 14th July 2010].
- Brenner J., 1973, 'Nuisance law and the Industrial Revolution', *Journal of Legal Studies*, 3, 403-433.
- Brubaker E., 1995, *Property Rights in the Defence of Nature*, Toronto: Earthscan Publications.
- Buckley R.A., 1981, *The Law of Nuisance*, London: Butterworths.
- Cavendish C., 2008, 'The Planning Bill: a bulldozer through democracy', *The Times*, 26th June. Available from: http://www.timesonline.co.uk/tol/comment/columnists/camilla_cavendish/article4214738.ece [accessed 20th July 2010].

- Coyle S. & Morrow K., 2004, *'The Philosophical Foundations of Environmental Law: Property, Rights and Nature'*, Oregon: Hart.
- Ellis, H., 2008, 'Planning and the people problem', *Journal of Planning & Environment Law*, 13 Supp (Delivery or Democracy?), 75-86.
- Communities & Local Government etc., 2007, *White Paper on Planning: Planning for a Sustainable Future*, HM Government, London.
- Elvin D. & Karas J., 2002, *'Unlawful Interference With Land'*, London: Sweet & Maxwell.
- Environmental Protection UK, 2008, Environmental Protection UK Briefing – Planning Bill and Nuisance (August). Available from: http://www.environmental-protection.org.uk/assets/library/documents/Planning_Bill_Briefing_Clause_151_Nuisance_Immunity.pdf [accessed 10th August 2010].
- Environmental Protection UK, 2008, Environmental Protection UK Briefing – Planning Bill and Nuisance (November). Available from: http://www.environmental-protection.org.uk/assets/library/documents/Planning_Bill_Nuisance_Immunity_Briefing.pdf [accessed 14th August 2010].
- Giliker P. & Beckwith S., 2008, *'Tort Textbook'*, London: Sweet & Maxwell.
- Guy S. & Marvin S., 2001, 'Reconnecting Networks and Buildings: The Development Process and the Reshaping of Water, Energy and Transport Demands', in Miller C. ed., *Planning and Environmental Protection: A Review of Law and Policy*, Oregon: Hart.
- Hope C., 2008, 'Hundreds of thousands will not be able to complain about noisy building sites' *Telegraph*, 13th August. Available from: <http://www.telegraph.co.uk/news/2553345/Hundreds-of-thousands-will-not-be-able-to-complain-about-noisy-building-sites.html> [accessed 14th July 2010].
- Linden A., 1966, 'Strict Liability, Nuisance and Legislative Authorization', *Osgoode Hall Law Journal*, Vol. 4 196-221.
- McLaren J., 1983, 'Nuisance Law and the Industrial Revolution – Some Lessons from Social History', *Oxford Journal of legal Studies*, 3(2), 155-221.
- Miner P., 2008, 'Unprecedented Alliance Forms Against Planning Bill Licence To Pollute' *ePolitix.com*, 14th October, Available from: <http://www.epolitix.com/stakeholder-websites/press-releases/press-release-details/newsarticle/unprecedented-alliance-forms-against-planning-bill-licence-to-pollute///sites/campaign-to-protect-rural-england/> [accessed 6th July 2010].

- Office of the Prime Minister, Queen's Speech summary. Available from: <http://www.number10.gov.uk/queens-speech/2010/05/queens-speech-decentralisation-and-localism-bill-50673> [accessed 29th June 2010].
- Penner J.E., 2000, 'Nuisance, the Morality of Neighbourliness, and Environmental Protection', in Lowry J. and Edmunds R. eds., *Environmental Protection and the Common Law*, Oregon: Hart.
- Pontin B., 2002 'Environmental rights, statutory authorisation and the Human Rights Act: is clarity possible?', *Environmental Law & Management*, 14(6) 350-356.
- Price H., 2008, 'CIEH lobbies Minister to prevent Planning Bill blunder', *ePolitix.com*, 15th August, Available from: <http://www.epolitix.com/stakeholder-websites/press-releases/press-release-details/newsarticle/cieh-lobbies-minister-to-prevent-planning-bill-blunder///sites/chartered-institute-of-environmental-health/> [accessed 23rd July 2009].
- Lord Reay, Hansard HL, 2008, Col.886, (16th October 2008). Available from: <http://www.publications.parliament.uk/pa/ld200708/ldhansrd/text/81016-0010.htm> [accessed 22nd April 2010].
- Samuels A., 2004, 'Developments that cause a nuisance: the legal significance of the grant of planning permission', *Journal of Planning & Environment Law*, Apr, 394-397.
- Spencer J., 1989, 'Public Nuisance – A Critical Examination', *Cambridge Law Journal*, 48(1), March, 55-84.
- Thompson P., 2009, 'Consultation and the authorisation of major infrastructure projects', *Journal of Planning & Environment Law*, 2, 174-189.
- Wilde M. & Smith C., 2010, 'R. v. Pease (1832)', in Mitchell C. and Mitchell P. eds., *Landmark Cases in the Law of Tort*, Oregon: Hart.
- Wilde M., 2008, 'Best Available Techniques (BAT) and coal-fired power stations: can the energy gap be plugged without increasing emissions?', *Journal of Environmental Law*, 20(1), 87-114.

A Principled Response to Pre-Contractual Remuneration

Kelvin Koh

Faculty of Law,
National University of Singapore

Email: kelvin_kohl1@hotmail.com

Abstract:

Under English law, there are uncertainties as to the circumstances in which a contractor can recover remuneration for any work done in anticipation of a contract that fails to materialize. The reluctance to adopt the Continental duty of good faith in the negotiation process has led to piecemeal responses to pre-contractual problems. Following *Yeoman's Row Management v Cobbe*, there seems to be a judicial gravitation towards unjust enrichment as the basis of liability. However, unjust enrichment as the sole justification cannot adequately address the rationale of imposing pre-contractual liability. This paper argues that the rationale is to remedy two types of harm: (i) the detrimental reliance suffered by the claimant and (ii) the wrongful transfer of value from the claimant to the defendant. Thereafter, the concept of risk will be explored and bifurcated into factual risk and normative risk to help us achieve a better understanding of the cases. This paper suggests that a model, consisting of promissory estoppel and unjust enrichment, will provide a better fit with the rationale. It will be shown that normative risk and fault can be assimilated into the two causes of action to form a composite picture of pre-contractual remedies. The discussion will culminate in a neat framework for analyzing anticipated contract scenarios that will hopefully bring consistency and order to the troubled realm of anticipated contracts.

Keywords:

Pre-contractual liability, Anticipated contracts, Unjust enrichment, Risk taking.

1 Introduction

An established developer decides to invest in a multi-billion dollar project and enters into negotiations with a contractor. Due to the complexities of the project and the fluctuating costs of building materials, the parties are unable to draw up a complete agreement on the costs and several important terms in the meantime. The developer wants to complete its mega project within four years and instructs the contractor to commence with the construction works. Subsequently, the working relationship between the parties turns sour and the contractor is anxious to recover his expenses on the project. Will he be able to do so? If the developer refuses to pay, what legal cause of action should the contractor take?

The courts have responded to such problems with piecemeal solutions such as the collateral contract theory¹, the law of restitution², promissory estoppel³, risk

¹ *Way v Latilla* [1937] 3 All ER 759 (HL).

analysis⁴ and the *quantum meruit* claim⁵. Although the different approaches may usually achieve a similar outcome for the parties, clear classification of the claim under a general principle is desirable for the “pursuit of rationality, consistency and order”.⁶ Nevertheless, in our attempt to achieve a consistent picture of pre-contractual claims, we should be careful not to conflate the different causes of actions without due care to their underlying basis. The application of different principles can sometimes lead to differing measures of damages and affect the availability of defences, and this could result in disparate outcomes.

The purpose of this paper is to present a principled response to pre-contractual problems. It begins by proposing a rationale based on the harm theory. Two types of harm pertinent to the pre-contractual context are identified, viz. the detrimental reliance suffered by the claimant and the wrongful transfer of value from the claimant to the defendant. Thereafter, the concept of risk will be explored and bifurcated into factual risk and normative risk to help us achieve a better understanding of the cases. This paper suggests that a model, consisting of promissory estoppel and unjust enrichment, will provide a better fit with the rationale. It will be shown that normative risk and fault can be assimilated into the two causes of action to form a composite picture of pre-contractual remedies. The discussion will culminate in a neat framework – constructed upon a gradation scale of risk and fault – for analyzing anticipated contract scenarios that will hopefully bring consistency and order to the troubled realm of anticipated contracts.

2 Unjust Enrichment as the Basis of Liability?

There has been widespread academic debate on the bases of liability for pre-contractual remuneration. Following *Yeoman’s Row Management v Cobbe*⁷, there seems to be a judicial gravitation towards unjust enrichment. There, the parties made an agreement-in-principle regarding the development of a parcel of land, however, negotiations fell through subsequently and the claimant sued to enforce the agreement. The House of Lords found that the agreement was merely “binding in honour”, and also rejected the proprietary estoppel argument for want of certainty. Although the claimant failed to

² *British Steel Corp v Cleveland Bridge and Engineering Ltd* (1981) [1984] 1 All ER 504 (Com Ct) [*British Steel*]; *Vedatech Corp v Crystal Decisions (UK) Ltd* [2002] EWHC 818; *Yeoman’s Row Management Ltd v Cobbe* [2008] 1 WLR 1752 (HL) [*Yeoman’s Row*]; *Whittle Movers Ltd v Hollywood Express Ltd* [2009] EWCA Civ 1189 [*Whittle Movers*].

³ *Walton Stores (Interstate) Ltd v Maher* (1988) 164 CLR 387 (High Court of Australia) [*Walton Stores*]; *Commonwealth of Australia v Verwayen* (1990) 170 CLR 394 (High Court of Australia) [*Verwayen*].

⁴ *Brewer Street Investments Ltd v Barclays Woolen Co Ltd* [1954] 1 QB 428 (CA) [*Brewer Street*]; *Jennings & Chapman Ltd v Matthews and Co.* [1952] 2 TLR 409 (CA) [*Jennings*].

⁵ *Sintalaw Engineering Pte Ltd v Nashimatsu Lum Chang JV* [1996] SGHC 274 (High Court of Singapore); *Vicson General Contractors Pte Ltd v Dai-Dun Co Ltd* [1997] SGHC 345 (High Court of Singapore); *Yeoman’s Row*, *supra* note 2. Although the term *quantum meruit* is frequently used in textbooks (See for example, D Wallace, *Hudson’s Building and Engineering Contracts, Vol 1* (11th edn Sweet & Maxwell, London 1995); S Furst and V Ramsey, *Keating On Construction Contracts* (8th edn Sweet & Maxwell, London 2006) 110), the legal basis of such claims is unclear (See *British Steel*, *supra* note 2, 509 (Goff J); J Bailey, ‘Repudiation, Termination and Quantum Meruit’ (2006) 22(4) Const LJ 217, 235), and one author has called for the term to be eradicated to clear grounds for a more lucid exposition of the liability model underpinning remuneration award (K Barker, ‘Coping with Failure – Reappraising Pre-Contractual Remuneration’ (2003) 19 JCL 105, 128. [*Barker*]).

⁶ E McKendrick, ‘Work Done in Anticipation of a Contract which does not Materialise’ in W.R. Cornish et al. (eds), *Restitution Past, Present and Future: Essays in Honour of Gareth Jones* (Hart Publishing, Oxford 1998), 163.

⁷ *Supra* note 2.

enforce the agreement, recovery of his pre-contractual remuneration was granted on the grounds of unjust enrichment.⁸ Most recently, in *Whittle Movers*,⁹ the Court of Appeal refused to imply a contract for the services performed in anticipation of a contract, and stated that any recovery must be effected under a restitutionary claim.

However, acceptance of unjust enrichment as the sole basis of liability is not without problems since it is a gain-based cause of action, which presupposes that the defendant has been enriched by the conduct of the claimant. An unjust enrichment analysis may hence be inadequate to justify recoveries in cases where the defendant does not receive any benefits. In *Brewer St*,¹⁰ the defendant, in anticipation of leasing the claimant's property, requested the claimant to perform some work on the premises. The lease failed to materialize and the defendant was unable to benefit from the work done. Nevertheless, the Court of Appeal allowed the claimant to recover for the value of the work done. The defendant was surely not enriched here, unless one interprets enrichment broadly to include any work requested for, even if the defendant did not receive the benefit of the work done. However, this might be seen as a distortion of the natural meaning of enrichment, so as to protect the claimant's reliance loss.¹¹ The objective of unjust enrichment is to reverse a defective transfer of value and it cannot apply where no actual transfer of value occurs. Reliance loss and wrongful transfer of value are two separate harms and cannot be conflated.

3 The Theoretical Bases for Pre-Contractual Liability

The issue of pre-contractual liability is often characterized by the question of which legal principles best form the bases of liability, and misses the more fundamental question of why pre-contractual liability should be imposed. It is beyond the scope of this paper to prescribe the perfect theoretical model for pre-contractual liability. Instead, this paper makes a humbler suggestion that the harm theory offers appropriate answers to the question of why pre-contractual liability should be imposed. In particular, the harm theory provides an explanation as to why there is a divergence of opinions with regards to the best liability model and shows that there is no single legal framework that can encompass all pre-contractual cases.

3.1 The Reluctance to Impose Pre-Contractual Liability

The rights of parties in a negotiation are shaped by two notions of freedom:

“the positive freedom of contract, which means that the parties are free to create a binding contract reflecting their free will, and the negative freedom of contract, which means that the parties are free from obligations so long as a binding contract has not been concluded”.¹²

⁸ It should be noted that Lord Scott appeared to be of the view that unjust enrichment and quantum meruit were separate common law remedies but produced co-extensive results. See [40] – [42].

⁹ *Supra* note 2.

¹⁰ *Supra* note 4.

¹¹ *Countrywide Communications Ltd v ICL Pathway Ltd* [2000] CLC 324 (QB) 349 [*Countrywide*].

¹² N Cohen, ‘Pre-Contractual Duties: Two Freedoms and the Contract to Negotiate’ in J Beatson and D Friedmann (eds), *Good Faith and Fault in Contract Law* (Clarendon Press, Oxford 1995), 25.

The premium placed upon freedom of contract under English law has resulted in its reluctance to embrace the Continental idea of duty of good faith in the bargaining process. The adoption of such a duty would also be “inherently repugnant” to the assumption at English law that parties take “adversarial positions” when involved in negotiations.¹³ This “aleatory view” of negotiations stands as the greatest obstacle to the imposition of pre-contractual liability: “a party that enters negotiations in the hope of the gain that will result from ultimate agreement bears the risk of whatever loss results if the other party breaks off the negotiations”.¹⁴ As Barry J stated in *William Lacey (Hounslow) Ltd v Davis*:¹⁵

“[The claimant] undertakes this work as a gamble, and its cost is part of the overhead expenses of his business which he hopes will be met out of the profits of such contracts as are made...”

One of the main concerns of imposing pre-contractual liability has been that it would have a “chilling” effect on parties.¹⁶ As the courts seek to protect the claimants from their pre-contractual losses, the courts have to respect the defendants’ autonomies. Otherwise, parties may adopt an overly defensive stance when entering into negotiations and stifle the spirit of cooperation that is desired. Hence, strong justifications must be provided for curtailing the defendants’ autonomies.

3.2 Harm Theory

The harm theory justifies the imposition of pre-contractual liability from a moral perspective. One can trace the roots of freedom of contract to the liberty of an individual, but liberty in itself has never been recognized as an unqualified right. In *On Liberty*, Mill theorises that:

“the only purpose for which power can be rightfully exercised over any member of a civilized community, against his will, is to prevent harm to others”.¹⁷

Therefore, each individual has the liberty to act as he wishes, so long as these actions do not harm others. In the anticipated contracts context, the defendant is free to withdraw from the negotiations unless in so doing, he would cause harm to the claimant. This paper suggests that such harm is manifested in two forms – the reliance loss suffered by the claimant and wrongful transfer of value from the claimant to the defendant. While these two types of harm have a degree of overlap, they are separate and distinct, and are thus addressed by two different normative theories: (1) the duty to ensure reliability of induced assumptions and (2) corrective justice.

3.2.1 Duty to Ensure Reliability of Induced Assumptions

A duty to ensure the reliability of induced assumptions can be derived from the general duty not to cause harm. Spence defines the duty as one which:

¹³ *Walford v Miles* [1992] 2 AC 128 (HL) 138.

¹⁴ E Farnsworth, ‘Precontractual Liability and Preliminary Agreements: Fair Dealing and Failed Negotiations’ (1987) 87 Col L Rev 218, 221.

¹⁵ [1957] 1 WLR 932 (QB) 934.

¹⁶ E Farnsworth, *Contracts* (4th edn Aspen Publishers, New York 2004), [3.26].

¹⁷ J S Mill, *On Liberty* (Yale University Press, New Haven 2003) at 80.

“... places primary and secondary obligations on a party who (i) induces an assumption in the mind of another party and (ii) induces the other party to rely upon that assumption. The primary obligation is that the inducing party must, in so far as he is reasonably able, prevent harm to the relying party. ‘Harm’ consists in the extent to which the relying party is worse off because the assumption has proved unjustified than he would have been had it never been induced. The secondary obligation is that, if the relying party does suffer harm of the relevant type, and the inducing party might reasonably have prevented it, then the inducing party must compensate the relying party for the harm he has suffered.”¹⁸

Spence’s formulation of the duty may favour the claimant unduly and a qualifier, that (iii) the assumption must be reasonable in the particular context, should be added. The reformulated duty and Spence’s definition of ‘harm’ thus narrow down Mill’s broad harm principle and demarcates clear boundaries that a substantive rule can operate within.

This duty also strikes a fair balance between a defendant’s negative freedom of contract and a claimant’s reliance interest. Where the defendant engages in negotiations and takes care not to induce any assumptions in the mind of the claimant, the defendant has the freedom to withdraw from negotiations any time prior to the formation of the contract. Even where he has induced an assumption, the defendant may still withdraw if he does so before the claimant has suffered any reliance loss. However, once the claimant has incurred expenses in reliance on his reasonable assumptions, the defendant will be liable to compensate the claimant for any reliance loss. Such a duty may be reminiscent of the continental duty of good faith, but they have two important conceptual differences. First, the duty of good faith suggests that both parties are expected to act in their common interests, while the duty to ensure the reliability of induced assumptions allows parties to take adversarial positions short of misleading the other party. Second, the duty to ensure reliability of induced assumptions is a theoretical duty and does not attract legal liability when it is breached per se. The claimant can only obtain compensation if he has a substantive cause of action against the defendant.¹⁹

It may be argued that a claimant who does work in the absence of a contract takes the risk of non-payment and hence cannot claim to suffer harm. However, such a view is overly broad and detached from commercial reality. For example, it is the industry norm for contractors to commence work once a letter of intent is received and it is unrealistic for contractors to delay performance until the contract is finalized. Giliker cites two common law studies²⁰ to show that parties are unlikely to “approach negotiations in an antagonistic manner” where they “share common interests, are dependent on each other or wish to develop a long-term relationship”.²¹ McFarlane explains that insistence on a commitment from the defendant may send the defendant

¹⁸ M Spence, *Protecting Reliance: The Emergent Doctrine of Equitable Estoppel* (Hart Publishing, Oxford and Portland, Oregon 1999), 2. [Spence]

¹⁹ Eg. contract, estoppel or restitution.

²⁰ S Macaulay, ‘Non-contractual relationships in business’ (1963) 28 Am Soc Rev 55 and H Beale and T Dugdale, ‘Contracts between businessmen: planning and the use of contractual remedies’ (1975) 2 Brit J Law and Soc 45 cited in P Giliker, *Pre-contractual Liability in English and French Law* (Kluwer Law International, The Hague 2002), 176–177. [Giliker]

²¹ Giliker, *ibid.*, 176 – 177.

unwelcome signs as to the claimant's suitability as a contracting party and commercial parties may prefer to trust that notions of reciprocal fairness will provide a sufficient incentive for performance.²² In such situations, it would be inappropriate for the claimant to insist on entering into a binding contract before making any pre-contractual investment.

Although a harm theory based on a duty to ensure reliability of induced assumptions presents a normative reason why a pre-contractual expenditure ought to be recoverable, it cannot adequately explain some anticipated contract cases. For instance, where Y does work on X's land under an agreement (expressly made '*subject to contract*'), Y may be reasonably led to assume that he had a good chance of obtaining the contract. Since the agreement was '*subject to contract*', Y ought to know that he risks non-payment for the work done. However, Y's risk-taking does not negate his assumptions that a contract might be forthcoming; Y still believes that a contract might be made eventually. How should one reconcile the theoretical duty with the concept of risk-taking? It would be unrealistic to say that risk-taking obviates any assumptions since it clearly does not. Risk is merely ancillary to any assumptions Y may make.

The duty also fails to justify recovery in instances where there is an absence of overt representations. By virtue of the smooth negotiations, both parties may share a common belief that a contract will eventuate in due course. However, since there was no clear act of inducement, Y's assumptions that a contract will materialize simply arose from his own perception of the negotiations. Therefore, X would not be liable to compensate Y for Y's pre-contractual expenditure. However, if X obtains a benefit as a result of Y's efforts, X would be unjustly enriched at the expense of Y. It has been argued that the theoretical basis of unjust enrichment is founded upon the idea of corrective justice.²³ Where the equality of the parties is disturbed in relation to a transaction between them, the imposition of restitutionary liability on one to the other would restore the parties to equilibrium.²⁴ The disruption to the equality of X and Y translates into a type of harm to Y, however the duty to ensure reliability of induced assumptions is too narrow to address such harm. It is then inaccurate to see the duty as synonymous with the harm theory. The duty is merely a sub-set of the harm theory. There is thus a need to supplement this duty with another variation of the harm theory, viz. corrective justice.

3.2.2 *Corrective Justice*

The idea of harm in the context of corrective justice is not based on wrongdoing, but on the broader notion of normative gains and losses. Weinrib perceives the fundamental concern of the Aristotelian notion of corrective justice to be the maintenance of transactional equality.²⁵ Where one party upsets an existing distribution of resources

²² B McFarlane, "The Problem of Pre-Contractual Reliance: Three Ways to a Third Way", Hauser Global Law School Program, GLWP 12/06, 35, http://www.law.nyu.edu/global/workingpapers/2006/ECM_DLW_015762, viewed: 1/7/2010

²³ E Weinrib, *The Idea of Private Law* (Harvard University Press, Cambridge, Massachusetts 1995), 133–134. [Weinrib]; K Barker, 'Unjust Enrichment: Containing the Beast' (1995) 15 OJLS 457, 468–471; L Smith, 'Restitution: The Heart of Corrective Justice' (2001) 79 Texas LR 2115; R Grantham and C Rickett, 'A Normative Account of Defences to Restitutionary Liability' (2008) 67 CLJ 92, 96–103. [Grantham and Rickett]

²⁴ Weinrib, *ibid.*, 117; Grantham and Rickett, *ibid.*, 98.

²⁵ Weinrib, *ibid.*, 114 – 120.

between himself and another, liability is imposed on the party to restore the parties to the equilibrium position. While Aristotle failed to justify the normative content of corrective justice, Weinrib borrows the Kantian concept of right to provide an explanation.²⁶ The Kantian right encompasses two types of legal rights – the right to the integrity of one's body and the right concerning the external objects of the will.²⁷ The former addresses the traditional concept of harm as commonly invoked under tort law; the latter includes rights to property and to contractual performance.²⁸ In the pre-contractual context, we are more concerned with the right concerning the external objects of the will.

Grantham and Rickett derive two conditions for restitutionary liability from Weinrib's theory of corrective justice.²⁹ First, the transfer of wealth or value by the claimant to the defendant must be defective in the sense that the claimant did not exercise his choice freely. Where Y does work on X's land on the basis that the parties would conclude a deal in due course, and the contract fails to eventuate, the basis for the transfer of wealth becomes defective. However, where Y knows that there is a clear risk of non-payment and he freely exercises his choice to do work despite the risk, he cannot thereafter claim that the transfer of wealth is defective. Second, there must be some reason or circumstance to implicate the defendant's autonomy. This accords respect to the defendant's interest in self-determination and his right over his personal wealth. Weinrib suggests that the defendant's autonomy should be implicated where the value was not transferred gratuitously. Where Y does work in expectation of remuneration through a future contract, X cannot establish the benefit as part of his personal wealth without making some form of payment for it. X cannot demand respect to his autonomy without first respecting the autonomy of Y. However, Y is taken to have done the work gratuitously where he knows that there is a risk of non-payment but proceeds with it anyway. The value accrued thus forms part of X's personal wealth and Y cannot insist on restoring the parties to the equilibrium position.

While corrective justice provides an alternative justification to the duty to ensure reliability of induced assumptions for the imposition of pre-contractual liability, it is crucial to view corrective justice only as a complementary theory. There are problems that cannot be adequately addressed by this theoretical model. Assume that after Y has done work on X's land, X decides not to develop his land. The work done on the land is of no value to X. Neither has X been saved any necessary expenses since X would not have paid a third party to do the work that Y had done. Assume further that Y's contribution to X's land had not increased its value in any way. In fact, the uncompleted works decreased the value of the land since any potential buyer would need to clear the land before developing it. Here, Y has incurred reliance expenses, but X has not received any benefits. There is no transfer of value. Instead, there is an actual diminution of the collective wealth of both parties. There is no transactional inequality here that can be addressed by the notion of corrective justice. However, it is instinctively unfair that X's liability to compensate Y is based on the adventitious circumstances of whether any value has been transferred. Corrective justice loses its

²⁶ *Ibid.*, 122.

²⁷ *Ibid.*, 128.

²⁸ *Ibid.*, 128.

²⁹ *Grantham and Rickett, supra* note 23, 100–101.

explanatory force in such situations, and must thus be supplemented by the duty to ensure reliability of induced assumption.

4 The Concept of Risk

The courts are generally reluctant to impose any pre-contractual liability on the defendant when the claimant does work under the assumption of risk. Despite the central importance of the concept of risk in anticipated contract cases, there have been no attempts to elucidate the content of risk. Stoljar has defined this risk as “simply the risk we run where an agreement in which nothing is said about remuneration might yet be construed as remunerative, especially where the work done during that agreement cannot be considered to be gratuitous”.³⁰ Ultimately, the issue of who should bear the risk is essentially a different manner of asking who should be liable to pay for pre-contractual preparations.³¹ The idea of risk is so intrinsically entwined with the fact-finding process that any attempts to give it substantive content would be tantamount to chasing the will-o'-the-wisp. Nevertheless, it may be helpful to distinguish the two different ways in which risk-taking may arise – the risk that has been allocated by the parties as a matter of fact and the risk that is imposed by law based on all the circumstances.

4.1 Allocation of Factual Risk

Whether there has been an allocation of risk is a matter of fact based on evidence. In the absence of a contract, courts have often relied on various factors to determine whom the risk of negotiations failing lies upon. Barker has helpfully distilled six evidential factors relevant to the determination of risk from his sample of pre-contract cases.³²

- (1) whether negotiations were ‘*subject to contract*’;³³
- (2) whether the work was an ‘accelerated part of contract performance’ or was merely ‘preparatory’ to the contract;³⁴
- (3) whether the work was for a purpose extraneous or collateral to the contract itself, so as to fall outside its ‘ambit’;³⁵
- (4) whether the defendant requested the work done;³⁶
- (5) whether the work was of a kind normally provided free of charge;³⁷
- (6) whether the claimant was acting in order to benefit the defendant, or out of his or her own ‘self-interest’.³⁸

³⁰ S J Stoljar, *The Law of Quasi-Contract* (2nd edn Law Book Co, Sydney 1989), 243–244.

³¹ *Spence*, *supra* note 18, 93.

³² *Barker*, *supra* note 5, 119–127.

³³ *Ibid.*, 120. This factor is a very strong indication that parties have allocated the risks. However, there could be instances where the parties used the term ‘subject to contract’ without appreciating its meaning. In such cases, the court may be more reluctant to find that term reflects the parties’ risk allocation.

³⁴ *Ibid.*, 122–123.

³⁵ *Ibid.*, 123–124.

³⁶ *Ibid.*, 124–125.

³⁷ *Ibid.*, 125.

The above factors are meant to provide useful guidance to the courts in finding where the risk lies, but they are by no means exhaustive. The court may take into account of other factors, and whether there has been an allocation of risk must ultimately be determined on a case-by-case basis. A finding that the claimant has in fact taken the risk generally precludes the operation of any substantive rules of recovery.³⁹

4.2 Normative Risk Imposed by Law

Even where there is insufficient evidence to support a finding that the parties have allocated the risk, the courts may nonetheless find that the claimant should be deemed to have taken the risk based on all the circumstances. In *Brewer Street*,⁴⁰ where work done in anticipation of a leasing agreement was wasted when the agreement failed to materialize, Denning LJ stated:

“The question is: on whom is the loss to fall? The parties themselves did not envisage the situation which has emerged and did not provide for it; and we do not know what they would have provided if they had envisaged it. Only the law can resolve their rights and liabilities in the new situation, either by means of implying terms or, more simply, by asking *on whom should the risk fall*.”⁴¹ (Emphasis added)

More recently, Lord Hoffman, in the context of a claim for mistaken payment, adopted the view that the risk is imputed:

“whether a party *should be treated as having taken the risk* depended upon the objective circumstances surrounding the [enrichment] as they could reasonably have been known to both parties”.⁴² (Emphasis added)

What are the considerations taken into account by the courts when deciding on the normative risk? Although there is no clear judicial guidance on this issue, a few factors can be inferred from the cases. The first is the relative gains and losses of the parties. Where a claimant derives a benefit from his labour, while the defendant could not benefit from it, the claimant will usually be made to bear the loss.⁴³ On the contrary, the courts will be more reluctant to find in favour of the defendant if the claimant does not obtain any gains.⁴⁴ Second, the circumstances in which the negotiations broke down are germane.⁴⁵ If the circumstances are beyond what reasonable parties would have directed their minds to, the court is more likely to find that they fall outside the scope of risk undertaken by the claimant at the outset. For example, if the project is a large one which is likely to take a long time to negotiate, the claimant is more likely to be deemed to have assumed the risk because parties can reasonably expect negotiations to be

³⁸ *Ibid.*, 126–127.

³⁹ See *Regalian Properties plc v London Docklands Development Corporation* [1995] 1 WLR 212 (Ch) [*Regalian*]; *Yeoman's Row*, *supra* note 2, [25]. Although Lord Scott, in *Yeoman's Row*, was referring to a proprietary estoppel claim, the reasoning applies *mutatis mutandis* to promissory estoppel.

⁴⁰ *Supra* note 4.

⁴¹ *Supra* note 4, 436.

⁴² *Deutsche Morgan Grenfell v IRC* [2007] 1 AC 558 (HL) [28].

⁴³ *Jennings*, *supra* note 4.

⁴⁴ *Brewer Street*, *supra* note 4.

⁴⁵ *Brewer Street*, *supra* note 4; *Sabemo Pty Ltd v North Sydney Municipal Council* [1977] NSWLR 880 (New South Wales Supreme Court, Australia) [*Sabemo*]; *Countrywide*, *supra* note 11.

hampered by changing economic conditions.⁴⁶ The third consideration is whether there is asymmetrical knowledge of the pre-conditions to the anticipated contract. In *Jennings*,⁴⁷ the lessee proposed to sublet the premises to a solicitor and agreed to convert the place into an office. The sublease did not materialize due to certain covenants in the head lease. The court imposed the normative risk on the lessee since he knew of the obstacles but did not alert the solicitor to them. Again, these factors are neither exhaustive nor conclusive and each problem must be approached individually.

5 A Principled Approach – The Gradation Model

5.1 The Root of Divergence

The divergent approaches towards pre-contractual cases are triggered by their fundamentally distinct starting points – the reliance detriment suffered by the claimant because his induced assumption has been proved unjustified and the transactional inequality when value is unjustly transferred from the claimant to the defendant. The estoppel argument addresses the former, while the unjust enrichment argument tackles the latter. There is also a perpetual tension between the need to respect the parties' freedom of contract and the desire to deter parties from abusing this freedom. The factual risk analysis and contractual solutions value the notion of freedom of contract. Fault-based decisions determine recovery in accordance with the degree of culpability of the defendant. Despite the different routes taken, they all lead to the common goal of serving justice.

5.2 A Gradation Model

This paper proposes a gradation model that attempts to incorporate the different approaches into a neat framework that achieves the rationale of the imposition of pre-contractual liability. The purpose of this model is not to prove that there is one substantive rule that fits all situations, but rather it is to show that the adoption of one approach to the exclusion of others would neglect the various objectives that the imposition of pre-contract liability is supposed to achieve. There is no one correct rule to use in pre-contractual cases, but only the correct rule to use in achieving the desired aims.

Since the imposition of pre-contractual liability is to remedy two different types of harm, the model must consist of two causes of action: promissory estoppel to protect the claimant's detrimental reliance⁴⁸ and unjust enrichment to reverse any transactional inequality. Although there may be a degree of overlap, the two harms are of a different nature and there is no single action that can appropriately address the harms concurrently. A claimant who wishes to protect his reliance expenses should bring an

⁴⁶ *Countrywide*, *supra* note 11, 348.

⁴⁷ *Supra* note 4.

⁴⁸ The exact legal nature of the estoppel is beyond the scope of this paper. While English law does not recognize promissory estoppel as a cause of action, the difficulties of using promissory estoppel as a sword is not insurmountable. American and Australian courts have favoured the doctrine of promissory estoppel as the answer to imposing pre-contractual liability. (See *Hoffman v Red Owl Stores*, 133 NW 2d 267 (1965); *Cyberchron Corp. v Calldata Systems Development, Inc.*, 47 F 3d 39 (2nd Cir. 1995); *Walton Stores*, *supra* note 3; *Verwayen*, *supra* note 3.) Promissory estoppel and variations of estoppel have also received substantial academic support (see *Giliker*, *supra* note 20, 151; *Spence*, *supra* note 18, 87; *McFarlane*, *supra* note 22, 11–12; R Goff and G Jones, *The Law of Restitution* (7th edn Sweet & Maxwell, London 2007), 672.)

action under promissory estoppel, while a claimant who desires to obtain the value unjustly transferred to the defendant should plead unjust enrichment. The model does not include a contractual approach because it simply expands the notion of contract so as to preclude the contract in question from failing. Hence, that is a separate question of how far should the ambit of contract law stretch and is beyond the scope of a model that addresses pre-contractual situations.

A sliding scale of fault and risk-taking constitutes the spine of the model. (See Diagram 1 below) On one end of the spectrum, we have a scenario in which Y has taken the greatest degree of risk by doing work under an agreement that was ‘*subject to contract*’ and in which X is relatively fault free. The court will be unlikely to impose any pre-contractual liability in this situation. On the other end of the spectrum, we have a scenario in which Y has taken minimal risk and in which X’s conduct was culpable (eg. inducing Y into providing work for free when X had no intention to contract with Y). The court will be more likely to impose pre-contractual liability in this situation. In any particular scenario, the chances of success of a claim for pre-contractual compensation will depend on the interplay between the elements of risk and fault, subject to the substantive principles underlying the two causes of action.

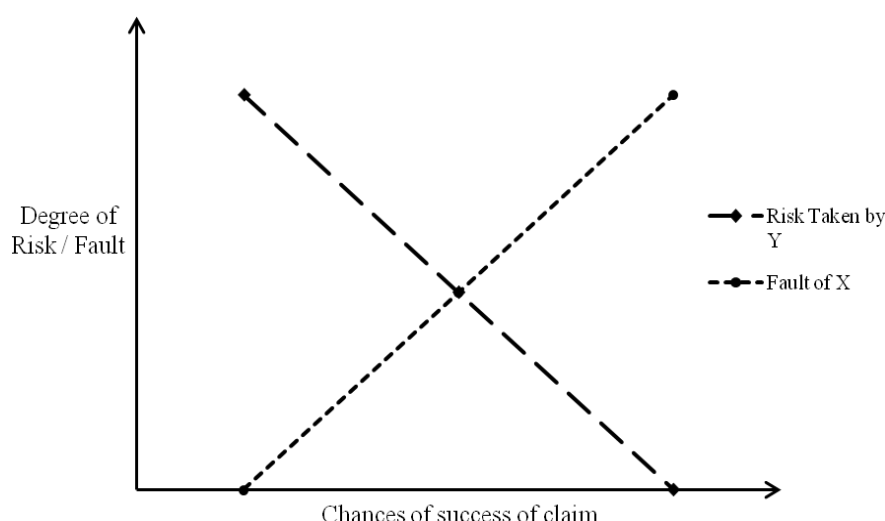


Diagram 1. Gradation Model

5.2.1 Reconciling the Causes of Action with the Gradation Model

The causes of action relate back to the theoretical bases of pre-contractual liability, therefore the gradation model cannot be presented in isolation from the causes of action. If risk and fault form the skeleton of the model, then promissory estoppel and unjust enrichment constitute the flesh of this theory. Their relationship can be illustrated through the use of a pyramidal structure (see Diagram 2). In order to establish pre-contractual liability, the claimant must pursue his claim under either of the two causes of action. The results of the causes of action will in turn depend on the degree of risk and fault present in each factual scenario. The connection between risk, fault and the causes of action will be expounded next.

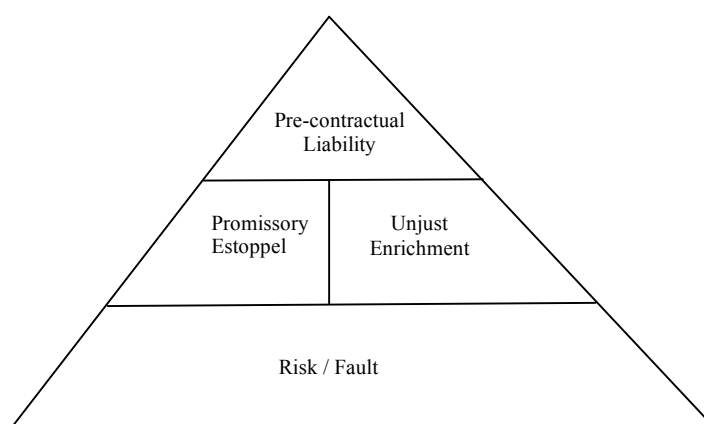


Diagram 2. Pyramidal Structure of Elements in Gradation Model

5.2.2 *Promissory Estoppel*

Promissory estoppel is often equated with a fault-based model and the language of risk-taking is conspicuously absent from the estoppel analysis. The concept of risk is an important aspect of establishing a claim under promissory estoppel but it has been masked under the unconscionability inquiry. Although the label “unconscionability” has commonly been associated with vagueness, Spence has provided us with a principled approach to determining unconscionability with his list of eight criteria:⁴⁹

- (1) How the assumption and reliance upon it were induced?
- (2) What is the content of the relevant assumption?
- (3) What is the state of relative knowledge of the parties?
- (4) What is the parties’ relative interest in the relevant activities in reliance?
- (5) What is the nature and context of the parties’ relationship?
- (6) What is the parties’ relative strength of position?
- (7) What is the history of the parties’ relationship like?
- (8) Whether the party who has induced the relevant assumption has already taken any steps to ensure that he has not caused preventable harm.

The focus here will be on factors (4), (7) and (8) because they are closely related to the element of risk.

The court is more likely to find unconscionability where the relying party has a greater interest in the relevant activities in reliance than the inducing party. The rationale is that “the greater the interest of the party against whom the estoppel is sought

⁴⁹ Spence, *supra* note 18, 59–66.

in the other party's reliance, the more likely it is that he induced it".⁵⁰ Spence contrasts the case of *Walton Stores* with *Powprop Pty Ltd v Valbirn Pty Ltd*⁵¹ to support this view. Both cases involved construction work undertaken in anticipation of a prospective lease. In *Walton Stores*, the inducing party, who was the lessor, had a stronger interest in having the relying party act upon the assumption (that the formalities for the lease would be completed in due course) to build a premise to certain specifications. But in *Powprop*, the relying party, who was the lessor, had a greater interest because he "was extremely anxious to enter into a lease with the [lessee], and was prepared to 'risk the expense of the alterations in the hope that it might achieve what it wanted.'"⁵² It thus appears that where the relying party takes the risk of expenditure to further his own interests, there is a greater likelihood that his expenses were not incurred in reliance on the inducement and it will not be unconscionable for the inducing party to resile from his representation.

Spence deems the history of the parties' relationship as relevant towards a finding of unconscionability because reliance will more readily be induced in a lengthy relationship.⁵³ Where the relying party has done work over a long period of time for the inducing party, the court is less likely to find that the relying party has taken the risk and provided his services gratuitously. In *Sabemo*,⁵⁴ the claimant did substantial work over a period of three years, in the absence of a contract, before the defendant aborted the building project. In arriving at its decision that the defendant was liable to compensate the claimant, the court took cognizance that work had been done over three years, and it would be "unthinkable that the [claimant] would have been prepared to do what it did if it thought that the defendant might change its mind about proceeding with the proposal".⁵⁵

The inducing party would be deemed to have taken steps to ensure that he has not caused preventable harm where he has given notice that the assumption is not to be relied upon.⁵⁶ The classical example to this is where pre-contractual dealings are expressly provided to be '*subject to contract*'.⁵⁷ As discussed earlier, the use of the term '*subject to contract*' is usually clear evidence of risk allocation by the parties and the loss would fall where it lies when a party takes the risk in doing work. Where parties have clearly expressed their intentions not to be bound, the court should be slow to find reliance upon any assumptions.

⁵⁰ *Ibid.*, 63.

⁵¹ [1990] ANZ Conv Rep 529 (Supreme Court of Queensland, Australia).

⁵² *Ibid.*, 530.

⁵³ *Spence*, *supra* note 18, 64.

⁵⁴ *Supra* note 45.

⁵⁵ *Sabemo*, *supra* note 45, 319.

⁵⁶ *Spence*, *supra* note 18, 65.

⁵⁷ *Attorney General of Hong Kong v Humphreys Estate (Queen's Garden) Ltd* [1987] 1 AC 114 (HL); *Regalian*, *supra* note 39.

5.2.3 Unjust Enrichment

Restitution scholars have oft mention that a risk taker should not be allowed to plead that the other party has been unjustly enriched.⁵⁸ While there is general acceptance of this view, the exact relationship between risk-taking and unjust enrichment has been relatively unexplored by academics and judges alike. In *Yeoman's Row*,⁵⁹ the House of Lords noted that the claimant knew that the "agreement in principle" was only "binding in honour" but he chose to provide his services notwithstanding this.⁶⁰ The language of risk-taking was expressly used by Lord Walker, who referred to the claimant's undertaking of risk in deciding to rely on the defendant's sense of honour.⁶¹ Despite the risk taken by the claimant, their Lordships imposed restitutionary liability on the defendant to reimburse the claimant's outgoings and pay him a fee for his services. The decision thus appears to be incompatible with the dominant academic opinion that risk taking is incompatible with unjust enrichment.

This paper made the distinction between factual risk and normative risk earlier on and it is submitted that this distinction is essential to the understanding of the relationship between risk-taking and unjust enrichment. The finding of a factual risk is an evidentiary finding that the parties had, as a matter of fact, allocated the risk and the courts should therefore be slow to interfere with the conscious decision of the parties. Where the evidence is not enough to support a clear finding that the parties had allocated the risk, the court may nevertheless, as a matter of law, deem that the claimant had taken a risk if the existing evidence and other circumstances warrant such a conclusion. A parallel can be drawn with the situation where the parties have made a contract. Unjust enrichment cannot operate in the existence of a contract because that would "disturb the legitimate hope and fears inherent in the bargain".⁶² However, where there is a gap in the contractual allocation of risk, restitution should be allowed since it would not unsettle the bargain of the parties.⁶³ Referring back to diagram 1 above, where the parties have provided that their negotiations are '*subject to contract*', it will be a clear case that the parties have expressly allocated their risk so as to preclude an unjust enrichment claim. If there is insufficient evidence to support a clear finding of risk, then whether the claimant should be deemed as a risk taker would depend on the application of the principles of unjust enrichment. The notion of a normative risk is thus inherent in the substantive content of unjust enrichment.

We shall now turn to the issue of reading normative risks into the principles of unjust enrichment. Barker suggests that several of his factors are relevant both to the

⁵⁸ A Burrows, *The Law of Restitution* (2nd edn Butterworths LexisNexis, London 2002), 377 [Burrows]; P Birks, *Unjust Enrichment*, (1st edn Oxford University Press, New York 2003), 129 [Birks 2003]; G Virgo, *The Principles of the Law of Restitution*, (2nd edn Oxford University Press, New York 2006), 343.

⁵⁹ *Supra* note 2.

⁶⁰ *Yeoman's Row*, *supra* note 2, [15] (Lord Scott) and [74] (Lord Walker).

⁶¹ *Ibid.*, [72].

⁶² J Beatson, 'The Temptation of Elegance: Concurrence of Restitutionary and Contractual Claims' in W Swadling and G Jones (eds), *The Search for Principle: Essays in Honour of Lord Goff of Chieveley* (Oxford University Press, New York 1999), 142, cited in P Birks, 'Failure of Consideration and Its Place on the Map' (2002) OJCLJ 1, 5.

⁶³ See *Roxborough v Rothmans of Pall Mall Australia Ltd* (2002) 208 CLJ 516 (High Court of Australia); J Beatson, 'Restitution and Contract: Non-Cumul?' (2001) 1 Theoretical Inquiries in Law 83, 94 cited in R Cunningham, 'Failure of Basis' (2004) LMCLQ 234, 250–251; J Beatson and G Virgo, 'Contract, Unjust Enrichment and Unconscionability' (2002) 118 LQR 352, 356; Tang H W, 'Unjust Enrichment within a Valid Contract: a Close Look at *Roxborough v Rothmans of Pall Mall Australia Ltd*' (2007) 23 JCL 201, 210.

issue of risk allocation, as well as to the unjust enrichment inquiry.⁶⁴ The type of work done may impact on the finding of an enrichment received by the defendant. Advance contractual performance is more likely to be beneficial to the receiver than preparatory work meant to help the claimant in his contractual bid.⁶⁵ Work collateral to the failing contract may also be more valuable than work closely connected with the proposed contract because it has a value independent of the venture's success.⁶⁶ The kind of work done may also affect the finding of injustice. If the work in question is usually provided free of charge in the industry, it may negate a claim based on failure of consideration or free acceptance since an unconditional donative intent indicates that the claimant had consented to the defendant receiving the benefit without payment.

The concept of risk may also be relevant in the determination of the basis on which the claimant did work. The identification of the basis is necessary for the unjust factor of failure of consideration to apply, but "in the absence of an express stipulation as to the basis upon which pre-contract work is carried out, it is not easy to discern the basis upon which work is done".⁶⁷ In such cases, the courts will have to infer the performing party's state of mind from the surrounding facts and circumstances. There are three reasons why a party would confer a pre-contractual benefit. First, the party may have an unconditional donative intent. Even if the party knew from the outset that he has no chance of ever being remunerated for his work, he would so proceed nonetheless. There may be other ancillary reasons why the party would commit unconditionally, such as to build up his reputation in the industry, but these reasons do not detract from the party's primary intention to confer the benefit out-and-out. Such situations are, however, unlikely to occur in commercial cases. Second, the party may have conferred the benefit in the *hope* that he would secure a contract with the receiving party subsequently. Third, the party may have done work under the assumption that it would be paid for under subsequent arrangements. The first situation will rarely arise in reality and can only be proven where there is clear and unequivocal evidence. The basis for the conferment of a benefit usually falls into the second or third categories; however, only the third situation will give rise to a failure of consideration. How do the courts then distinguish between the two bases? Birks addresses this problem in the following passage:

"There is a fine line between, on the one hand, wanting a contract to come into existence but taking the risk that it may not, and, on the other, acting solely on the basis that it will. In general this line is found by asking whether the claimant made it clear to the recipient that he was not taking that risk or *reasonably believed* that the recipient knew that he was not."⁶⁸

Burrows proposes a similar test to the one adopted by Birks above:

"In line with the standard meaning of failure of consideration, the claimant alleging that the rendering of the benefit was conditional on an unpromised future event can only avoid being treated as a risk-taker not entitled to restitution

⁶⁴ *Barker*, *supra* note 5, 122–127.

⁶⁵ *Ibid.*, 123.

⁶⁶ *Ibid.*

⁶⁷ *Ibid.*, 184.

⁶⁸ *Birks* 2003, *supra* note 58, 129.

if he either expressly communicated the contingent condition to the defendant or was *reasonably led to believe* by the defendant that he (the defendant) knew that the contingent condition was the basis for the benefit being rendered.”⁶⁹

The essence of the two tests can be restated as a 2-stage inquiry. First, did the claimant communicate his expectation to be remunerated under subsequent arrangements to the defendant? If he did, there is a failure of consideration. Otherwise, the second question must be asked: Did the claimant have a reasonable belief that the defendant knew that he (the claimant) expected to be remunerated under subsequent arrangements? If the answer is affirmative, there is a failure of consideration. However, if the question is answered in the negative, then the claimant is a risk-taker and all he has is merely a *hope* for remuneration. The reasonable belief of the claimant and his risk-taking are thus two sides of a same coin and depend on the same factual inquiry. Failure of consideration is inextricably linked to the element of normative risk.

The next obstacle to reconciliation with the gradation model is the apparent incompatibility of unjust enrichment and fault. Unjust enrichment has always been seen as a strict liability claim and is generally independent of fault. However, it is an overstatement to say that fault is always irrelevant.⁷⁰ There are a few exceptional situations where the fault of the defendant is an essential ingredient of an unjust enrichment claim. For example, bad faith will deprive the defendant of the availability of the change of position defence. Fault can also be relevant in the identification of an enrichment and a defendant is prevented from subjectively devaluing a benefit after he has freely accepted the benefit. Furthermore, fault plays an important role as a ground of restitution under free acceptance. Although free acceptance is a controversial unjust factor due to its fault-based nature,⁷¹ once its efficacy is denied, it follows that either unjust enrichment is unable to justify many pre-contractual cases or that the cases which refer to fault are wrongly decided.⁷² Birks clearly valued free acceptance as an unjust factor and persuasively argued for a residual role for fault-based restitution when the inquiry into plaintiff-sided factors has produced a negative result.⁷³ Without free acceptance, a risk-taker (in the factual sense) cannot protest that the other party has been unjustly enriched, even if the benefit had been extracted unconscientiously. However, free acceptance:

“works because of the unconscientiousness of the recipient in not availing himself of the opportunity to save the intervener from the risk. The unconscientious nature of the resulting receipt is what justifies tipping the balance in favour of the intervener and allowing the *quantum meruit*.”⁷⁴

⁶⁹ Burrows, *supra* note 58, 408.

⁷⁰ See G Virgo, ‘The Role of Fault in the Law of Restitution’ in A Burrows and A Rodger (eds), *Mapping the Law: Essays in Memory of Peter Birks* (Oxford University Press, Oxford; New York 2006), 83.

⁷¹ See P Birks, *An Introduction to the Law of Restitution* (Clarendon Press, Oxford; Oxford University Press, New York 1985), 265; A Burrows, ‘Free Acceptance and the Law of Restitution’ (1988) 104 LQR 576; P Birks, ‘In Defence of Free Acceptance’ in A Burrows (ed), *Essays on the Law of Restitution* (Clarendon Press, Oxford 1991), 105–146 [Birks 1991]; Burrows, *supra* note 58, 402–407.

⁷² Rattee J appears to be of this view when he doubted whether the fault principle exists in English law in *Regalian*, *supra* note 39, 230–231.

⁷³ Birks 1991, *supra* note 71, 144–145.

⁷⁴ *Ibid.* at 111. It is unfortunate that Birks uses the term ‘*quantum meruit*’ since its legal basis is unclear (*supra* note 5). But *quantum meruit* here is clearly used in the restitutionary sense.

The unique fault-based nature of free acceptance allows it to justify recovery in instances where the claimant was a risk taker (as a matter of fact).

5.3 The Gradation Model Framework

This paper has established the conceptual relationship between the two causes of action and the basic ideas of risk and fault. Next, a framework to approach pre-contractual cases shall be proposed. The question of pre-contractual liability can be broken down into a four-stage inquiry:

- (1) Do the circumstances of the case support a clear finding of risk allocation by the parties?⁷⁵
- (2) Is the conduct of the defendant of a sufficient degree of fault to negate any factual risk taking on the part of the claimant?
- (3) What is the harm that the claimant desires to remedy?
- (4) Are the elements of the cause of action satisfied on the facts of the case?

In stage (1), a clear finding that the claimant has taken the risk of non-recovery of his investment will generally preclude the imposition of pre-contractual liability. The parties have struck their bargain and the courts should refrain from interfering unless there are exceptional circumstances that warrant intervention. Whether there is a valid case for judicial intervention falls under the stage (2) inquiry. Unless the behaviour of the defendant falls far below the standards of reasonable businessmen, the court should be slow to disregard the risks allocated by the parties. The standard is neither one of good faith nor complete honesty since it has been recognized that parties do take “adversarial positions” in negotiations⁷⁶ and business people cannot be expected to put their “cards face upwards on the table”⁷⁷. If the parties have not allocated the risks under stage (1) or if they have done so, but the case falls under the exceptional circumstances in stage (2), the court shall proceed to stage (3). Otherwise, the pre-contractual claim fails. In stage (3), the court will need to identify whether the claimant is pleading to protect his detrimental reliance or for the reversal of a transactional inequality. The type of harm will necessarily determine the cause of action. Finally, the courts will apply the principles of promissory estoppel or unjust enrichment in stage (4) to determine if the case for imposition of pre-contractual liability has been made out.

⁷⁵ See Section 4.1 for the list of factors that may provide useful guidance.

⁷⁶ *Walford v Miles*, *supra* note 13, 138.

⁷⁷ *Interfoto Picture Library Ltd v Stiletto Visual Programmes Ltd* [1989] QB 433 (CA) 439 (Bingham LJ).

6 Conclusion

There is a need to bring consistency and order to the troubled realm of anticipated contracts, but that cannot warrant the sacrifice of substance for form. An approach founded upon one single liability model is prone to the distortion of legal principles and must be rejected for its artificiality. While the imposition of pre-contractual liability is desired from an efficiency and moral perspective, the legal response should always correspond to the type of harm in question. The paper proposes that promissory estoppel should be invoked to remedy detrimental reliance, and unjust enrichment should be adopted to correct any transactional inequalities between parties. Risk and fault are not separate legal categories from the traditional models of liability, but the building blocks of the legal principles. Cases can be arranged along a gradation scale based on the degree of risk and fault involved, and that constitutes the starting point of every pre-contractual legal inquiry. Although the gradation model admittedly does not eliminate the uncertainties that arise from the fact-sensitive nature of pre-contractual problems, it reconciles the piecemeal solutions into a coherent picture. The contractor in dilemma may henceforth be able to enforce his legal rights on a more principled basis.

7 Acknowledgements

I would like to express my deepest gratitude to Associate Professor Tang Hang Wu (Faculty of Law, National University of Singapore), without whose guidance through the world of restitution I would have been hopelessly lost.

8 References

- Bailey, J., 'Repudiation, Termination and Quantum Meruit' (2006) 22(4) Const LJ 217
- Barker, K., 'Coping with Failure – Reappraising Pre-Contractual Remuneration' (2003) 19 JCL 105
- Barker, K., 'Unjust Enrichment: Containing the Beast' (1995) 15 OJLS 457
- Beale, H. and Dugdale, T., 'Contracts between businessmen: planning and the use of contractual remedies' (1975) 2 Brit J Law and Soc 45
- Beatson, J., 'Restitution and Contract: Non-Cumul?' (2001) 1 Theoretical Inquiries in Law 83
- Beatson, J., 'The Temptation of Elegance: Concurrence of Restitutionary and Contractual Claims', In: Swadling, W. and Jones, G. (eds), *The Search for Principle: Essays in Honour of Lord Goff of Chieveley* (Oxford University Press, New York 1999)
- Beatson, J. and Virgo, G., 'Contract, Unjust Enrichment and Unconscionability' (2002) 118 LQR 352
- Birks, P., *An Introduction to the Law of Restitution* (Clarendon Press, Oxford; Oxford University Press, New York 1985)
- Birks, P., *Unjust Enrichment*, 1st ed. (Oxford University Press, New York 2003)
- Birks, P., 'Failure of Consideration and Its Place on the Map' (2002) OJLS 1
- Birks, P., 'In Defence of Free Acceptance' in Burrows, A. (ed), *Essays on the Law of Restitution* (Clarendon Press, Oxford 1991)
- Burrows, A., *The Law of Restitution* (2nd edn Butterworths LexisNexis, London 2002)
- Burrows, A., 'Free Acceptance and the Law of Restitution' (1988) 104 LQR 576
- Cohen, N., 'Pre-Contractual Duties: Two Freedoms and the Contract to Negotiate', In: Beatson, J. and Friedmann, D. (eds), *Good Faith and Fault in Contract Law* (Clarendon Press, Oxford 1995), 25
- Cunnington, R., 'Failure of Basis' (2004) LMCLQ 234
- Farnsworth, E., *Contracts* (4th edn Aspen Publishers, New York 2004)
- Farnsworth, E., 'Precontractual Liability and Preliminary Agreements: Fair Dealing and Failed Negotiations' (1987) 87 Col L Rev 218
- Furst, S. and Ramsey, V., *Keating On Construction Contracts* (8th edn Sweet & Maxwell, London 2006)
- Giliker, P., *Pre-contractual Liability in English and French Law* (Kluwer Law International, The Hague 2002)
- Goff, R. and Jones, G., *The Law of Restitution*, (7th edn Sweet & Maxwell, London 2007)
- Graham, R. and Rickett, C., 'A Normative Account of Defences to Restitutionary Liability', (2008) 67 CLJ 92
- Macaulay, S., 'Non-contractual relationships in business' (1963) 28 Am Soc Rev 55

- McFarlane, B., 'The Problem of Pre-Contractual Reliance: Three Ways to a Third Way', Hauser Global Law School Program, GLWP 12/06, http://www.law.nyu.edu/global/workingpapers/2006/ECM_DLW_015762, viewed: 1/7/2010
- McKendrick, E., 'Work Done in Anticipation of a Contract which does not Materialise' in Cornish, W.R., et al. (ed.), *Restitution Past, Present and Future: Essays in Honour of Gareth Jones* (Hart Publishing, Oxford 1998), 163
- Mill, J. S., *On Liberty* (Yale University Press, New Haven 2003)
- Smith, L., 'Restitution: The Heart of Corrective Justice' (2001) 79 Texas LR 2115
- Spence, M., *Protecting Reliance: The Emergent Doctrine of Equitable Estoppel* (Hart Publishing, Oxford and Portland, Oregon 1999)
- Stoljar, S. J., *The Law of Quasi-Contract* (2nd edn Law Book Co, Sydney 1989)
- Tang, H. W., 'Unjust Enrichment within a Valid Contract: a Close Look at *Roxborough v Rothmans of Pall Mall Australia Ltd*' (2007) 23 JCL 201
- Virgo, G., *The Principles of the Law of Restitution*, (2nd edn Oxford University Press, New York 2006)
- Virgo, G., "The Role of Fault in the Law of Restitution", In: Burrows, A. and Rodger, A. (eds), *Mapping the Law: Essays in Memory of Peter Birks* (Oxford University Press, Oxford; New York 2006)
- Wallace, D., *Hudson's Building and Engineering Contracts, Vol 1* (11th edn Sweet & Maxwell, London 1995)
- Weinrib, E., *The Idea of Private Law* (Harvard University Press, Cambridge, Massachusetts 1995)

Function and efficiency of the anticipated recording in China's real estate transactions

Bin LI¹

¹School of Law,
Harbin Institute of Technology, Harbin
China

Email: libienvenue@gmail.com

Abstract:

Anticipated recording is for the first time provided for by China's Property Law 2007. It is further detailed by the House Record Measures 2008 adopted by the Ministry of Construction and the Land Record Measures 2008 adopted by the Ministry of Land and Resources. The purpose of the paper is to analyze the anticipated recording system with regard to its function and efficiency in protecting real estate purchasers' rights and interests. This paper will explain the purpose of anticipated recording, its scope of application, legal effect, and mechanics, etc. Considering that anticipated sale of commercial residential housing is a wide and important transaction in China's real estate market, and that the Property Law 2007 allows the mortgage of construction-in-progress, this paper concludes that the anticipated recording system should be further ameliorated in several important aspects in order to be more efficient for realizing its purpose.

Keywords:

Anticipated recording, Property Law, China,

1 Introduction

China's Property Law 2007 lays down the legal regime of real estate recording¹. From Chinese civil lawyers' dominant point of view, recording system based on the registry is not only a means of publicity, but also a substantial requirement on the valid transfer of real property (Wang, 2007, p.282). The Property Law 2007 confers on recording more important legal value than just evidential function. It clearly provides that the possessor of a real property and its title's content are defined on the basis of registry². A possessor can only prove its title by the certificate issued by recording agency in accordance with the items contained in registry; when certificate conflicts with the items in the registry, the latter prevails except that errors in the registry are proved by evidence³. In the case of transfer of real property, although China's Property Law 2007 does not follow the German law doctrine of *Trennungsprinzip* -i.e., the principle of separation, according to which sale and disposition are two distinct contracts (Von Bar, 2004, p. 332)-, recording

¹ Property Law 2007, Chapter 2, Section 1.

² Property Law 2007, Art. 16.

³ Property Law 2007, Art. 17.

is an independent step from the conclusion of contract stipulating the transfer of property (Guo, 2007, p.170). Under Chinese law, the transfer of real property requires, beyond the conclusion of a sale contract, the modification of corresponding information in the registry before recording agency. Recording is mandatory and prerequisite to the transfer of real property, while physical occupation alone can not justify the title.

It is evident that recording agency intervenes and plays a vital role in the transfer of real property. But the Property Law 2007 doesn't provide for clear and operative rules on those issues, such as who is the recording agency within the current administrative body system, how the recording agency deals with modification of registry and according to what procedures⁴. The gap is to a large extent filled by the Houses Record Measures adopted by the Ministry of Construction in 2008 and the Land Record Measures adopted by the Ministry of Land and Resources in 2008. In so far as property titles relating to houses are concerned, there are four types of recording including, the recording of ownership title, the recording of mortgagor's right, the recording of support right and the anticipated recording of legal titles relating to houses and other real estates. This article analyzes mainly the anticipated recording, which is a special feature of China's real estates recording regime, intended to solve practical problems in real estate market.

2 The Function of Anticipated Recording

2.1 Scope of Application

Article 20 of the Property Law 2007 provides for the general rule on anticipated recording, in terms of which the parties of a sale contract of houses or other real estates could stipulate in their contract that anticipated recording is to be made in order to ensure the future realization of property titles over the houses or other real estates. The wording "future realization of property titles" is a key element to understand the scope of application of anticipated recording. The recording is described as "anticipated", because its legal effect on recorded property titles is deferred to some time in future. Or in other words, the act of recording anticipates the legal effect that it gives.

In practice, anticipated recording is often applied to houses or other real estates under construction. In accordance with Art.30 of Property Law 2007, property titles over real estate under construction come into being only at the moment when the construction work is actually completed. Its rationale lies in China's civil law doctrine refusing to admit buildings and other facilities under construction as forming an object of property right: they are regarded as construction work in progress instead of a complete and independent property (Liang, 1998, p.48); the same is true for the house units composing the construction work. Correspondingly, under Chinese law, the sale of house units under construction is defined as presale contract of residential houses in order to characterize the special nature of the transaction by contrast to normal real estate sale contracts. Urban Real Property Administration Law, promulgated in 1994 and revised in 2007⁵, lays down a special legal regime on presale contract of residential houses. Compared with other real estate transactions, presale contract of residential

⁴ Property Law 2007, Art. 20.

⁵ Urban Real Property Administration Law, Art. 45, Art. 46.

houses under construction is subjected to more governmental regulations⁶. For anticipated recording to be applicable, there shall be an agreement between transferor and transferee to submit the house unit as an individual part of the construction work to anticipated recording. That agreement is often contained in the presale contract of residential houses. Similarly, through contractual arrangements between project owner and constructor, a whole construction work in progress can be subjected to anticipated recording, in view of ensuring the possession of property by the project owner in the future when the construction work is completed.

The anticipated recording regime provided for by the Property Law 2007 applies both to buildings and land. In accordance with Art. 20 of the Property Law 2007, parties to the sale contract of houses *and other real estates* may apply for anticipated recording in view of ensuring the possession of property titles in the future. It is worth of noting that anticipated recording applies also to property titles over land, particularly land use rights. The general rules on anticipated recording of land use rights and buildings are the same. However, detailed rules on anticipated recording of land use rights are contained in the Land Record Measures 2008⁷. Moreover, mortgage is considered under Chinese law as a kind of restricted property right (Jiang, 2007, p.393). The Property Law 2007 permits that houses, houses under construction, land use rights, as well as construction work can be legally mortgaged⁸. Consequently, anticipated recording can also be used for ensuring the future possession of mortgaged property by mortgagor under certain conditions stipulated by contract.

2.2 The Purpose of Anticipated Recording: Ensure the Possession of Property Titles in the Future

Article 20 of the Property Law 2007 is somewhat misleading for the expression of “ensure the future realization of property titles”. Firstly, the word “ensure” does not stand for that anticipated recording is a legal arrangement like insurance or mortgage. It means that anticipated recording confers, by its legal effect, to the creditor of a contract superior status in getting possession of legal titles due to that party. Similarly, “future realization” means that legal titles over houses or other real estates will be possessed in the future when the conditions stipulated by contract or defined by law are met. For example, in the case of presale of houses under construction, the transferee of presale contract will obtain its ownership title of the house when the construction work is completed. The word “presale” demonstrates that ownership title is not transferred at the moment of formation of sale contract, but deferred to the completion of construction work. In fact, at the moment when presale contract is concluded, the object of transaction is still uncertain due to ongoing character of the construction (Gao et al. 2006, p.219). As a consequence, “ensure the future realization of property titles” is an expression of the purpose of anticipated recording: it aims at protecting the creditor of a contract from impediments to the possession of property titles, such as ownership and mortgage right, in the future.

⁶ Ministry of Construction, Administration Measures on Presale of Commercial Residential Houses, adopted in 1994 and revised in 2004.

⁷ Land Record Measures, Art. 62.

⁸ Property Law 2007, Art. 180.

2.3 The Effect of Anticipated Recording

To analyze the effect of anticipated recording, it is worth pointing out the practical reason for anticipated recording. For example, in the case of a presale contract of houses under construction and in the absence of anticipated recording, the transferee could not normally possess property titles at the moment of conclusion of the sale contract. Once the construction work is completed, the transferor and transferee shall file recording. Only when the relevant information is recorded in the registry, could the transferee effectively possess the expected property title. However, it is possible that the transferor sell the same houses under construction twice or even more, the first transferee's possession of property title would be obstructed. In practice, it is not rare that transferor mortgages the houses under construction to a third person -often banks, for getting loans (Liang, 2006). As a result, at the moment when construction is completed, the transferor refuses to or can no longer file recording with the first transferee. Although the transferee of presale contract could claim damages for breaking of presale contract by the transferor, he can not claim any remedy relating to property, for example, return of property, without having filed recording after the construction work is completed. Anticipated recording is a solution targeting on the above situation, in view of giving special protection to transferee of presale contract against transferor who purports to transfer or mortgage wholly or partially the same house or other real estates under construction to a third party. In this regard, Property Law 2007 defines the legal effect of anticipated recording as the following: after a sale contract having been concluded and anticipated recording having been filed, if the same property or construction work is transferred to a third party without prior consent of the creditor, the transaction can not have any legal effect on the title of property. However, the wording of "can not have any legal effect on the title of property" deserves some explanations, which reflects the special feature of China's anticipated recording system.

While in some countries, like the U.S., recording system has the effect of giving recorded transferee priority over non recorded transfers (Bernhardt et al. 2000, p.327), China's Property Law 2007 does not clearly indicate the priority of recorded transferees against third party. In providing that transfer which takes place after the anticipated recording can not have any legal effect on the title of property, Art.20 of the Property Law 2007 intends to impose a constraint on the transferor, which prevents the latter from effectively transferring the property to a third party following the anticipated recording. It seems that the negation of legal effect of subsequent transfer of property had the same function of giving priority to recorded transferee. Nonetheless, there are still some important nuances. On one hand, the anticipated recording does not confer any property title on transferee; it is possessed only in the future, for example, when the construction work is completed in the case of presale contract of residential houses. Consequently, what is recorded at the moment of anticipated recording is transferee's contractual right, which is however specially protected by law by analogy as property right (Liang, 2006). If transferor resells or mortgages the construction work, the contract of resale or mortgage are not void *per se* with regard to Property Law 2007. In the case where both recorded transferee and third party bring claims against the transferor resulting from the concurrent transfers, the recorded transferee will prevail. In fact, the third party has only contractual claim, while the recorded transferee is entitled to somewhat property claim on the basis of anticipated recording. The competing claims between third party and recorded transferee is settled on the basis of China's civil law

doctrine, in according to which property right prevails over contractual right, so do the property claim over contractual claim (Guo, 2007, p.157). By means of Chinese civil law doctrine “property claim prevails over contractual claim”, recorded transferee is given a superior status. On the other hand, however, Art.20 of the Property Law 2007 deals only with the situation where transferor resells or mortgages the property for a second or more times following anticipated recording without transferee’s prior consent, but it says nothing on how to deal with the situation where two or more sale contracts have been successively concluded by the same transferor with different transferees, while anticipated recording is filed for a subsequent sale contract. A literature interpretation of Art.20 can not afford useful guide, given that the article itself deals only with subsequent transfers following anticipated recording. It is then hard to say that the subsequent recorded transferee has a preferential status in relation to the foregoing transferees in possessing property titles. In this regard, the legal implication of Art.20 differs very much from the recording system which expressly gives priority to recorded transferee.

2.4 Mechanics of Anticipated Recording

General rules of anticipated recording are set down by the Property Law 2007, while the Houses Record Measures 2008 and the Land Record Measures 2008 are far more detailed on the mechanics of anticipated recording. Given that the mechanics of anticipated recording of houses and of land are very similar, this article will focus on the mechanics of anticipated recording of property titles relating to houses.

2.4.1 Types of anticipated recording

Art.20 of the Property Law 2007 does not distinguish different types of anticipated recording, while Art.67 of the Houses Record Measures provides for four types of anticipated recording respectively applied to four kinds of real estate transaction, including presale of commercial residential houses; mortgage of presold commercial residential houses, i.e., transferee of presale contract of residential houses mortgages the presold houses; transfer of ownership or of the mortgage of houses; and other cases defined by laws and regulations. Consequently, there are anticipated recording of ownership title in the case of presold commercial residential houses⁹, the anticipated recording of mortgage of presold commercial residential houses¹⁰, the anticipated recording of the transfer of ownership title of houses¹¹, and anticipated recording of mortgage of houses¹². Note, however, that the above four types could be divided into two groups, the first two types belong to the category of anticipated recordings of property titles over houses and other real estates under construction. The last two types belong to the category of anticipated recording of ownership title and mortgage of existing houses and other real estates. The wording of Art.20 of the Property Law 2007 “ensure the future realization of legal titles” is interpreted as covering both houses and other real estates under construction and those existing, although practically, the main objective of anticipated recording is to protect the transferee of presale contract of commercial residential houses (Cheng, 2008). However, it is excluded from anticipated recording the transfer of ownership title of presold commercial residential houses and

⁹ Houses Record Measures, Art. 70.

¹⁰ Houses Record Measures, Art. 71.

¹¹ Houses Record Measures, Art. 72.

¹² Houses Record Measures, Art. 73.

the transfer of mortgage of presold commercial residential houses (Wang, 2009). Although those transactions relate also to the future realization of property titles, the transactions themselves are strictly controlled by governmental regulations on national level.¹³ Consequently, a chain of presale contracts is thus broken, the main purpose of which is to restrict real estate speculations.

The Houses Record Measures distinguish three consecutive steps in the procedure of anticipated recording. The first step is filing, the second is examination by the recording agency and the third step is the decision-making on registration.

2.4.2 Recording Agency

The Property Law 2007 is not clear on the issue, i.e., which governmental department is designated as recording agency. Art.4 of the Houses Record Measures provides that houses recording agencies are either governmental departments in charge of construction of the direct-controlled municipality, municipality and district or the bodies established by those governmental departments for recording purpose. Although Art.10 of the Property Law 2007 lays down the principle that state establishes a uniform real estate recording system, the same article refers to other laws and regulations to enact detailed rules on the issue. In fact, beyond Houses Record Measures, land recording is separately governed by Land Record Measures adopted by the Ministry of Land and Resources in 2007 and entered into effect in 2008. The Land Record Measures precise that land recording is made before district or higher government's department in charge of land and resource administration. As a result, houses recording and land recording are separately governed by two different governmental departments and the two recording procedures are independent one with the other. Whether or not the two recording procedures shall be combined into one is an open question at the current stage of Chinese legislation (Liang et al., 2009). Some local governments manage to combine the two recording procedures into one operation through recording agencies' reform, but the practice is not well established on national level (Jin, 2009).

2.4.3 Filing of Application

The filing is made by both parties of the real estate transaction, with the exception that the transferee of a presale contract can unilaterally file the recording application when the transferor refuses to file in accordance with the stipulation of the presale contract¹⁴. It is evident that the above exception is intended to protect transferee's interests. Note, however, that transferee's unilateral filing of anticipated recording application is not mandatory, because whether or not transferee could unilaterally file the application depends primarily on the fact that anticipated recording has been agreed upon by the two parties in the presale contract. If there is not a clause in the presale contract on anticipated recording, transferee has no way to file the anticipated recording application.

There are documents that shall be submitted to the recording agency in all types of recording: they are application form for recording, identity documents of applicants, evidence of the applicants' agreement on anticipated recording, and other necessary

¹³ Opinions on the stabilization of residential houses' price, adopted on May 30th, 2005, by Ministry of Construction, Ministry of Finance, Ministry of Land and Resources, People's Bank of China, State Administration of Taxation, China Banking Regulatory Commission.

¹⁴ Houses Record Measures, Art. 69.

documents required by the recording agency. Besides the abovementioned common documents, different types of anticipated recording requires some special documents to be submitted. In the case of anticipated recording of ownership title of presold commercial residential houses, there should be presale contract which has already been duly recorded in accordance with Art.44, (4) of the Urban Real Property Administration Law. In the case of anticipated recording of mortgage of presold commercial residential houses, there should be mortgage instrument, presale contract and the certificate of anticipated recording of the presold commercial residential houses. In the case of anticipated recording of the transfer of ownership title of house, there should be the sale contract of the house and the title certificate of the house to be sold. In the case of anticipated recording of the mortgage of house, there should be mortgage instrument, the contract on the secured debts, and the title certificate of the house or the certificate of anticipated recording of ownership title of the existing house. In the last case, anticipated recording of mortgage could be filed following the anticipated recording of transfer of ownership title of the house. However, it is arguable whether the successive anticipated recordings -the first for transfer of ownership, the second for mortgage- can be filed on an existing house. Negative opinion considers that the transferee of an existing house under anticipated recording does not possess ownership title in a strict sense, then he or she cannot proceed to mortgage the house in accordance with Art.67 of the Houses Record Measures (Jin, 2008). The issue shall be clarified by competent governmental department in order to guide the practice.

2.4.4 Examination

The second step that follows the filing of documents is recording agency's examination on recording application. Art.12 of the Property Law 2007 confers recording agency such powers as verifying the authenticity of documents filed by applicants, enquiring the applicants, requiring applicants to submit supplementary materials, as well as leading field investigation. Art.18 of the Houses Record Measures indicates that recording agency shall verify the existence of mutual agreement between applicants over recording, the consent of current recorded title holder on the modification of recorded items. Note that China's recording regime determines the valid transfer of real property, by contrast to French civil law where recording can only make the concerned transaction binding against third party (Cornu, 2007, p.296). Correspondingly, recording agency in China addresses the legality of transaction relating to property titles. For example, in the case of transfer of house, recording agency examines the validity of the sale contract which is considered to be the legal basis of title transfer (Chang, 2009). Art.20 of the Houses Record Measures 2008 precise the issues that shall be examined by recording agency, including applicants' identity which shall be attested by the filed documents; the conformity of the house to be recorded with the urban plan for the first-time recording -i.e., for those newly-built houses having never been recorded before-, or the consistency of the house with existing information in the registry for subsequent recording applications following the first-time recording; the consistency between items to be recorded and facts attested by the submitted evidential materials. Moreover, recording agency shall verify that the items to be recorded do not enter in to conflict with existing property titles recorded in the registry, and there is not any other situations for which recording agency could refuse registration. By contrast to common law practices, for example, in the U.S. recording system where title insurance performs a function of insuring against a wide variety of off-record risks (Malloy et al. 2003, p.

482), China has not such insurance. It seems that recording agency's examination on the legality and validity of real estate transactions could ensure title possessor's security. That function is even consolidated by recording agency's responsibility of compensation for errors made in the recording process, including errors made in examination¹⁵.

The Houses Record Measures 2008 have no special provision on the issues to be examined by the recording agency for anticipated recording. However, there are two points worthy of pointing out. Firstly, due to the fact that anticipated recording application is filed on the basis of mutual agreement between applicants, recording agency would normally pay much attention in verifying the existence and validity of the agreement, particularly when one party unilaterally files the application in the case that the other party breaches its contractual obligations. Secondly, as far as anticipated recording of ownership title or mortgage of houses or other real estates under construction is concerned, whether or not the contract between the parties, particularly the presale contract of commercial residential houses, is in conformity with state's and local government's regulatory measures is a key issue to be examined by recording agency.

2.4.5 *Decisions on Registration*

Art.22 of the Houses Record Measures 2008 provides for the reasons by which recording agency shall decide to refuse registration. Precisely, recording agency shall refuse registration where, houses or other real estates have been built without obtaining in advance urban plan certificate and construction certificate, or construction work has been done in contradiction with the requirements set down by urban plan certificate; applicant can not submit legal and valid materials proving the origin of its title or house titles to be recorded are inconsistent with submitted evidential materials; items to be recorded are in conflict with information recorded in the registry; the house to be recorded is not specific or have no independent value of utilization; the house to be recorded has already been expropriated or confiscated. In an open-ended manner, Art.22 requires recording agency to refuse registration for other situations defined by law, regulations and the Houses Record Measures themselves.

Houses Record Measures 2008 requires that recording agency shall make its decision within 10 working days from the date where filed documents are received. Moreover, in accordance with Art.22 of the Houses Record Measures 2008, recording agency that refuses registration shall notify the applicants its decision in writing and specify the reasons for refusal. However, neither the Property Law 2007 nor the Houses Record Measures 2008 clearly provides for legal remedies against recording agency's decision. It is widely recognized that judicial review should be available to applicants, and administrative litigation is an appropriate way to redress recording agency's illegal activities in recording process (Chang, 2010). Note, however, that Art.20 of the Property Law 2007 and Art.92 of the Houses Record Measures 2008 clearly lay down the principle that recording agency and its staff shall be liability for compensation, when their violation of legal rules, particularly, the Houses Record Measures, causes damages

¹⁵ Property Law 2007, Art. 21, paragraph 2; Houses Record Measures 2008, Art. 92.

to others: it is the recording agency that shall firstly pay the compensation and then it has recourse to its staff having made intentional faults or serious negligence.

2.4.6 Issuance of Certificates

Once recording agency decides to accept recording application, it shall record relevant items in the house registry and then issue corresponding title certificates. In accordance with Art.24 of the Houses Record Measures 2008, house registry shall contain information on physical character of the recorded house, the titles relating to it, and other items that shall be recorded as required by law and regulation. Art.25 of the Houses Record Measures 2008 provides that recording agency shall issue title certificate to its possessor in conformity with information recorded in the registry. There are different types of certificate, including ownership title certificate and other property titles certificates, such as usufruct certificate, mortgage certificate, etc. In the case of anticipated recording, after relevant items having been recorded in the registry, recording agency shall issue corresponding certificates, although the recorded title will be possessed only when the construction work has been completed. In a strict sense, the certificate issued under anticipated recording is not a title certificate, but a recording certificate. The difference is shown by the wording of Art.25 of the Houses Record Measures.

As far as the relationship between title or recording certificate and the registry in terms of their legal value, Art.26 of the Houses Record Measures 2008 lays down the principle according to which the registry prevails in case of inconsistency between certificate's content and items in the registry, with the exception that if there is any evidence proving that items in the registry have been recorded by mistake. However, the operation of that exception would be difficult, given the provision's abstract and ambiguous character. It is not clear on the question how opposition could be raised against recorded items in the registry, by whom and when. It seems that recording agency would exercise a very large discretionary power in determining the issue.

2.4.7 Expiry of Anticipated Registration

As explained before, anticipated recording has no function of creating property titles, but aims at ensure their future realization. Consequently, anticipated recording is provisional and shall be followed by formal title recording procedures. Art.20, paragraph 2, of the Property Law 2007 provides that anticipated recording will be invalid either after three months starting from the time when recording of property titles become possible or when the debts relating to anticipated recording cease to exist. For example, at the time when the construction work has been completed and there had been an anticipated recording of the presold house unit, transferee and transferor of the presale contract shall apply for ownership title recording within three months starting from the moment of completion of the construction work. Given that transferor may refuse to cooperate with transferee to file the recording of ownership title, it is necessary to permit transferee to unilaterally file recording in view of protecting transferee's interests. Nevertheless, neither the Property Law 2007 nor the Houses Record Measures 2008 are clear with regard to unilateral filing. By Art.68, paragraph 2, of the Houses Record Measures 2008, when for formal title recording is filed within the aforementioned time limit, recording agency shall record information in the registry in accordance with items contained in the anticipated registration. It is not clear whether the recording shall

be jointly filed by transferee and transferor, or transferee of the presale contract could unilaterally file. Some local governments require joint filing. For example, the Houses and Land Record Measures adopted by the Shanghai Municipality in 2008 deal with the issue through a two-steps arrangement: in the case of presale contract of houses under construction, once the construction work is completed, there shall be a first-time recording of ownership title -which has the effect of conferring ownership title of the real estate to the transferor of presale contract-, then it is followed by the ownership title transfer recording jointly filed by transferee and transferor of the presale contract of house unit¹⁶. However, if transferor refuses to cooperate with transferee in filing ownership title transfer recording as stipulated by the presale contract, it seems that transferee of presale contract can only have remedies against breaches of presale contract by transferor. As a result, anticipated recording is hardly effective in protecting transferee's interests with regard to its purpose as described by the Property Law 2007, i.e., to ensure the future realization of property titles. Consequently, a useful solution would be to substitute the expiry of anticipated recording with a new mechanism in according to which anticipated recording would be automatically transformed into formal title recording, except that the parties agree otherwise (Xu, 2009).

3 The Efficiency of Anticipated Recording

The efficiency of anticipated recording relates to the regime's capacity in realizing its purpose, i.e., to ensure the future realization of property titles. Some of its drawbacks shall be cured in order to better answer practical needs.

3.1 Inexistence of National Registry

The Property Law 2007 only lays down general principles on real property recording, including anticipated recording. However, until now China has not established a national property registry. It is up to governments of provincial level to enact detailed rules and measures on recording issues. Correspondingly, there is not a single national registry, but registries with in every territory of provinces, autonomous regions and direct-controlled Municipalities, established and managed by different recording agencies. The Houses Record Measures 2008 require that recording shall be filed at the place where the house is located. Consequently, every person intending to acquire an interest in house shall first search the registry before local recording agency in order to determine its priority. In the absence of cooperation between different provincial recording agencies, particularly with regard to the communication of recorded information, trans-provincial inquiry is cumbersome given that physical displacement from one province to another would increase transaction costs for dealers from different provinces. A more crucial issue relating to trans-provincial inquiry is the discrepancy among local recording measures. It is possible that dealer from one province may not understand exactly the recording measures of another province where is situated the house under transaction. The peculiarity of local rules and measures could be a pitfall to "aliens". Much due diligence should thus be done, in order to respect the legality during real property transactions. For example, although the sale and mortgage of presold commercial residential houses under construction are officially prohibited at national level, they are in principle permitted by the Houses and Land Record Measures of the

¹⁶ Houses and Land Record Measures of the Shanghai Municipality, Art.59.

Shanghai Municipality, with some exceptions which however are not clearly defined¹⁷. As a result, transferee of presale contract of commercial residential house under construction could transfer -through sale or mortgage- the presold house under construction, and file anticipated recording jointly with the subsequent transferee or mortgagee before the Recording Office of Shanghai Municipality. However, for Shanghai's local investors who intend to enter into presale contract of houses under construction but located outside Shanghai, they should make due diligence on whether recording measures of other municipalities or provinces are in the same position of Shanghai Municipality.

3.2 Separation of Houses Record and Land Record

Before the promulgation of the Property Law 2007, it was possible to own land but not the buildings due to the separate land and premises title certificate system (Alsen, 1996). The situation is put to an end by Art. 146 and Art.147 of the Property Law 2007, which are considered to affirm the integrated principle, i.e., land use right and ownership title of premises on the land shall be transferred together (Gao et al. 2006, p. 82). However, houses record and land record keep being two separate recording mechanics. As a result, in the case of transfer of house, transferor and transferee shall file two recordings, one concerning house ownership title, the other on land use right. However, in the case of presale contract of houses, there is not any clear requirement that transferee and transferor of the presale contract shall also file anticipated recording of land use right, besides the anticipated recording of houses under construction. In other words, anticipated recording of presold house will not necessarily be effectuated together with anticipated recording of land use right, although the future realization of house ownership title by the transferee will result in the transfer of land use right, in accordance with the integrated principle. In fact, as an anticipated recording aims at ensuring the future realization of property right title, rather than creating or proving property title, and it has a function of "notice" against third parties, simultaneous anticipated recordings of house and land use right seem to be unnecessary. However, an anticipated recording of presold house without anticipated recording of transfer of land use right would give rise to some difficult situations. For example, transferor of presale contract is restricted to transfer the house under construction to a third party, but he or she is not similarly restricted to transfer the land use right; if the transferor of presale contract of house transfers the land use right to a third party, the validity of the contract for transfer of land use right could be a debatable issue. In the above case, legal effect of anticipated recording shall be interpreted in a manner that legal interests of transferee under presale contract can be sufficiently protected: the anticipated recording of presold house has the same function of notice against third parties with regard to not only transfer of the house but also transfer of the land use right. From transferee's point of view, it will be desirable that anticipated recording of presold house and anticipated recording of transfer of land use right be filed together, because uncertainties resulting from legal provisions' interpretation would be thus effectively avoided.

3.3 Voluntary Nature of Anticipated Recording

Anticipated recording was intended to protect transferee of presale contract of houses under construction, who is a weaker party compared to transferors taking into account

¹⁷ Houses and Land Record Measures of the Shanghai Municipality, Art.51, Art.54, paragraph 3.

China's real market's situation (Liang, 2006). However, contents of legal provisions governing anticipated recording seem to contradict the aforementioned purpose. Although anticipated recording has its function as a "notice" against third parties, the question whether or not parties of presale contract engage in anticipated recording depends on their mutual agreement. Only when anticipated recording is stipulated in the presale contract that bears transferor an obligation to cooperate with transferee, otherwise, transferee could not unilaterally file anticipated recording. In other words, anticipated recording is not mandatory but results from the contractual arrangement between parties. Consequently, the usage of anticipated recording is determined by the parties' negotiating power. In the case of presale contract of house under construction which is generally concluded between individuals and real estate developer as transferor, it is nearly impossible for individual transferee to make transferor accept anticipated recording arrangement in presale contract. In fact, presale contracts of house under construction are usually standard contracts unilaterally drafted by real estate developer as transferor.

On the other hand, in the case of construction contract, project owner and constructor could also stipulate on anticipated recording of the construction work, the purpose of which is to protect project owner's interest against the transfer of construction work by constructor to third parties. However, by contrast to presale contract of house under construction concluded between individual and real estate developer, project owners are more often than not more powerful than constructors. In practice, many disputes have arisen from project owner's breaches of contract for non-payment or late payment. Remedies available to constructors include, among other things, transfer of the construction work to a third party by means of auction in accordance with Art.286 of the Contract Law 1999 and then constructor gets preferential payment from the value of sale of the construction work (Cui, 2007, p.406). Nevertheless, anticipated recording - accepted by constructor as project owner insists upon it - could prevent the constructor from exercising its right to preferential payment through transfer of the construction work, due to anticipated recording's function in prohibiting the transfer of construction work in progress without project owner's prior consent. Under current Chinese law, it is not clear whether or not legal effect of anticipated recording could be overthrown by constructor's legal claims against breaches of contract by project owner, particularly on payment issues (Sun, 2009, p.351). If constructors can not transfer the construction work even though project owner refuses to perform its obligation of payment, anticipated recording would become an impediment to the practical need of striking a balance between project owner and constructor's interests. Some exceptions shall be given to legal effect of anticipated recording, taking into account particular situations in different kinds of transaction.

3.4 Recording Agency's Examining Power

Given that recording is a prerequisite on valid transfer of property, recording agency exercises a very important power in addressing issue of validity of the transfer contract between parties. Art.22 of the Houses Record Measures provides for a rather wide legal basis on which recording agency could refuse registration. As a principle, in filing anticipated recording, applicants are required to submit legal and valid documents to justify the origin of the property titles to be recorded. For example, the contract of sale of house shall be subjected to examination by recording agency, when purchaser and

vendor jointly file documents for purpose of recording the transfer of ownership title. According to the Houses Record Measures 2008, recording agency will not only verify the existence of sale contract, but also make evaluation on its legality and validity¹⁸, while such issues shall be however appropriately dealt with by tribunals. In fact, recording agency's scope of examination is not well defined and it has much flexibility on the issue (Jiang, 2007, p.321). It is possible that recording agency refuses registration, by reason that the transaction is unlawful and invalid. Such a decision means that recording agency refuses to give effect to the transfer of ownership, but also makes the sale contract void. As real estate transactions are more and more subjected to governmental regulations, the compatibility of contracts leading to transfer of property titles with governmental regulations fall within the scope of examination by recording agency. As a consequence, recording agency plays the role of governmental regulations' enforcement organ and exercise somewhat a police power over real estate transactions. Some local recording measures contain ambiguous conditions on recording, especially on anticipated recording, thus conferring recording agency with rather wide discretion in making decisions on registration. For example, the Houses and Land Record Measures of the Shanghai Municipality requires that recording agency should refuse to admit anticipated recording of subsequent transfers of commercial residential houses under construction following conclusion of presale contract, if the commercial residential houses under construction have not yet been recorded, or *subsequent transfers of commercial residential houses under construction is not in conformity with state's or Shanghai Municipality's regulations thereof*¹⁹. It means that all the conditions on anticipated recording are not completely specified in the current Houses and Land Record Measures of the Shanghai Municipality; they could be determined by reference to other regulations that maybe adopted in the future. As a result, modifications of governmental regulations at some moment in the future can render unlawful and void subsequent transfers of commercial residential houses under construction following conclusion of presale contract, while such subsequent transfers were legal in the past and keep being legal at current stage. The open nature of recording agency's examining power, even more reinforced by its power of adopting recording measures, is thus a serious menace to the predictability and legal security of real estate transactions.

4 Conclusion

Anticipated recording in the Property Law 2007 is intended to protect transferee's interest in possessing property title in the future. However, anticipated recording mechanics established by administrative measures at both national and local levels have some important drawbacks that undermine its efficiency with regard to the realization of its purpose. This article's analysis on issues relating to anticipated recording shows that the Property Law 2007 is just a first and primary step of legislation on China's real estate market, the healthy development of which depends on the forthcoming, more reasonable and operative legal rules.

¹⁸ Houses Record Measures, Art.22, (2).

¹⁹ Houses and Land Record Measures of the Shanghai Municipality, Art. 53, paragraph 3.

5 References

- Alsen, Jonas. (1996), 'An Introduction to Chinese Property Law', *Maryland Journal of International Law and Trade*, Spring 1996, p 54.
- Bernhardt, Roger H., Burkhardt Ann M. (2000), *Real Property*, 4th Edition, Law Press China, Beijing.
- Chang, Peng'ao. (2009), 'The Recording Procedure of Houses Record Measures (Fangwu dengji banfa zhong de dengji chengxu)', *Journal of Jinan University (Ji'nan xuebao)*, Issue 31, p 7.
- Chang, Peng'ao. (2010), 'Principle and Prospective of Real Estate Recoding (Budongchan dengji de lifa dingwei yu zhanwang)', *Legal Science (Faxue)*, Issue 3, 2010, p 107.
- Cheng, Xiao. (2008), "Issues of Anticipated Recording in the Houses Record Measures (Fangwu dengji banfa zhong de yugao dengji wenti)", *China Real Estate (Zhongguo fangdichan)*, Issue 330, p 17.
- Cornu, Gérard. (2007), *Droit Civil, Les biens*, 13th Edition, Montchrestien, Paris.
- Cui, Jianyuan. (2007), *Contract Law (Hetongfa)*, 3rd Edition, China Law Press, Beijing.
- Gao, Fuping., Huang Wushuang. (2006), *Real Property Law (Fangdichan faxue)*, 2nd Edition, Higher Education Press (Gaodeng jiaoyu chubanshe), Beijing.
- Guo, Mingrui. (2007), *Civil Law (Min Fa)*, 3rd Edition, Higher Education Press, Beijing.
- Jiang, Ping. (2007), *Civil Law (Minfaxue)*, China University of Political Science and Law Press (Zhongguo zhengfadaxue chubanshe), Beijing.
- Jin, Shaoda. (2008), 'Can Anticipated Recording of Mortgage be Done on the Basis of Anticipated Recording of Transfer of Existing House's Ownership Title (Xianfang maimai shifou keyi ping yugao dengji zhengming banli diyaquan yugao dengji)?', *China Real Estate (Zhongguo fangdichan)*, Issue 334, p 28.
- Jin, Shaoda. (2009), 'An Ideal Model of Real Estate Recording (Budongchan Dengji jigou de lixiang moshi)', *China Real Estate (Zhongguo fangdichan)*, Issue 343, p 45.
- Liang, Huixing. (2006), 'Debates on several legal issues of the draft property law (Guanyu wuquanfa caoan jige falv wenti de shangque)', *Worker's Daily (Gongren Ribao)*, September 18, 2006, p 7.
- Liang, Huixing. (1998), *Studies on Chinese Property Law (Zhongguo wuquanfa yanjiu)*, China Law Press, Beijing.
- Liang, Yarong., Wang, Chongmin. (2009), 'Analysis about Real Estate Register Agency Setting (Budongchan dengji jigou shezhi tanxi)', *Legal Forum (Faxue luntan)*, Issue 1, p 133.
- Malloy, Robin Paul., Smith, James Charles. (2003), *Real Estate Transactions: Problems, Cases, and Materials*, 2nd Edition, CITIC Publishing House, Beijing.
- Randolph, Patrick A., Lou, Jianbo. (2002), *Chinese Real Estate Law*, Kluwer International Law, The Hague.
- Sun, Xianzhong. (2009), *General Principles on Real Rights Law of China (Zhongguo wuquanfa zonglun)*, 2nd Edition, China Law Press, Beijing.
- Xu, Huiying. (2009), 'On the Amendment of Anticipated Recording Regime: Comments on Art.20 of the Property Law (Yugao dengji zhidu de wanshan: ping wuquanfa di ershitiao)', *Essays of Law and Business (Fashang luncong)*, Volume 2, p 75.

- Von Bar, Christian., Drobni, Ulrich. (2004), *The Interaction of Contract Law and Tort and Property Law in Europe: A Comparative Study*, European Law Publishers, München.
- Wang, Dexiang. (2009), 'Can another Anticipated Recording be Made on the Basis of the Already Existed Anticipated Recording ? (Yugao dengji shang nengfou zai sheli yugao dengji)', *China Real Estate (Zhongguo fangdichan)*, Issue 343, p 12.
- Wang, Liming. (2007), *Studies on Property Law (Wuquanfa yanjiu)*, 2nd Edition, China Renmin University Press (Zhongguo renmin daxue chubanshe), Beijing.

Where will European property law go next? Tenancies and property after *Hutten-Czapska v Poland* and *FEANTSA v Slovenia*

Jane Ball
University of Sheffield, School of Law

1. Introduction

Two recent and quite contradictory decisions cast doubt on the coherence of the European rights approach to property: *Hutten-Czapska v Poland*¹, under the European Convention on Human Rights (ECHR)² and *FEANTSA v Slovenia*³ under the Revised European Social Charter (RESC).⁴ These cases concern the property and tenancy law imported into Eastern Europe in the transition from communism. The European Union has insisted that new Eastern European members should adopt both treaties, as a condition of membership as well as developing a market economy (Commission of the European Communities, 1991). Unfortunately, there is no agreement in Western Europe as to the nature of landlord and tenant rights or the economic theory to be applied.

A divergence in the approaches under the ECHR and the RESC have led to heavy duties being imposed on the new member states, without allowing the means whereby these duties can be satisfied. In *Hutten-Czapska*, the Grand Chamber of the European Court of Human Rights (ECtHR) ruled strongly in favour of landlords in Poland recovering their property after the fall of communism. The decision disapproves strong rent control, proposes government restitutionary compensation for landlords and requires an effective eviction law, leaving the responsibility for the tenants affected by this to the state. In contrast, *FEANTSA* is a decision of the European Social Committee (ESC), concerning the RESC. This supports Slovenian tenants who were suffering deteriorating conditions and who did not benefit from the discounted right-to-buy extended to most former tenants. This time the state should protect tenants from eviction and probably pay compensation to tenants who did not enjoy the right to buy. State responsibility and compensation is a feature of both cases, which it might be unable to discharge in the current economic climate.

The parallels between the situations in the two cases are striking, although the outcomes are very different. These cases might be exceptional attempts to create a market in property after the fall of communism. Nevertheless, some problems, such as a high level of home ownership or how to move from a system of strong rent control are an experience shared by long-established members of the European Union.

These cases raise a series of problems, questions: firstly, about the way in which the decisions of the ECtHR and the ESC pose practical problems, particularly limiting the solutions which can be applied by national governments, secondly about the nature of the landlord and tenant law imported, in a context where the basis for these rights is not settled across Europe; and thirdly about the institutional dilemmas of who should be ultimately settling the form of European property and tenancy law, if this should be done at all.

There are conflicting institutional approaches between the ECtHR and the ESC, the problems with which are exacerbated within the European Union for several reasons. This raises wider questions about the lack of agreement by the Council of Europe Courts which increasingly are used as a

¹ *Hutten-Czapska v Poland* (2007) 45 E.H.R.R. 4; 20 B.H.R.C. 49.

² *European Convention for the Protection of Human Rights and Fundamental Freedoms*, 4th November 1950 available at <http://www.coe.int>.

³ *Fédération Européenne des Associations Nationales Travaillant avec les Sans-abri (FEANTSA) v. Slovenia*, Complaint no. 53/2008, decision of 8 December 2009.

⁴ *European Social Charter (Revised)*, Council of Europe, 3rd May 1996, available at the same website.

constitution for the whole of the EU. Interaction between all three institutions is becoming important, particularly since the economic model applied to tenancies is also uncertain. This paper generally raises questions rather than providing answers. The originality of this article lies in its concern for any emerging European-wide property and tenancy law, rather than simply the ECHR and the RESC. The existence of a common European model for these is implicit but does not exist.

This focus is appropriate considering recent acceleration towards harmonization of European law, and the sheer intrusiveness into national law of the both *Hutten-Czapska* and *FEANTSA* decisions. This paper firstly looks at these cases on transition from communism to a 'free' rental market and how similar cases are dealt with differently inconsistent approaches according to the institution concerned. Then the fact that there is no common legal model to apply to these countries by the ECtHR, and the RESC, either in private law or divergent public law norms around property which also affect interpretations under the ECHR. Finally conflicts between the ESC, the ECHR and EU approaches are considered with their consequences for property law and tenancies, including divergent models of the market. This risks interfering with these in incoherent ways hampering housing policy. A recent example is that a recent Competition Commission decision also limits housing policy to assist tenants by limiting subsidy to social housing.

2. The Eastern European States in Transition

The Eastern European countries represented a challenge after the fall of communism after 1989. Sunstein (1991) describes this transition as having three aspects: markets, democracy and constitutionalism. In both Poland and Slovenia, the transition to democracy and constitutionalism is well underway, but the transition to the market in housing is problematic.

An essential element for developing markets is the right to property, under article 1 of Protocol 1 of the ECHR⁵. The approach of the ECtHR to Eastern European countries in transition has been to ease the transition to a private law system by accepting temporary schemes which would normally infringe human rights, whilst 'respecting domestic sensitivities linked to injustices suffered in the past' (Varju, 2009: 3). This could be found in cases such as *Velikova v Bulgaria* legitimizing a scheme within the context of transition which would otherwise fail ECHR criteria,⁶ and in *Jahn and others v Germany*, property rights were denied for people not carrying out an economic activity⁷.

The ECHR will intervene either where sufficient time has elapsed in transition or where fair balance concerning fundamental rights is not respected, such as in *Hutten-Czapska*, and *Broniowski v Poland*⁸. These were dealt with under the pilot procedure under article 46 of the ECHR, allowing the law pronounced to apply to multiple applicants, a very strong response.

This section particularly looks at the cases of *Hutten-Czapska* in Poland and *FEANTSA* in Slovenia in their context. This means first looking at points in common and then how the situation was handled in each case, before considering inconsistencies. There is an overall question about why the ECtHR favoured landlords and the ESC favoured tenants, to the detriment of balance.

2.1. Similarity and Difference in Eastern Europe

⁵ Protocol to the Convention for the Protection of Human Rights and Fundamental Freedoms, Council of Europe, signed 20 March 1952, Paris.

⁶ Para. 170 in *Velikovi and others v. Bulgaria* (41488/98) judgment of 15 March 2003, (commented on in Varju, 2009) and Paras. 92 and 125; *Jahn and others v. Germany* (2006) 42 E.H.R.R. 49 (46720/99) the ECtHR (Grand Chamber), 30 June 2005.

⁷ *Ibid* para. 123-124. Lack of activity relating to food, forestry or agriculture was a justification for effective expropriation.

⁸ *Broniowski v Poland* (2005) 40 E.H.R.R. 21: (31443/96) ECH, (Grand Chamber), 22 June 2004.

Hutten-Czapska and *FEANTSA* both concerned landlord and tenant relations in Eastern European countries after the fall of communism. In both cases people deprived of their property under communism after the War and recovered this during the transition towards a free market, but recovered their properties with sitting tenants. These new landlords sought sufficient rent to pay for repairs, to make a profit and for eviction to be possible. After 1990, there was a serious problem of disrepair in both owned and tenanted homes (Kerestes, 2002, Gromnicka and Zysk 2004). Most residential tenants from the communist era after 1990 were allowed to buy their homes at a large discount. Under Slovenian communism, many people had retained occupancy of their own homes, even though nationalized, not generally the case in Poland.

Both the Polish and Slovenian cases concerned tenants who did not buy their homes and who became tenants in denationalized homes. The transition to a free market meant loss of security and increased rents. The Slovenian complaint concerned tenants who were not given the opportunity to buy their homes because of restitution to former owners. The Polish complaint was from landlords who recovered their homes because the tenant did not exercise the right to buy, due to poverty or other reasons.

In *Hutten-Czapska*, the applicant to the ECtHR was a landlord, effectively representing 100,000 landlords in Poland. This affected 600,000-900,000 tenants of denationalized housing, around 5.2 percent of housing stock. In *FEANTSA*⁹, a Federation of European Homelessness Organizations supported 13,000 tenants or ex-tenants¹⁰ of denationalized housing by collective complaint to the ESC. Poland's population is around 38 million people, whilst Slovenia's is only two million, so the importance for Slovenian housing is more similar than the numbers might suggest¹¹.

Both Poland and Slovenia had effectively high level of homeownership with a small private rental sector. Statistics are unclear about the exact size of the private rental sector in both countries.¹² Poland had 57 percent owner-occupied homes in 2004. The 18 percent of co-operative housing was in transition to something closer to ownership, with around 250,000 sold to occupants in 2006-7 alone¹³. Diverse types of public or social homes were 25 percent of stock but only 12% was recorded as social¹⁴.

Some Polish public or social tenants had very favourable conditions,¹⁵ but other "substitutive" housing was of inferior quality, with short fixed-term tenancies at a low rental. This municipal stock was for vulnerable people suffering poverty and social exclusion. In *Hutten-Czapska*, landlords were often unable to evict, because judges stayed eviction, declaring evictees entitled to substitutive housing and there were long waiting lists.

In Slovenia, around 84 percent of housing stock was owner-occupied in 2004, with only nine percent rental stock. Rented property was of generally low quality, split between firstly free market rentals with high rents and low security of tenure and secondly not-for-profit rentals with indefinite contracts and low rents. Eight percent of the population lived in property owned by parents or relatives, in preference to renting (Cirman, 2010). In 2004, not-for-profit or social housing was six percent of housing stock, a small amount to meet needs.

⁹ *FEANTSA* (European Federation of National Associations Working with the Homeless).

¹⁰ This figure is from *FEANTSA*, whilst the Slovenian government said that this affected 4,700 properties now, of which 2,566 had found housing - para. 6, in *FEANTSA*.

¹¹ Gromnicka and Zysk (2004: 16).

¹² Sendi (1999) and the (Polish) Central Statistical Office (2007). Sendi suggests that figures for private rentals were deliberately suppressed by socialist governments across Europe as contrary to their philosophy.

¹³ Polish Central Statistical Office (2007).

¹⁴ From Minister of Infrastructure of the Italian Republic (MIIR) and Federcasa (2006). It is difficult to characterize public or quasi-public stock as a single type of social housing, because rental stock could be owned by local authorities (gminas), the State Treasury, social building societies, other public entities as well as private companies and individuals.

¹⁵ Gromnicka and Zysk (2004: 17) and Central Statistical Office (2007) showing less than 2 percent of this stock in some cities.

In both Slovenia and Poland, after 1946, “tenants” had been imposed by an administrative decision¹⁶. Tenants imposed in both countries enjoyed a “housing right”, strong residence rights and low rents under communism arising from socialist theory. In 1964, the Polish Civil Code accepted a socialist typology of unity of public ownership. Even when amended in 2001, the new Constitution: ‘... did not differentiate between various types resident, reflecting a common view that this protection should be established for all occupants of premises, regardless of the legal form of a “lawful possession”’. (Gromnicka and Zyst, 2004: 7). In Slovenia, tenants prior to 1991 also enjoyed a “housing right”¹⁷ which *FEANSTA* described as like a property right. .

The ECtHR and ESC faced three conflicting moral demands: The demand for restitution from property owners, the demand for the right to property by the landlords and the demand for the right to housing from tenants. Even allowing for differences in history, economics, culture and policy approaches in Eastern European countries,¹⁸ this is essentially the same legal problem. Sendi (1999: 813) described the Slovenian position:

‘The ‘inherited’ sitting tenants insist on continuing to enjoy the same rights as before (unlimited term of tenure, and government controlled rents) while the new landlords (restituted owners) regard their acquired dwellings as private property which they ought to be able to trade freely on the open market.’

The ECHR faced an almost impossible dilemma in resolving these conflicting demands, problems arising before many people affected were born. People whose property was taken sought reparation or restitution within established principles stated in UN documents¹⁹. Restitution was a principle ‘to restore victims to the condition they would have enjoyed had no violations of their rights occurred’ (Williams, 2007: 6). Possible damage included intangible harms to victims’ mental or physical integrity and dignity.

2.2. *Hutten-Czapska v Poland*

The ECtHR in *Hutten* described the poverty of Poland with 58 percent of people below the poverty line²⁰ and thus the importance of social rights in the Polish constitution during the transition from communisms²¹. The court also emphasised the power of appreciation of national governments in matters of property, although this did not prevent censure of the Polish government for what it described as a “structural and systemic problem” affecting private landlords with pre-1990 tenants²².

Mrs. Hutten-Czapska’s parents’ home was compulsorily let in 1948, becoming part of a “special lease scheme” in 1974. This did not improve her position, and neither did her recognition as legal heir to her parents in 1990. She wished to recover the property, now divided into three flats, for use as the headquarters of a private foundation and to charge a higher rent. She had complained or litigated at every opportunity since 1974.

Poland began the transition to a market economy in land by a 1994 statute relaxing rent control in commercial and residential premises, but this did not apply to older tenancies under the special lease scheme. The legislation commuted these leases into contractual leases signed for an indefinite period

¹⁶ Para. 69 in *Hutten-Czapska*, and para. 39 in *FEANTSA*. In Poland letting could also be by contract, see Gromnicka and Zysk (2004) p.5.

¹⁷ This had additional features, see below p.6.

¹⁸ See, for example, Balchin (1996) Hegedüs, *et al.* (1996) and Pilcher-Milanovich (2001).

¹⁹ ‘Basic Principles and Guidelines on the Right to a Remedy and Reparation for Victims of Gross Violations of International Human Rights Law and Serious Violations of International Humanitarian Law, UN Doc. A/RES/60/147 (March 21, 2006) and preliminary principles in ‘Principles on Housing and Property Restitution for Refugees and Displaced Persons,’ UN Doc. E/CN.4/Sub.2/2005/17 (June 28, 2005), part of general responsibilities to provide individuals with remedies.

²⁰ *Hutten-Czapska*, para. 58.

²¹ Article 75 and 76 of the Polish constitution obliges the government to protect tenants.

²² *Hutten-Czapska*, para. 234.

with the existing tenant and restricted rent. The situation was worsened by the imposition of repair obligations on landlords, without any government incentive or subsidy and with a disadvantageous tax regime. Eviction was still substantially limited because of lack of “substituted” housing. Later acts of 2001 and 2004 were passed to remedy this situation following censure from the national Constitutional Court, but made little difference.²³

Parliament had showed reluctance to remove tenant protection, whilst the Constitutional Court censured this inaction, in major decisions of 2000, 2002 and 2005. In 2000 the Court ruled that the landlords’ obligations were incompatible with the constitutional principles of the protection of property rights and social justice, because, in particular it placed a heavy financial burden on landlords “...which was in no way proportionate to the income from controlled rents”.²⁴ During the ECHR case the rent was higher than originally, but was still around 60 percent of repair costs, so properties were deteriorating.

The ECtHR Grand Chamber substantially confirmed a decision of the lower chamber, which in turn supported the Polish Constitutional Court. The government had consistently expressed a desire for reform, but this did not seem to have democratic support. Owners were deprived of “the slightest substance of their property rights”²⁵. This was not a question of complete deprivation of housing rights, or legality, but a question of proportionality.

The ECtHR does not prescribe the means by which states should satisfy convention requirements, but the inability of the applicant to recover her property by eviction was criticised²⁶ in the absence of mitigating factors of government support for landlords. The inability to raise the rent was censured²⁷, particularly when the government had promised no rent restrictions.

The consequences of raising the rent are serious in a poor country where housing benefits for tenants were described as “practically non-existent”²⁸. Effects can be seen in government statistics. *Hutten-Czapska* records around nine percent of tenants in arrears with their rent in 2003. Government figures for 2007 showed that 45 percent of tenants were in arrears. This may be due to non-eviction, but eviction on that scale would be a threat to public order²⁹, and action could deepen resentment between landlords and tenants.

2.3. FEANTSA v Slovenia

The collective complaint to the ESC in Slovenia is the mirror image of the Polish application to the ECtHR. The complaint was on behalf of tenants and former tenants of denationalized homes claiming breach of their right to housing under article 31 of the ESC. In Slovenia, high levels of effective owner-occupation even under communism meant that only 13,000 homes were affected by restitutionary rights of private owners. As in Poland, the majority of the small rental market comprised public or social stock with very secure tenants, thus probably unavailable to evictees³⁰.

Article 31 of the RESC requires that government should: promote access to decent housing; prevent and reduce homelessness; and make the price of housing accessible. The collective complaint against the Slovenian government included discrimination (article E) and failure to provide social, legal and economic protection for families (article 16). The RESC collective complaints system creates a quasi-

²³ The 2004 act might have provided for a higher rent but this was limited by amendment by the second parliamentary chamber.

²⁴ Para. 84, *Hutten-Czapska*.

²⁵ Para. 200, *ibid*, quoting the lower chamber.

²⁶ Described throughout *Hutten-Czapska*, particularly paragraph 224.

²⁷ See *ibid*, para. 176.

²⁸ *Ibid*, para. 150.

²⁹ A point made by the Grand Chamber at para. 106.

³⁰ See above, section 2.1. for the statistical position.

jurisprudence (Brillat, 2005) or case law about social rights (Quinn, 2005)³¹. This imposes obligations on governments including progress towards the goals of the RESC on a reasonable timetable³².

The ESC decision in *FEANTSA* expressed a concern to comply with relevant provisions of the ECHR, presumably including the right to property. Nonetheless, the ESC cited *Mellacher v Austria*³³, a case which approved rent control on social grounds, rather than *Hutten-Czapska* which did not. In *Hutten-Czapska*, the ECHR had distinguished these two cases, approving *Mellacher* because there were government measures compensating landlords, allowing at least a cost rent³⁴. In *FEANTSA*, Slovenia had cost rents of the type approved in *Mellacher*, but the ESC still disapproved of the cost-rents recoverable from Slovenian tenants as not affordable. The ESC did not consider *Hutten-Czapska* at all, which disapproved rents lower than repair costs, something which would have satisfied the ESC in *FEANTSA*. This is an inconsistency between the courts.

In Slovenia, the government had done what the ECHR had asked the Polish government to do in raising rents to a level which would pay repair, in providing grounds for eviction and limiting inheritance. Former land owners were offered either a discounted price in compensation for the loss of their property, or to recover their property with a tenant. Tenants of denationalized properties owned by private landlords, enjoyed a “housing right” as in Poland, but supplemented by provisions that only cost rents should be charged, by the possibility of a small grant to buy a home elsewhere or the right to purchase their tenanted home if the landlord agreed. These tenants were unable to buy their home because of this requirement for landlords’ agreement and claimed discrimination compared to holders of the housing right in public or social stock. *FEANTSA* had claimed that the government had systematically created a housing problem for people not suffering difficulties before³⁵.

The operative part of the ESC decision is quite short, most of the decision containing the rival assertions of the parties, making it difficult to establish which claims are accepted by the ESC and which not. An example is that *FEANTSA* claims rents were raised by 613% over 12 years³⁶, whilst the government claimed the increase was no more than 128 percent after inflation, or 16.5 percent of average income³⁷ and that this was still only a cost rent³⁸. Both could be true, for example, the Polish below-cost rent at only 60% of repair expenses had still risen by 100% since the beginning of reform when rents were 30% of repair costs³⁹. There was no guarantee that Polish incomes would increase to cover this 60%, even though below cost.

FEANTSA proposed that the 13,000 tenants or former tenants of private restituted property were in a situation of vulnerability⁴⁰, including people who had already lost their homes. The government suggested that more than half tenants with the housing right had found other housing by the purchase and grant arrangements, 2,566 out of 4,700 tenants, a much lower number.

So far as it can be judged, *FEANTSA*’s claim was that tenants had a property-like “housing right” was accepted. This was said to be a usufruct, a conventional Roman law fragmented real-property ownership right, but one which is usually for life and not subject to any rent which does not fit this description⁴¹. The housing right might be more like an emphyteusis, a long real-property type lease

³¹ Something found throughout this collective work edited by De Búrca and de Witte (2005) although authors identify difficulties and inconsistencies.

³² See *FEANTSA*, generally for an exposition of these obligations.

³³ Judgment of 19 December 1989, Series A no. 169.

³⁴ *Hutten-Czapska*, paras, 177, and 224.

³⁵ *FEANTSA*, Case Document No. 3, para. 177.

³⁶ *FEANTSA*, para. 6.

³⁷ *Ibid*, para. 7.

³⁸ *Ibid*, para. 67.

³⁹ *Hutten-Czapska*, para. 14.

⁴⁰ *FEANTSA*, para. 6.

⁴¹ This is the position in France, Ball (2010) but also in Spain and likely to be the case anywhere where the civil law is based on Roman law, although not yet fully researched.

described below⁴². It was argued in support in both Poland and Slovenia that tenants spent their own money improving their homes, so would Polish tenants have property-type rights too?

The Slovenian assertion that the “housing right” was a property-like right, was facilitated by a decision of the Constitutional Court (quoted in the ESC decision), ruling on the time property rights were created. The original Slovenian owners recovering property were treated as acquiring entirely new rights after communism, so the tenants’ rights predated this and would prevail⁴³. In contrast, in Poland the possibility of the tenants having a property-like housing right was not considered. This was supported by the ECtHR’s view that although the court could take into account the Polish position before communism and accession to the ECHR. This position was a continuing difficulty in *Mrs. Hutten-Czapska* recovering her property, so effectively her rights predated the tenants’ housing right⁴⁴. This is inconsistent between the two cases.

The ESC found against the Slovenian government because rents were not affordable (not because they were cost rents) something inherent in article 31, particularly affecting poorer and older tenants who lost both the prospect of buying their home or a reasonable rent. This also caused a breach of obligations to provide family protection. The treatment of tenants in denationalized property was also discriminatory because there was nothing in their situation to justify treating them differently from other tenants who had the right to buy.

The ESC found that a significant number of households in a very precarious positions were prevented from effectively exercising their right to housing caused by insufficient measures of access to a substitute flat, the change in rules on occupancy and rent increases. Amongst other things governments should adopt the necessary legal, financial and operational means of ensuring steady progress towards the goals laid down in the Charter. The result of neither of these cases is unjust towards the individuals, who benefitted from the decisions of the ESC or the RESC, but in this situation, the decisions were always going to be unjust to the losers, and the judgements cause other problems.

2.4. Problems from the Decisions

These decisions in combination limit the scope of national housing policy, in ways that might be mutually inconsistent. The nature of decisions in both the ECHR and the RESC is to disapprove government policies for breach, rather than provide positive support. However, in the current state of European recession, the normal policies of the EU to support the Council of Europe by “aid, association and membership” (Moravcsik, 1995:166) are considerably constrained.

In addition to these limitations on aid, the financial burdens that the ECHR imposed on the Polish government include the €30,000 non-pecuniary damages plus loss of rent suffered by *Mrs. Hutten-Czapska* multiplied by 100,000 other landlords. In addition, rent control should be lifted, so either the increased housing benefit should be paid to cover increased rents, or there is the social cost of helping the evicted. In Slovenia, the government should probably compensate tenants for the lack of a right-to-buy, pay benefits and find a place for the poor and dispossessed. It is not surprising that there is some evidence of non-compliance with the judgement in the high level of rent arrears in Poland, who were apparently not evicted. These heavy burdens on poorer members of the EU do not lead to respect for European law.

⁴² See below p. 10.

⁴³ *FEANTSA*, para. 20.

⁴⁴ A question of *ratione temporis*, *Hutten-Czapska*, paras 147-153. Polish property owners names had stayed on the land registers under communism, not the case in Slovenia, so Polish landlords had some benefits of ownership even though costs of repair exceeded rent. Despite this formal situation, in Poland and Slovenia the substantive deprivation of rights was very similar.

There are implications, too for housing policy across the EU and for the shape of future rights to property and to housing. The ECtHR has ruled against heavy rent control, and in favour of eviction for failure to pay higher rents. This constrains housing policy and the insistence on a market rent and discourages tenures which are not either a short full-rental tenancy or absolute ownership, with narrow exceptions. Common Roman law fragmentary rights might be accepted like the usufruct or emphyteusis, but this limits other diversity or innovation in tenures. The fragmentary property rights in England and other parts of the UK⁴⁵ might be particularly affected.

The Polish fragmentary “housing right” was not accepted, so is there a risk to English long leases, as in *James v the UK*⁴⁶? Here, tenants paid a large sum upfront for a real property right with low fixed rents and the lease was mortgageable. In that case, tenants has a right to renew the lease or buy the freehold, which the ECtHR rather grudgingly saw as a social exception to property law, even though new leaseholder had paid something close to the price of ownership for this since 1967⁴⁷. Long leases of between 18-99 years rather similar to this are possible in France, but the Roman long lease or emphyteusis (below) cannot generally be used for residential tenancies there because of a similar rule to that in *Hutten*. This is because French long leases are not exempt from compulsory break clauses to increase rent (Ball, 2010). There are likely to be other unusual tenures elsewhere affected.

At best there is a narrow range of government action possible without infringing one of the treaties and at worst they conflict. In Slovenia under *FEANTSA*, how is the Slovenian government going to facilitate a lower rent, when this would take the rent below the minimum “cost” level prescribed in *Hutten-Czapska*? If Slovenia must facilitate eviction under *Hutten-Czapska*, how will it satisfy obligations to evictees under *FEANTSA* at a time of shortage?

Both governments could pay more housing benefit (not prominent in either country), which would be a large leap to the full cost of a welfare state at a time of recession. There is also an economic problem in both Slovenia and Poland that housing benefit as a preferred solution increases liquidity in the market, tending to increase both rents and the price of homes based on those rents (Granelle, 1998). In Slovenia, there is the additional burden of compensating people who did not get the right to buy, even though this was because the government was satisfying duties towards former landowners favoured in *Hutten-Czapska*.

Conversely, Poland must be vulnerable to a collective complaint from tenants if they put into effect *Hutten-Czapska*. The 45 percent of tenants who are in arrears with their rent in Poland are in a similar situation of vulnerable position of tenants in Slovenia, worse if they are all evicted. The Polish government was enjoined to provide for them alongside the high costs of paying damages and higher benefits. Polish tenants who happened not to exercise their right to buy are suffering deteriorating conditions as in Slovenia. Should the council of Europe bodies be so intrusive into property law, where they are not applying a consistent western European model of property and tenancies?

3. Transition to What?

It is implicit in the process of transition that there is a western European model both of property law and of the market but there is considerable scope for doubt about the unity of any such model adopted under the ECHR, the RESC and by the EU. It is becoming increasingly reasonable to consider these institutions together, in the light of the use of ECHR principles by the European Court of Justice of the EU. Increased rapprochement of European institutions will happen when the EU signs the ECHR

⁴⁵ The law of Scotland is different so it is not possible to generalize about the whole of the UK.

⁴⁶ Application 8793/79 (1986).

⁴⁷ The date of the Leasehold Reform Act 1967 which allows tenants of longer leases to buy the freehold or obtain another long lease. Leases can be as long as 999 years and such tenants regard themselves as property owners (Ball, 2010).

and RESC as a contracting party⁴⁸, even though the influence of the rights treaties is still limited to areas of EU competence (Grousot and Pech, 2010).

A lack of agreement on property rights this was apparent from signing of the ECHR. Allen (2007: 289) says the contracting parties were “seriously divided over the very nature and function of a human right to property”. The UK in particular did not agree to a guarantee of property rights. Article 1 of protocol 1 to the ECHR was later signed in 1952, protecting the peaceful enjoyment of “possessions” with the right not to be deprived of these except under certain conditions, particularly in accordance with the general interest, a quite vague formulation.

This section seeks to explain the inconsistent approach of the ECHR and the RESC to the law in the light of conflicting approaches to the market, and to variable concepts of tenant protection in private law and public law norms. The next section will deal with the interface with the EU and the concept of the market.

3.1. Differences in Private Tenancy Law

The difference in property rights between nation states is most apparent in landlord and tenant law, even though there is a common western European experience of tenant protection. This protection against arbitrary or unfair treatment by landlords has been a concern of many member states after the industrial revolution, brought poor living conditions after migration to cities. However, the way states achieved this differs considerably, a portfolio of policy responses including background protection of public health, planning and building standards, with benefits, subsidized construction and social housing.

Security of tenure for residential tenants is a traditional component of the welfare state, and long-term rent control such as that in *Hutten-Czapska* has been an experience of many western countries. From the Great War, a freeze in rents was seen in England⁴⁹, France⁵⁰, Spain⁵¹, Poland⁵² and probably elsewhere, commonly accompanied by effective security of tenure. This led to loss of rental stock with landlords anxious to withdraw from the market. However, tenants’ security was reduced very slowly and some properties affected still exist after more than 50 years⁵³. In Spain, reform of 1920s strong tenant protection is only happening now (Orti Vallejo, 2010)⁵⁴. In England and France, older secure tenancies were closed to newcomers but succession was limited, something allowing longer for family to find another home. Limits to succession were a cause of complaint in *FEANTSA*. Why then do Eastern European countries have to adapt so quickly, particularly when the background protection of residents above may not be fully in place? Manipulation of the boundaries of tenures to suit local conditions is the norm rather than exceptional.

Although tenants’ security of tenure is common across Europe to some extent, underlying conceptions of property rights and the regulation of tenancies are quite different. A tenancy is commonly an agreement, but this is often supplemented either by some protective status (whether by property law, consumer law or collective rights) or by statutory protection

As a matter of principle, English tenants are protected because they have a bundle of property rights, the result of a common law conception of fragmented property rights which is generally contrary to principle in civil law countries⁵⁵. This applies to all tenancies in the UK (excepting Scotland⁵⁶), and to

⁴⁸ Article 6 of the Treaty of the European Union as amended by the Treaty of Lisbon of 9 May 2008.

⁴⁹ Malpass and Murie (1999).

⁵⁰ Shapiro (1987).

⁵¹ Orti Vallejo and Jiménez Linares (2010).

⁵² Gromnicka and Zysk (2004: 4).

⁵³ A few properties remaining mainly in Paris under a statute no.48-1360 of 1 September 1948.

⁵⁴ Even then, there is still a compulsory five year term of letting and required standard terms.

⁵⁵ An exception is long leases, see Ball (2010).

⁵⁶ Scotland is a mixed jurisdiction with both civil law and common law elements.

slightly longer leases in Ireland (possibly)⁵⁷ and the US⁵⁸. This theoretically brings benefits such as inheritability, although severely limited by statute⁵⁹. In fact, there is a European tradition for secure long leases in the ‘emphyteusis’ a long lease from the late Roman period with a fixed term which is real property and can be mortgaged by the tenant, originally used to bring agricultural land into use (Terré and Simler, 1998: 699). This exists in France⁶⁰ and is the later basis of the construction lease⁶¹.

In contrast, a common feature of the 25 civil law countries in Europe is a concept of property ownership which precludes most tenants from having property rights. This dates from the French revolution, creating an absolute and unitary right to property⁶², to avoid the property fragmentation of feudalism.

Erp (2009) proposes a common European property model based on the ‘classical’ model from that 1804, with a *numerus clausus*, a strictly limited number of property rights. This is a conservative model because most European tenancy protection postdates this. It is awkward that the English *numerus clausus* effectively includes tenants’ rights and other hybrid property rights⁶³. Erp also proposes a principle that property rights should be either property or contract and not both. English tenancies are both of these. There is no common European model in this field, and no good case has been made to change the rights of property owners and tenants across the European union to very limited tenure forms.

The protection of tenants in civil law countries can be seen as a social response to tenants’ exclusion from property rights which is inherent in the unitary notion of property. Consequently, this exclusionary property is much more political than the UK scheme, with the rise of collective rights of workers and tenants against ‘capitalist’ employers and landlords. This approach (with the prospect of collective complaint) is embodied in the RESC, containing the right to work at article 1 and the right to housing at article 31. The tenants have social rights, either through the state in the general interest, or through direct participation of tenants in the negotiations with landlords, with the backing of trade unions as in France⁶⁴.

There is little pure law material comparing European tenancies, although there is a wealth of interdisciplinary material on rental markets.⁶⁵ There is one large EU-funded comparative study of 18 countries (‘The EU study’) from the European University Institute (2004). In this, the nature of the tenancy is not defined, and the study includes indefinite and lifetime tenancies. The levels of tenant protection vary both over time and between countries. The study, perhaps inevitably suffers problems with superficiality, with a lack of equivalence between national reports from considerably different traditions.

Legrand (2006) suggests that comparative research should be based on immersion in two countries’ legal systems for ten years. The author has done this for France and England, being respectively civil law and common law systems, but this too has its limits, in a Europe of 27 countries. However, the UK and French examples are already divergent and bringing in more knowledge from other countries is likely to decrease these documented divergences, not reduce them.

⁵⁷ See the (Irish) Law Reform Commission (2003) before the 2004 reforms, on the ambivalent status of tenancies.

⁵⁸ For leases longer than a year.

⁵⁹ See the Law Commission (2003) for a general account of this.

⁶⁰ Code Rural (Nouveau), arts. L.541-1 to L.541-14.

⁶¹ Code de la Construction et de l’Habitation, arts. L.251-1-L.251-9 and R.251 *et seq.*

⁶² Article 17, of the Declaration of the Rights of Man and the Citizen of 26 August 1789.

⁶³ Tenancies are classified as ‘estates’, alongside freehold ownership in s.1(1) of the Law of Property Act, whilst there is a category of equitable interests in land found in s.1(3) including rights such as the so called ‘equitable leases’, beneficial interests under a trusts and options.

⁶⁴ See Ball (2008, 2010).

⁶⁵ For example, the profuse interdisciplinary literature arising from the 20 year history of the interdisciplinary European Network for Housing Research, often researching tenancies within economics, sociology, planning, geography or cultural, with accumulated published examples found in their online archives.

In the general report of the EU study, Schmid (undated) suggests that the diversity of European tenancies is such that this should not be harmonized, due to national variability and embeddedness in national systems. Despite this Schmid tends to dismiss the importance of collective bargaining approaches to tenant rights (Finland, Sweden and France), as well as tenants' property rights (the UK and Ireland) and consumer approaches from Europe. This underestimates both difficulties in comparing tenancies in depth when there is a disciplinary mismatch but also role of tenancies in the formulation of divergent national public law norms governing property

3.2. Tenant Protection in the Public Law

These divergent approaches to property and tenancies means there are different normative bases for tenant protection which also affect the way the ECtHR and the ESC support tenants. In England, tenants are also property owners as a matter of law, although they may well not realize this. This means that when the UK Parliament legislated to protect tenants there was no requirement for a generic social principle as a legal basis for protecting tenants, separately from property rights. This would be inconsistent because, for example, in London where subletting is common, people can be both landlord and tenant at the same time. In Germany too, tenants enjoy a constitutional right to property reflecting strong security of tenure⁶⁶, although this is controversial. Property can thus protect tenants to a limited extent, for example in assumptions of rights to inherit and transfer tenancies, even though this is severely limited by statute in the UK. This property status can be seen in the ECtHR jurisprudence in isolated cases of long leases, *James and Iatridis v Greece*⁶⁷.

In most civil law countries, tenants find protection in constitutional principles outside property. In France, Dugout (1912) proposed that property has a social function. This was to provide a normative basis for limitations to property rights. Probably a majority of European countries limit property rights like this with tenants enjoying a 'social' right to housing (Center on Housing Rights and Evictions [COHRE], 2000), although the right to property was never modified like this in France⁶⁸. Instead French tenants' rights were based in the States power to intervene in the general interest and on a developing right to housing, which was formally based on the right to human dignity in 1994 (Ball, 2010).

The right to human dignity, is also a basis for welfare in Germany⁶⁹, and might be thought to provide some common basis for protection of landlords and tenants in a balanced way. There is, however little sign of a developing major welfare ECHR jurisprudence around that principle to protect tenants, although that the ECtHR has protected tenants in exceptional cases based around the right to the home (article 8)⁷⁰ or due process (article 6).

There are consequently divergent approaches to tenancies in constitutional law, which also affect the ECHR. German and English lawyers might expect tenants' rights protection under the ECHR right to property, whilst article 31 of the RESC corresponds to the French support for tenants, where it would not be expected to apply under the ECHR. The article 31 approach opposes landlords within the ECHR and tenants within the RESC in conceptual split between the political left and right, and between individual and collective rights.

4. Conflict within European institutions

⁶⁶ Article 14(2) of the *Grundgesetz*, elaborated by two cases, Beschluss des Ersten Senats vom 26 .Mai 1993 - 1 BvR 208/93- NJW 1993, 2035 and Pressemitteilung des BvrtG nr.46/2000 vom 13.4.2000, NJW-Informationen Heft 19/2000.

⁶⁷ (2000) 30 E.H.R.R. 97.

⁶⁸ The first draft post-war constitution included such a modification, but this was rejected by referendum.

⁶⁹ Particularly developed in articles 1-6 of the Bundesgesetzbuch, but a strong German approach generally.

⁷⁰ Decision 267/2004 27 May 2004.

If the rights treaties are to effectively form the constitutional court of the EU in any future acceleration of harmonization, this poses considerable problems in conflicting interpretations of tenancy law and its place in the legal order. This section looks first at the conflicts between the ECHR and the RESC and then the contribution of EU law to the burgeoning European regulation of tenancies, whilst raising questions about closer harmonization and integration of European institutions.

4.1. Conflict between the ECtHR and the ESC

Some conflicts between the ECtHR and the ESC have already been described, with the ECtHR tending to support landlords and the ESC tending to support tenants, with each court proposing different requirements. Eviction is a requirement of the court in *Hutten-Czapska*, but a specific cause of complaint in *FEANTSA*. There are conflicting measures of allowable rent. In *FEANTSA* this should be affordable whilst in *Hutten-Czapska* it should cover the landlords' costs at a minimum, apparently whether affordable or not.

Extra support for landlords from the ECtHR can be found in *Hutten-Czapska*, because the judgement of the lower chamber mentions a landlords' "right to profit",⁷¹ apparently approved by the Grand Chamber, which also referred to an "entitlement to derive profit".⁷² This is something new, more than the possibility of profit. In a partly dissenting opinion, Judge Zupančič⁷³ said 'the right to derive profit by merely owning an apartment building cannot be seen as a human right' and "a sheer profit for the landlord – in other words, income not derived from his services – is, for the tenant, of necessity a payment that is not reciprocated by a benefit. How can that be a landlord's human right?" In general European rent regimes force rents downward, not upwards to a level ensuring a profit, which in Poland could be unaffordable.

Extra support for tenants from the ESC can be found in *FEANTSA*, who were said to have suffered discrimination relative to other tenants. However, there was a difference between Slovenian private and public tenants because former owners of private stock had a restitutionary claim. The extent of this tenants' right to equality is not clear, based on discrimination relative to public and social tenants. This seems to imply that they should be assimilated to the more favourable position of tenants of the public and social sector. Equality could be created in Poland and Slovenia by reducing the security of public and social tenants but is doubtful that this was what the ESC meant, since the insecurity of private tenants was a complaint. This equality is a new proposition favouring tenants, probably inconsistent with the Polish decision deregulating private renting only.

This is an unresolved opposition between the political left and right. Landlords in Poland were the strongest lobby in *Hutten-Czapska*⁷⁴, and in Slovenia the only well organized group in renting are tenants of denationalized properties (Kerestes, 2004). In most western countries there are mechanisms to resolve landlord and tenants disputes, by both parties having property rights (in the UK), by collective bargaining (in France, Sweden, Finland⁷⁵ and perhaps other countries) and by a single constitutional court to balance landlord and tenants rights everywhere.

⁷¹ From a Constitutional Court judgement of 22 April 2005, quoted by the lower chamber in *Hutten-Czapska* (Application no. 35014/97), at para. 86, repeated with approval at para. 176. Para. 86 also spoke of assuring a profit.

⁷² Para. 239 of *Hutten-Czapska*, the Grand Chamber judgement.

⁷³ *Hutten-Czapska*, the partly dissenting judgement of Judge Zupančič, part III.

⁷⁴ Activity and provision of information by landlords' organizations is mentioned in *Hutten-Czapska*, paras 59, 99, 100, 135, and 336.

⁷⁵ Schmid (undated). Although the EU project questionnaire asked whether there were negotiations with landlord and tenant organizations affecting tenancies, not everyone seems to have answered it and the project only covered 18 countries. The questionnaire did not cover situations where tenants or unions were strong lobbies and consulted as a matter of course, such as in Spain.

4.2. Questions about European Rights and the EU

If the ECtH and the ESC together are gradually becoming a constitutional court for the EU after Lisbon this raises questions. Why are there two such constitutional courts? There is a real danger of unresolved conflict either between institutions or between the texts and objectives of the different treaties. Merging the two would be politically unlikely but it might be better to see the prospect of an appeal from the ESC to the ECtHR as ultimate constitutional court. The ECtHR might have to become more hospitable to tenants, emphasising the integration of these rights with others such as human dignity, the right to family life or the right to a fair trial. This would encourage a co-ordinated policy.

Another question concerns that the ECtHR allows a margin of appreciation to national governments in applying property rights. Since the Council of Europe organizations and the EU pursue common objectives in the free market, is the ECtHR using the 'margin of appreciation' allowed to national governments as a proxy for subsidiarity? Subsidiarity is the EU rule that decisions should be taken at the lowest level⁷⁶. The margin of appreciation also allows members room for manoeuvre, but this is not the same juridical concept⁷⁷. The margin of appreciation implies that there is a common concept of property which is assessed nationally in different ways (Whittaker, 2008, Ball, 2008). Firstly, the competence of the EU does not extend to property, another treaty mismatch, and secondly, there is no common approach to property, certainly so far as tenancies are concerned within the diversity already described.

If a common property policy were pursued, this should be co-ordinated between the ESC, the ECtHR, and the EU, even though property law is outside the EU treaties but is a full competence of the ECtHR. Will this interaction drag real property law further into EU regulation? The new TFEU still provides that 'The Treaties shall in no way prejudice the rules in Member States governing the system of property' (article 345, formerly 295). Housing is also not strictly an EU policy, but there are already policies impacting on both housing and tenancies and thus on property. Free movement of persons is affected by the availability of rented property, particularly if tenants are consumers. The Unfair Terms in Consumer Contracts Directive⁷⁸ directly affects tenancies, although Schmid (undated) suggests this is of little importance.

The EU-sponsored Draft Common Frame of Reference could also affect tenancies⁷⁹. This large project moves towards harmonization of European contract law drawing on 20 years of comparative European study. The DCFR excludes real property but would automatically affect tenancies, were it ever to have effect within the EU. The website of the EU study suggests that application of this to tenancies would result in a 'conflict-prone and disintegrative situation' arising from different national values and from the superimposition of general rules on a special contract (European University Institute, 2004). Should tenancies accordingly be explicitly excluded from such harmonization at least until tenancy and property law differences are more studied?

EU law also affects whether national governments such as Slovenia and Poland can construct social housing to assist tenants displaced by a policy promoting eviction or to help citizens generally affected by housing shortage. The European Competition Commission recently moved to limit government housing construction subsidies except for people on a low income or in need. This followed around three years' negotiation following complaints from property owners in France, the Netherlands and Sweden, causing those countries to change traditions of universal access to social housing (Ball, 2008). A Commission decision on the Netherlands limits state aid to people in need of housing on a low

⁷⁶ Protocol No. 2 to the TFEU, article 1.

⁷⁷ There are procedures for EU states to object to subsidiarity being insufficiently taken into account in the EU (*ibid*) but no such appeal is possible from the ECHR.

⁷⁸ Directive 93/13/EEC.

⁷⁹ Published as Acquis Group (Study Group on a European Civil Code/Research Group on EC Private Law) (eds.) Principles, Definitions and Model Rules of European Private Law, The Draft Common Frame of Reference, Full Edition, (München: Sellier, 2009).

income⁸⁰. There are other state-aid exemptions, but this limitation of state financial aid for construction could be a difficulty in housing policy generally.

4.3. What Economic Model is in the General Interest?

It might be thought that the ECtHR is adopting a liberal, politically right-leaning approach to the free market where most tenants do not have substantial rights. The court's concern with social exceptions as in *Mellacher* and *James*, suggests this is not necessarily so. Nevertheless it is difficult for a moral court to propose more eviction and higher rents in countries with no cultural tradition of eviction, as an instrument to create a free market. Reluctance to evict could be linked to severe stigmatization of debtors in national credit markets, the effect of which on individuals should not be underestimated.

It can be questioned whether the ECtHR and the ESC are appropriate instruments for this economic policy in *Hutten-Czapska* and *FEANTSA*. No evidence on economics was taken, nor does the court explicitly refer to economic studies and policies. This might not be possible where these institutions do not normally explicitly take the ability of a state to pay into account, something of a problem in finding a working economic model. This results in severe financial burdens on Slovenia and Poland.

An inconsistent economic model is applied concerning the general interest. This utilitarian idea is common legal approach within the ECHR, RESC and the EU when deciding whether to support tenants. This is a basis for tenant support in the French Constitution,⁸¹ and a derogation from the right to property in the ECHR. The essentially similar 'common interest' is also the basis for exceptions to competition law in the EU.⁸² In the early days of strong tenant protection described above, it might be assumed that the social protection of tenants was in the general interest, and an example of social progress. This approach is embodied in the RESC, a basis for tenant support in *FEANTSA* under article 31, which was new in 1996.

Today, more liberal ideas concerning the effective functioning of a housing market tend to propose that the general interest can require reduced rights for tenants, as can be seen in *Hutten-Czapska*. This can only be explained in economic terms, something for which there is little space in this article. This is not simply a question of the politics of left and right but of different economic theories. On top of basic doctrinal divergences, the heavy or light regulation of tenant security in favour of landlords or tenants is frequently due to governments adopting a particular economic model.

Kemeny (1995) observed two major economic approaches to rental markets in Europe, a unitary and dualist approach. The dualist or 'Anglo-Saxon' model from the US and the UK segregates public or social renting (cost renting) from private profit renting to prevent competition between the two rental sectors. This 'free market' approach seems to be the model adopted by the ECtHR in *Hutten*, by deregulating private but not public or social housing, and by the Competition Commission in limiting investment to a residualized social sector. Kemeny says:

'Limiting the availability of cost rental housing effectively forces the great majority of households to choose between insecure and high rent profit renting on the one hand and owner occupation on the other. The result is a rapid expansion of owner occupation.' (Kemeny, 1995: 179)

Policy problems included glut-famine cycles in a fragile market, and increasing pressure to subsidize low income people to purchase and stay in tenure.

The alternative unitary or integrated model, conceived in Germany before the War, promotes competition between subsidized cost renting and private profit renting. This tends to reduce rental

⁸⁰ before Lisbon, the general interest was an exception to the limitations to state housing subsidy in EU competition law (Article 87 TEU). The New Treaty on the Functioning of the EU (TFEU), article 107, now allows aid 'of a social character' or 'not contrary to the common interest' (Article 107, para. 2(a) TFEU).

⁸¹ In article 17 of the Declaration of the Rights of Man and The Citizen of 26 August 1789, still part of the 1958 Constitution (This concerns public utility), and in articles 544 and 545 of the Civil Code.

⁸² Above note 80.

prices and increase quality in the private profit sector. This approach requiring equality between tenants seems to have been adopted by the ESC in *FEANTSA* deciding different treatment of tenants was discrimination. Kemeny suggests that the policy problem with this is that, there is increased regulation over time and command approaches from the State to maintain equality between sectors. In legal terms, this is a limitation to freedom to contract, and freedom for property owners to deal with their property. Kemeny suggested that the dualist approach led to more diverse tenures and national approaches. Kemeny (2007) identifies Germany, Sweden, Denmark, The Netherlands, Switzerland and Austria as working like this, using the not-for-profit sector to temper the free market. What then is the way forward for European property and tenancy regimes?

5. Conclusion

In rental markets, there is a diversity of economic thought in Europe to which this article cannot do justice as well as substantial diversity in legal principles. There is no obvious way forward in the development of a coherent European regime for property either in law or economics. This diversity could be regarded as a kind of wealth to be drawn on by national governments in the myriad economic, cultural, social and historical circumstances facing them. A lack of common European property principles means the ECHR already allows a substantial margin of appreciation in the right to property, due to a lack of common European principles. Perhaps this margin of appreciation should be larger, when it comes to tenancies.

Despite this legal and economic variability of principle, the ECtHR, the ESC and the EU competition commission, have all acted to limit housing policy in ways seeming uncoordinated and inconsistent. The European Commission and the ECtHR are pushing towards a “free market” or dualist model of rental markets. However, public order legislation such as public health, building regulations, procedural protection on eviction and other background principles (such as property rights for tenants in the UK) mean no market is entirely free⁸³. In contrast, the ESC is pushing towards a unitary economic model by an implication that tenants in private denationalized property should have their position improved upwards to that of public tenants, something requiring heavy continuing regulation.

The ECtHR has pushed forward landlord rights to a profit whilst the ESC has promoted tenant rights to equality. Both these approaches imply heavy regulation rather than a free market, and in ways that are mutually inconsistent. The problem at the heart of the social *acquis* is that improved tenant rights mean worse landlord profits, although up to a point a better tenancy product could improve both landlord and tenants positions in the longer term, by making tenancies more attractive⁸⁴. A concern is the political dimension of this with the ECtHR supporting the politics of the right and the ESC supporting the politics of the left, in circumstances where there is no cross-appeal between these institutions. This lack of co-ordination risks reviving left-right disputes long settled in Western Europe.

The EU tenancy project, the only large scale legal study, has strongly recommended not harmonizing tenancy law and this should not be disregarded. The potential for damage to national housing markets is considerable where national government do not have the full range of housing policies available to them to tackle problems. If the trend of limiting available housing policies continues, national governments should pay housing benefit, rather than subsidize not-for-profit construction other than for the very poor. Rent control was to be removed in Poland, even when this would make rent unaffordable. Even then, Polish tenants could complain about the higher rents breaching their right to housing as in *FEANTSA*. Then Poland and Slovenia also face paying compensation to landlords and tenants respectively, in addition to funding higher housing benefit due to lifting of rent control.

⁸³ Stiglitz (2000) sees the public and private markets as interdependent, a more modern approach.

⁸⁴ In *Hutten-Czapska*, both landlords and tenants were looking to improvements to the rental problem because lack of repair affected both groups.

Implementing this does not seem very likely given current low benefit levels, the relatively low national wealth and the recession. The state is thus responsible for satisfying both landlords and tenants with reduced tools for doing so.

The only reasonable solution to these dilemmas is increased co-ordination between the various European institutions to produce a transparent and consistent light-touch approach to national property and tenancy law. It is arguable whether this intrusion should happen at all, except as a last resort. Property law has evolved in Europe for centuries, not the short time-scale allowed by the ECtHR for transition to the free market. There is need for considerably more research into successful national property laws, before regulating these. This might allow delight at the many ideas, with diversity as a resource to be properly assessed rather than harmonization automatically being the preferred approach. National property laws are a delicate balance at the heart of some very different national legal cultures. Housing markets are also notoriously multi-factorial, and in a context of legal variability and economic disagreement, a long-time scale for any European involvement in tenancies is essential.

References

- Allen, T. (2007) 'Compensation for Property under the European Convention on Human Rights,' *Michigan Journal of International Law*, 28(2): 287-335.
- Balchin, P. (Ed.)(1996) *Housing Policy in Europe*, London, Routledge.
- Ball (2010) *Towards a Deeper Understanding of Property, Tenancies and Land Occupancy in the Face of Legal Doctrinal Divergence*, paper given at the ENHR conference in Istanbul, 6 July 2010.
- Ball, (A.) J. (2008) *The Limitations to the Right to Housing for the Poor and Disadvantaged in France: Insiders and Outsiders in Social Housing Allocation*. PhD thesis, Sheffield, Sheffield University.
- Brillat, R. (2005) 'The Supervisory Machinery of the European Social Charter: Recent Developments and their Impact' in De Búrca and de Witte (2005), below, pp 31-44...
- Granelle, J-J. (1998) "Le Logement Comme Bien de Consommation", in Coloos, B. (ed.)(1998) *Comprendre les Marchés du Logement*, Paris, l'Harmattan, pp. 25-60.
- Central Statistical Office (2007) *Housing Economy in 2007*, Warsaw, Central Statistical Office.
- Cirman, A. (2006) 'Housing Tenure Preferences in the Post-privatisation Period: The Case of Slovenia', *Housing Studies*, 21: 1, pp 113-134.
- COHRE (2000) *Housing Resources for Housing Rights, International and National Standards*, 2nd edition, downloadable from <http://www.cohre.org/>
Geneva, COHRE,
- Commission of the European Communities('CEC') (1991) *Annual Report*, Luxembourg, CEC.
- De Búrca and de Witte (eds.)(2005) *Social Rights in Europe*, Oxford, Oxford University Press.
- Duguit, L. (1912) *Les transformations générales du droit privé depuis le Code Napoléon*, Paris, LeBon.

Erp, J.H.M. van (2009) 'From "Classical" to Modern European Property Law'. In *Essays in Honour of Konstantinos D. Kerameus/Festschrift für Konstantinos D. Kerameus, Vol., I* Athens/Brussels, Ant. A. Sakkoulas/Bruylant, pp 1517-1533.

European University Institute (2004) *Tenancy Law and Procedure in the EU*, Florence, European University Institute.
<http://www.eui.eu/Documents/DepartmentsCentres/Law/ResearchTeaching/ResearchThemes/EuropeanPrivateLaw/TenancyLawProject/TenancyLawGeneralReport.pdf> (last visited 25 March 2010). (The date of 2004 is approximate because only a few of the component national reports are dated).

Gromnicka, E. and Zysk, P. (2004) 'Polish Tenancy Law and Principles of European Contract Law', in European University Institute (2004) above.

Groussot, X., and Pech, L (2010) *Fundamental Rights Protection in the European Union post Lisbon Treaty, Fondation Robert Schuman Policy Paper, European Issue no. 173*, Brussels/Paris, Fondation Robert Schuman.

Hegedüs, J., Mayo, S. & Tosics, I. (1996) "Transition of the housing sector in the East-Central European countries", *Review of Urban and Regional Development Studies*, 8, pp 101-136.

Quinn, G (2005) 'The European Social Charter and EU Anti-discrimination Law in the Field of Disability: Two Gravitational Fields with One Common Purpose' in De Búrca and de Witte (eds.)(2005), above, pp.279-304.

Kemeny J. (2007) 'Corporatism and Housing Regimes' in *Housing, Theory and Society*, Vol. 23, No. 1, 2006, pp 1-18.

Kemeny, J. (1995) *From Public Housing to Social Market, Rental Policy Strategies in Comparative Perspective*, London, Routledge.

Kerestes (2004) 'Report on Tenancy Law in Slovenia', part of *Tenancy Law and Procedure in the EU*, above, European University Institute (2004).

Law Commission (2002) *Renting Homes I; Status and Security*, Consultation Paper no. 162, (England), The Stationery Office.

Law Reform Commission (2003) Consultation Paper on General Law of Landlord and tenant, LRC CP 28-2003 (Ireland), The Law Reform Commission, 35-39.

Legrand, P. (2006) 'Antivonbar' in *Journal of Comparative Law*, Vol. 1, Issue 1, pp 13-40.

Malpas, P. and Murie, A. (1999) *Housing Policy and Practice*, Basingstoke, MacMillan Press.

Ministry of Infrastructure of the Italian Republic and Federcasa (2006) *Housing Statistics in the European Union 2005/6*
http://www.federcasa.it/news/housing_statistics/Report_housing_statistics_2005_2006.pdf (visited 16 June 2010).

Moravcsik, A. (1995) 'Explaining International Human Rights Regimes: Liberal Theory and Western Europe', in *European Journal of International Relations*, 1995, 1, 57, pp 158-189.

Orti Vallejo, A and Jiménez Linares, M..J. (2010) *Latest Measures to Promote Growth of Spanish Rental Housing*, paper delivered at the ENHR conference in Istanbul on 6th July 2010.

Pilcher-Milanovich, N. (2001) “Urban Housing Markets in central and Eastern Europe: Convergence, Divergence or Policy Collapse?”, *International Journal of Housing Policy*, Volume 1, Issue 2 August 2001 , pp. 145-187.

Schmid, C.U. (undated) ‘General Report’ (European University Institute 2003) above.

Sendi, R (1999) “Housing Construction in the Transition Period: Slovenia's Non-starter Situation” in *Housing Studies*, 14: 6, pp 803—819.

Shapiro, Ann-Louise (1980) *Housing the Poor of Paris, 1850-1902*, Madison, University of Wisconsin Press.

Shorts, E. and De Than, C (2001) *Human Rights Law in the UK*, London, Sweet & Maxwell.

Stiglitz, J.E. (2000) *Economics of the Public Sector* (third edition), New York/London, W.W. Norton and Co.

Sunstein, C.R. (1991) ‘Constitutionalism, Prosperity, Democracy: Transition in Eastern Europe’ in *Constitutional Political Economy*, Volume 2, Number 3 / September, 1991, pp 372-394.

Terré and Simler (1996) *Droit Civil, Les Biens*, 5th Edition, Paris, Dalloz.

Varju, M (2009) *Transition as a Concept of European Human Rights*, on the Social Science Research Network website http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1329978, viewed 15/06/2010.

Whittaker, S. (2008) ‘Burden of Proof in the Consumer *Acquis* and in the Draft Common Frame of Reference: Law, Fact and Things in Between’, *European Review of Contract Law*. Volume 4, Issue 3, pp 411–444.

Williams, R.C. (2007) *The Contemporary Right to Property Restitution in the Context of Transitional Justice*, New York, International Center for Transitional Justice.

The challenges of integrating the carbon reduction scheme into commercial leases ¹

Susan Bright
New College
Oxford

Abstract

This paper explores the potential impact that the introduction of the CRC Energy Efficiency Scheme will have on a) energy use in the tenanted commercial built environment and b) the idea of the net lease. The paper reviews various background documents preceding the implementation of CRC in order to identify the abatement incentives established. Against this, the common structure of commercial leases, and the early property market reaction to CRC, the paper considers how effective the CRC scheme is likely to be in achieving the twin goals of carbon saving and landlord-tenant collaboration.

Key to the success of the CRC scheme will be the way in which the financial and reputational drivers established by the CRC scheme incentivise landlords and tenants to make technological and social changes to reduce energy consumption. Given the variety in the ways energy is supplied to tenanted commercial property, the complexity of the CRC scheme, the traditionally adversarial relations between landlords and tenants and the ‘split-incentive’ of commercial leases, energy abatement opportunities are significantly more limited in the leasehold context than for owner-occupied properties. Nonetheless, the introduction of the CRC scheme has begun an important conversation and is an important step towards tackling energy efficiency.

Paper

The UK took an important global lead in April 2010 by implementing a compulsory carbon trading scheme through the CRC Energy Efficiency Scheme Order 2010 (the ‘CRC scheme’).² The government states that the objectives of the CRC scheme are ‘to encourage carbon savings within large organisations, primarily by driving uptake of cost-effective end use energy efficiency measures and fuel switching – in order to help the UK meet its medium and longer-term greenhouse gas emission reduction targets’ (Department of Energy and Climate Change, 2010). The scheme works by requiring qualifying organisations to register as CRC participants and purchase annual carbon allowances to cover their anticipated emissions during each reporting year running from 1 April to 31 March. The CRC scheme applies to organisations whose total annual half hourly metered electricity use in 2008 was above the qualification threshold of 6,000 MWh; this roughly equates to an annual electricity bill of more than £500,000. Small energy users are excluded because the administrative and transaction costs of including

¹ My thanks to Peter Williams of Eversheds, and Sue Highmore for their considerable help in understanding this topic, and for commenting on an earlier draft.

² SI 2010/768.

them would outweigh any energy savings (NERA Economic Consulting and Enviro Consulting, 2006). The Department of Energy and Climate Change estimates that around 5,000 organisations will qualify for full participation, and around a further 15,000 will have to make an ‘information disclosure’ (which relates to metering and energy use). During the early years of the scheme the price of carbon is fixed, but from 2013 allowances will be sold on an auction basis and the imposition of a government cap will lead to a gradual reduction in overall emission levels. It is not only through the pricing and cap mechanisms that the scheme aims to drive down emissions; there are also provisions for the ‘recycling’ of payments at the end of each accounting year, which will reward those organisations that have made the most strides towards improving energy efficiency. Performance is made transparent through the publication of a CRC ‘Performance League Table’. Although it has not been specifically designed to target emissions from the commercial built environment, it is clear that the CRC scheme will have a significant impact on the property sector.

Of the 5,000 or so organisations that will be required to participate under the CRC scheme, many will be landlords of commercial property. If these organisations are to reduce energy emissions to reflect the goals of the CRC scheme it is therefore necessary that the structure and terms of leases permit appropriate abatement measures to be taken. It is, however, widely acknowledged that the majority of UK leases present serious barriers and disincentives to the implementation of energy efficient measures. This has been recognised throughout the policy history of the CRC scheme. In the first major publication to promote a mandatory cap and trade scheme, the Carbon Trust’s 2005 report on the UK Climate Change Programme, the ‘landlord-tenant divide’ was identified as one of the key barriers (The Carbon Trust, 2005). The challenges to the CRC scheme generated by the division of ownership and management responsibility within the landlord and tenant relationship and what is commonly referred to as the ‘split incentive’ (which is explained below) have been referred to in all of the key policy documents since then. Notwithstanding this, the CRC scheme makes no express provision for landlords and tenants and the only indirect acknowledgement of the difficulties that there may be in delivery of the scheme in this context is the statutory imposition on tenants of what the government has described as a duty to cooperate with the landlord organisation for the purposes of complying with the CRC scheme (Department of Energy and Climate Change, 2009). On what is a central issue, the ability of landlords to pass on the costs of participating in the CRC scheme to tenants, the government conspicuously decided not to legislate but to ‘develop voluntary “good practice” guidance (in collaboration with both landlords and tenants)’ (Department for Environment, 2008). This work is ongoing, but as yet there is no sign of industry consensus on either the principle of whether tenants should be required to contribute to their landlord’s CRC costs, or (if so) how to reflect this in the leasehold drafting.

Much of the property industry is critical. Liz Peace, the Chief Executive of the British Property Federation (the lead organisation representing property owners and investors in the UK) has said that:

... the scheme has not been properly thought through. Large areas of uncertainty remain over how you apply the CRC around existing leases. Landlords have no

legal remit to influence how tenants use energy. (Property Industry Working Party, 2009)

This paper will consider the challenges in applying the CRC scheme so as to deliver the desired policy outcomes in the context of commercial landlord and tenant relationships. There is a subsidiary issue that is also very important in practice, to do with how the CRC scheme can be mapped onto the wording of existing leases, and provided for through the wording of new leases. This subsidiary issue will be touched on below but the primary focus of this paper is upon the challenge of using the CRC scheme to achieve serious energy reduction against the constraints of the traditional structure and pattern of commercial leases in the UK.

An Overview of Energy Payment in UK Commercial Leases

Much commercial property in the UK is leased. It is difficult to find reliable figures, but a report of the Investment Property Forum estimates that 43 per cent of the total stock by value is in the 'investment market', comprising just over 60 per cent of the retail and office sectors, but a little less than 25 per cent in the industrial sector (Investment Property Forum, 2005, 51). It is also hard to find how much is let to more than one tenant, but one statistic suggests that 70 per cent of tenanted commercial buildings are multi-occupancy (NERA Economic Consulting and Enviro Consulting, 2006, 43). The IPF report is confined to 'investment grade' property and although the commercial property sector is broader than this, CRC participants are most likely to be owners or occupiers of this level of property. Further, it is in this market that what is referred to as the 'institutional lease' is dominant.

This idea of the institutional lease is important in relation to this paper for two reasons. First, the fact that for many landlords the lease is primarily an investment requires the lease to be readily marketable as an asset, and this is a driver towards investment leases taking a common format and not containing unusual provisions. As Williams and Eddington explain, standard lease terms carry the advantages of acceptability in the marketplace, easier negotiation and familiarity (Williams and Eddington, 2009). Leases that reflect a standard pattern are much easier to attach a capital value to (Edwards and Krendel, 2007, 142-143). This leads to the second aspect of the institutional lease that is important. For an investor landlord the ideal is a lease that delivers a long-term secure income. One outworking of this principle, important in the CRC context, is that all unknown servicing costs (such as repair and insurance) are borne by the tenant. This means that the landlord receives a net rent, one that is not vulnerable to deduction of variable costs. In most cases, therefore, it is the tenant that pays for energy costs for the premises.

Although it is the tenant who will usually pay for energy, it does not follow that the tenant is in charge either of energy supply, or of how much energy is used. A sample of the variety of arrangements found in practice can be seen from looking at a case study based on the property portfolio of Land Securities, the UK's largest Real Estate Investment Trust (Burgess Salmon and Ecofys, 2008, 67). The summary below builds on

this case study (making some further assumptions) to describe a range of tenancy and energy arrangements that exist:

- a. The whole building is let to a sole tenant, which has its own energy supply.
- b. Multi-let property scenario 1: the landlord supplies energy for the common parts and for heating and air conditioning throughout the property (with the costs being passed through to tenants by a service charge) but the each tenant has its own electricity supply to its premises.
- c. Multi-let property scenario 2: all energy is supplied by the landlord (with costs passed through to tenants by a service charge and the tenant's share is calculated by reference to the floor area occupied by the tenant).
- d. Multi-let property scenario 3: all energy is supplied by the landlord but there is sub-metering for tenants (and the energy costs are passed through by a service charge which takes account of each tenant's individual energy use within its premises).
- e. Multi-let property scenario 4: this contains a mix of retail and office space. The retail tenants have their own electricity supply, but gas is supplied by the landlord. The office tenants have all energy supplied by the landlord. (It is more common for retail tenants to have direct electricity supplies than office tenants.) The landlord supplies all energy for the common parts. Service charge provisions are complex to reflect the mix of arrangements.

Under each arrangement it is the tenants who (individually or collectively) pay for energy. With multi-occupancy, even if the tenant has a direct supply of electricity to its own premises, there will usually still be a service charge to cover energy supplied to communal areas and the landlord's supplies to the tenant's premises for services such as heating and air conditioning. Where energy costs are included within the service charge, they can account for up to 20 per cent of the service charge in commercial offices and up to 6 per cent in the retail sector (NERA Economic Consulting and Enviro Consulting, 2006, 42).

The net rent concept might suggest that the tenants should also pay for the additional costs incurred by their landlord under the CRC scheme. It is not, however, so straightforward. One consideration, at the level of principle, is that the costs of being a CRC participant are not related to the energy costs of any particular building. Not all landlords will be required to register under the CRC scheme – it will depend on the level of energy supplies for which they are responsible. Those who are required to participate in the CRC scheme will cross the threshold because of the activities of their organisation as a whole. It is by no means obvious that the net rent idea means that a CRC registered landlord should be able to pass on the costs of being a CRC participant when those costs are neither specific to the building nor to the tenant's business.

Even if the principle of pass through is accepted, there are then very complex further issues. Should only the costs of carbon allowances be passed through, or also the additional costs of participating in the CRC scheme, such as administrative and registration costs? If costs are passed through should recycling payments be passed

through as well, or should they be retained for investment in energy efficiency measures? Should costs and refunds passed through reflect individual tenant's personal energy use? Even if in principle the answer is yes, it may not be technically possible if there is no sub-metering: at the time of the Land Securities Case Study, only 3 of the 22 shopping centres owned and managed by Land Securities provided tenants' electricity via sub-meters (Burgess Salmon and Ecofys, 2008, 66). And what happens if there is a change in ownership – either of landlord or tenant? These issues are extremely complex. Within existing leases, it is doubtful if CRC costs can be passed through to tenants, although it will depend on the precise wording used in individual leases. For new leases, if CRC costs are to be provided for, the drafting challenges are extraordinarily complex (Williams and Highmore, 2010).

The fact that tenants pay for the energy costs is part of the widely discussed 'split-incentive' problem. The other dimension of it relates to responsibility for the provision of energy supply equipment under the lease, providing for core services such as heating, air conditioning and electricity. As seen above, it is usually the landlord that supplies energy to a multilet property, at least in part. The opportunity to change or upgrade technical equipment to 'greener' technology is usually only available to the landlord as the equipment will not usually fall within the tenant's premises, and even if it does there will usually be restrictions in the lease preventing the tenant from making structural alterations. Indeed, the Carbon Trust state that up to 88% of cost-effective measures are within the remit of the landlord to change (seemingly referring to technological measures) (The Carbon Trust, 2009, 111). To the extent that the tenant is able to invest in energy efficient works, it will only be motivated to do so if there is sufficient pay back. Given that relatively short leases have become the norm in the UK, there may not be a strong enough return.³ This means that the energy management opportunity lies with the landlord, but the financial incentive is with the tenant who pays for the energy (of course, the CRC scheme may begin to change the way incentives work). Further, the cost of any equipment upgrades will usually fall to the landlord who may be unable to pass the capital costs through the service charge. As the RICS Code of Practice for Service Charges in Commercial Property illustrates, the general principle found in service charge arrangements is that the landlord can recover the costs of maintenance and repair of plant and equipment, but not initial or improvement costs, and can only recover the costs of replacement where the equipment is beyond economic repair.

This 'split-incentive' is not simply the outcome of 'lawyer's drafting' but is core to the way that the UK commercial property market has developed. Attempts to move away from the net lease model to accommodate occupier preferences have generally not taken off because investors find it difficult to attach value to buildings let in this way. By way of illustration, in recent years Land Securities marketed a 'Landflex lease' with all-inclusive rental but this is no longer on offer 'for several reasons...including the fact that the investment market failed to come to terms with how to value buildings that were subject to "managed office" leases.' (Punshon and King, 2010). To think of shifting away

³ In 2006/07 the average length of the commercial lease was around six years; 2/3 of leases were for five years or less; and fewer than 3% of leases were for more than 15 years. Sources: The Annual Lease Review published by the British Property Federation (BPF) and the Investment Property Database (IPD).

from the net lease model involves disturbing core principles that underpin the investment value of property; yet it is commonly stated that without such a paradigm shift it is not possible to make significant progress in improving the environmental performance of the built environment (Dowden, 2009; Lovell, 30 April 2010). The difficulties and challenges in adapting traditional lease structures and clauses to accommodate CRC principles may have an unsettling effect on the market. The Royal Institution of Chartered Surveyors (RICS) has commented on the possible market adjustments:

Many aspects of the property market will be affected by the confusion created ... by the unfortunate and unintended consequences of the legislation. ... Transactions are likely to be increasingly illiquid and the market may segment into regulated parts and unregulated parts, with heavy energy users in leased properties being most likely to fall into the unregulated parts since regulated landlords will seek to distance themselves from such uses. (RICS, 2010)

There have also been concerns expressed that implementation of the CRC scheme will generate greater conflict between landlords and tenants. This relationship is often portrayed as adversarial in nature. What is needed to improve energy efficiency in shared space is cooperation, but many consider that working out how to deal with the CRC scheme, particularly its costs, may introduce further friction into already tense relations.

An overview of CRC goals

Commercial, public and industrial buildings account for 18% of total UK carbon emissions. The government's impact assessment refers to a 'large body of evidence suggesting strong potential for reducing carbon emissions cost-effectively' (Department of Energy and Climate Change, 2010, 8). More specifically the Carbon Trust report of 2005 found that by 2020 emissions in non-domestic buildings could be reduced by at least 20% by using existing cost-effective carbon abatement technology (The Carbon Trust, 2005, 3). The aim of the CRC scheme is clearly to reduce carbon emissions, but to do so only in part by financial drivers. Indeed, the role of the financial levers is in large part to attract senior 'management attention'. An important goal is therefore to change organisational behaviour by changing the attention that carbon emission levels receive within management. It aims to do this through the cost of CRC allowances, but also by the strict reporting, accounting and measuring systems that will need to be put in place (NERA Economic Consulting and Enviro Consulting, 2006). Alongside this is the reputational driver of the Performance League Table which will rank participants annually based on their performance as measured by defined metrics.

This is an interesting approach. Given that there are cost-effective abatement measures already available that are not being taken up, it is clear that it is not economic factors alone that explain the slow uptake of energy-efficiency improvements. In non-energy intensive businesses energy costs represent a very small portion of an organisation's costs – around 1-2% - and so, even if a measure is cost-effective, it will seldom be a significant enough saving to get on the management radar. Biggart and Lutzenhiser have drawn attention to the fact that economic sociology can help to explain why organisations frequently fail to implement cost-effective investments in higher efficiency equipment (Biggart and Lutzenhiser, 2007). Price alone often does not shape behaviour, particularly

in relation to environmental issues which tends to be driven by beliefs, aspirations and community norms. The focus of organisations is likely to be upon their core businesses that have nothing to do with energy, and they may be unaware of, or unconcerned about, energy costs. The aim of the CRC scheme is to encourage such organisations to readjust their focus.

Delivering CRC goals in the Tenanted Sector

The CRC scheme is part of a bigger picture within which energy reduction and sustainability are increasingly becoming the focus of corporate attention. Other drivers include various policy initiatives (such as energy performance certificates, display energy certificates, and revision of building regulations) and market factors (such as corporate social responsibility agendas, and employee recruitment and retention). The issue addressed in this section, however, is the extent to which the CRC scheme in itself is likely to contribute to the goal of reducing carbon emissions given the acknowledged barrier that the landlord-tenant relationship presents. In the 2006 report produced by NERA Economic Consulting and Enviros Consulting there was an attempt to put a figure on the extent to which the misalignment of incentives within the landlord-tenant relationship reduces abatement potential. The figure is very 'rough and ready' and involves making several assumptions about the property market, and property use, which may not be terribly robust, but the conclusion is that for all tenanted commercial properties combined the 'reduction in cost-effective abatement potential could be in the order of 44 percent' (NERA Economic Consulting and Enviros Consulting, 2006, 43). Although the figure is only an estimate it does suggest that the structure of landlord-tenant relationships reduces the potential for energy reduction significantly, in some cases by almost half.

RICS is even more doubtful as to the likely effectiveness of the CRC scheme. In response to the government consultation in 2009 it commented:

As the legislation and draft order stand, however, the distance between responsibility and payment, cause and effect, and even blame and guilt means that there can be little certainty of success in driving down the use of carbon in tenanted buildings. (RICS, 4 June 2009)

To evaluate this claim, it is helpful to think about the potential for energy reduction by addressing separately abatement 'opportunities' and abatement 'incentives'. Abatement opportunities are presented through the technical possibilities for more efficient equipment and infrastructure, and through behavioural change. Abatement incentives are those factors that incentivise a party to pursue abatement opportunities.

The interrelationship between abatement opportunities and abatement incentives under the CRC scheme is complex, and will vary according to the particular physical and leasehold configuration involved.

At a simple level it will broadly be the case that in most leases, tenants have an existing abatement incentive provided through energy costs, but as we have seen these are generally not significant enough as an overhead to produce much behavioural change. *If,*

in the new CRC world, tenants are required to contribute to CRC costs then they *may* have an additional financial incentive to reduce energy use. The reason for caution about this conclusion is because the incentive may further depend on the way the tenant's share of energy costs is calculated. If there is sub-metering and the CRC costs passed on to the tenants are based on individual tenant energy use then the incentive is highest. A consultation exercise within the property industry supported this idea that the division of CRC costs amongst tenants should reflect actual energy use (Brown, 2010). Land Securities has also stated a preference for pass-through based on individual tenant use, basing this on the 'polluter pays' principle being fundamental to the use of this kind of economic instrument to manage environmental impacts.⁴ Of course, this option is available only if there is sub-metering in place to record actual energy use. If costs are instead calculated on a simple floor area basis, the incentive is less: why should one tenant make great efforts to reduce energy use if it has to pay towards the profligacy of another? This has been referred to as the 'commons dilemma' but interestingly (and encouragingly) when this was explored in a residential context and incentive structures introduced to reward specific conservation behaviours it was discovered that 'collective welfare can be substituted for self-interest, even when the financial incentive structure might tend to reward the latter' (Lutzenhiser, 1993, 257).

As lease terms will not usually permit tenants to make technological changes, such as upgrading energy supply equipment, the abatement opportunities available to tenants are mainly confined to the behavioural, such as switching off lighting and computers. On the assumption that the landlord's CRC costs are passed on to tenants, this means that the tenants are incentivised to examine how they can use the property differently so as to reduce energy consumption. Given that tenants, on this assumption, are carrying both the CRC and energy costs there are no financial incentives for the landlord to think about abatement opportunities. It may be, however, that reputational factors stemming from position in the league table will be sufficiently important to landlords that they will consider technological improvements even though this carries a financial cost. This means that if tenants bear the CRC costs, there is the possibility of both forms of abatement opportunities being delivered. Predicting how the CRC scheme will incentivise energy reduction is complex however. Some have argued that passing on the CRC costs to tenants will not:

...address the key purpose of the CRC, namely to reduce carbon emissions. It could in fact lead to tenants cushioning the blow that CRC will deliver to energy-inefficient landlords. (Jayson and Redman, 2010)

Under existing leases, before the implementation of the CRC scheme, landlords were unlikely to have any significant abatement incentives (putting aside the possibility of corporate social responsibility). As energy use is usually paid for by the tenant (either directly or through the service charge) the landlord has, at least financially, no reason to spend money upgrading energy supply equipment or improving the infrastructure to provide for greater energy efficiency. There are some indications that capital values and

⁴ Land Securities, CRC energy efficiency scheme treating the customer fairly, Powerpoint presentation 26th October 2009. See also Department for Environment, Food and Rural Affairs 'How is the successful qualification of EPC organisations ensured?' (London, May 2007) p. 19.

rents are beginning to be impacted by the green credentials of a building, but the evidence is ambivalent, particularly as 'green buildings' tend to be the more modern and attractive ones in any case. In the new world of the CRC scheme, however, *if* the cost of participation remains with the landlord, the incentives shift. The landlord will now have a financial incentive to drive down energy use in the property (through the cost of carbon), as well as a reputational incentive. Under this model, however, the landlord has no lever – beyond the tenant's share of energy costs – to encourage behavioural changes by tenants. Some major landlords are, however, implementing programmes to encourage full tenant engagement. So, for example, to encourage tenants to enter into a 'memorandum of understanding' that incorporates 'green' goals, Land Securities operates an incentive scheme including free energy audits, training, awareness materials, and other environmental services (UNEP Finance Initiative Property Working Group, November 2009, 7).

There is, thus, a confusing picture. If CRC costs are placed on the tenant, this may encourage behavioural changes that result in energy use reduction but there is unlikely to be much opportunity for the tenant to make equipment changes. Depending on the strength of reputational drivers, the landlord may, however, choose to make technical improvements. If CRC costs are carried by the landlord, the landlord has both financial and reputational incentives to make technical improvements, but has no direct way of controlling energy use that is driven by behaviour. Under either model, the impact of the financial incentive will depend upon the price of carbon allowances. Once the initial, fixed price, phase is over this is an unknown although the Conservative – Liberal Democrat coalition government has stated that it plans to put a floor on the carbon price. Some predict that carbon allowances will be very expensive once the auction procedure starts, while others are more circumspect.

It is therefore clear that the impact of the CRC scheme will be affected by who will carry the costs of carbon allowances. There is no consensus within the property industry as to whether CRC costs should be borne by landlords, or passed on to tenants. A consultation organised by an industry working party early in 2010 showed substantial support for the costs of CRC allowances being subject to tenant contributions (Brown, 2010). In this vein, a number of major landlords, including Land Securities and British Land, have declared a desire to pass on the costs. It may be, however, that the logistical challenges involved in drafting for this (Dowden, 2009; Williams and Highmore, 2010) and the administrative costs involved in running internal trading schemes mean that many landlords choose to simply carry the cost themselves. This was the conclusion reached by Hermes early on. On the basis of a modelling exercise it conducted across 103 of its directly managed properties it found that during the fixed price phase it may be cheaper to simply write off the cost of carbon allowances rather than incur the administrative and legal costs of engaging with tenants.⁵ It has also been suggested that leases that enable

⁵ Property Industry Working Party (2009), "Press release accompanying launch of Guide". See also a Case Study of BAA in which BAA concluded that the rules and contractual arrangements necessary to run an internal trading scheme were too complex to justify the energy savings: NERA Economic Consulting and Enviro Consulting (April 2006), "Energy Efficiency and Trading Part II: Options for the Implementation of a New Mandatory UK Emissions Trading Scheme" London, p 45.

landlords to recover CRC contributions from tenants may be regarded as onerous, which would have a negative impact on the value of the landlord's interest (Jayson and Redman, 2010; Williams and Highmore, 2010). There has been a great deal of consultation and discussion amongst property experts in the hope that a common approach would emerge, which would enable standard lease clauses to be drafted, but for now there is simply no consensus on this.

The impact will also be influenced by the extent to which it is possible to achieve significant abatement through either technical strategies alone or behavioural changes alone. Although technical changes are critical to success, it is people that use the energy in buildings and the choices and behaviour of people and groups are also fundamentally important (Janda, 2009). This is often overlooked. A report prepared for DEFRA by NERA Economic Consulting and Enviro Consulting in 2006 states that in offices approximately two thirds of the *cost effective* abatement potential is available to the tenant as either investment measures (such as changing to energy efficient light bulbs) or behavioural ones (such as switching off lights and IT equipment) (NERA Economic Consulting and Enviro Consulting, 2006, 39). In the retail sector it was lower, but still almost half of the potential lay with tenants. Even behavioural changes are not, however, always available to occupiers. For example, sometimes a local retail outlet that forms part of a national chain will have its energy use (such as lighting) controlled centrally rather than at the local level. Whilst, again, the figures may not be terribly robust, they do show that to achieve significant energy abatement there must be both technical change *and* behavioural change. Under a 'landlord carries the cost of CRC' model, there is no incentive for behavioural change placed on the tenant, although it is likely that landlords will become much more actively concerned to work with tenants in relation to energy management. Paul Edwards, Head of Sustainability at Hammerson, comments:

Ultimately, we have no legal right to ask a tenant to turn their lights off, so we need to improve relationships and increase collaboration to ensure that everyone takes responsibility for their emissions. (Property Industry Working Party, 12 June 2009)

This may, in turn, promote a greater interest in the use of 'green leases' or memoranda of understanding which provide a leasehold structure that will facilitate and support the property being used in an environmentally sensitive manner. Current take-up is patchy and occupiers, in particular, appear resistant to them. It may be that the 'unknown' quality of green leases and memoranda will delay their widespread use but the behavioural change necessitated by the CRC scheme will likely lead to environmental management plans being adopted more widely on an informal basis. Over time, the market and occupiers may become more comfortable with the idea of formal commitments and leases may more regularly be sensitive to environmental concerns.

A More Co-operative Working Relationship?

One of the very real challenges in this field stems from the traditionally adversarial relationship between landlords and tenants. Tenants are often suspicious of landlord motives and distrustful in relation to expenditure. Tenants understandably wish to contain expenditure on their premises, and may have made careful considerations when deciding whether to take a property as to whether the costs of occupation will enable them to make

a business profit. Any unplanned increase in expenditure may put this at risk, and landlords are often perceived as opportunistically seeking to make money out of the provision of services to tenants rather than to provide them on a 'no profit, no loss' basis. In order to improve the energy performance of tenanted commercial property it is, however, desirable for there to be a genuinely co-operative working relationship between the landlord and tenant that goes beyond the simple provision of information about energy use and embraces joint endeavours to address both the technical and the social aspects of energy use. The importance of tenant engagement is illustrated by case studies conducted by Land Securities:

Land Securities cannot make significant reductions in building energy consumption, or reduce waste or improve recycling, without a high level of cooperation from the tenants. Its own case studies show that on their own, they can reduce energy usage in a building by maybe up to 10 per cent over 3 years through improved management, but with tenant buy-in they can achieve 15-20 per cent reductions in one year. (UNEP Finance Initiative Property Working Group, November 2009, 8)

Some commentators have suggested that the aim of the scheme in a landlord and tenant context 'is for the parties to collaborate to improve a building's energy efficiency' (Williams and Highmore, 2010). If this collaboration is the goal then it may well be best for the landlord to bear the costs on the basis that it then has the incentive both to carry out energy management improvements, and to work with tenants to reduce energy use. Yet one of the concerns expressed in relation to the CRC scheme is that operating it may generate yet further friction in an already fraught relationship, whether a scheme that passes on the costs to tenants is adopted (Jayson and Redman, 2010) or (as RICS prefer) the landlord bears the cost but attempts to improve energy performance by engaging in dialogue with tenants (RICS, 2010).

It is perhaps in recognition of this that the government decided to introduce a 'duty to co-operate'. The government paper that first flagged this duty suggested that this duty has the broad and ambitious goal of encouraging a deep and meaningful cooperative working between landlords and tenants:

... in order to encourage tenants to work with landlords Government is including an additional provision in the revised CRC Order which requires the tenant to cooperate with the landlord organisation for the purpose of complying with CRC. Government will not provide means by which to apportion costs and benefits of CRC in the revised CRC Order as this is a matter for the commercial arrangements between landlords and tenants. However our approach does not prevent landlords from passing through the costs and benefits of CRC to the tenants where current leases would allow this, and it would encourage landlords and tenants to revise and agree terms for leases accordingly. Our intention is that CRC will also encourage good practice such as sub-metering so that costs and benefits can be fairly distributed. (Department of Energy and Climate Change, 2009, 36)

The relevant paragraph of the CRC Scheme Order, para 65, is much more narrowly defined. It provides that the landlord may request the tenant to provide the landlord with

‘such reasonable assistance’ as the landlord requires to comply with the landlord’s obligations as a CRC participant.

There are various ways in which this phrase – such reasonable assistance – could be read. Narrowly it could mean assistance that is *necessary* or *essential* for the landlord to fulfil the registration, reporting and record keeping requirements. As landlords need to know how many allowances to buy to cover their tenants’ energy use, this means that tenants will need to provide certain information about energy use and meter readings, and perhaps access to the tenant’s premises (of course, landlords generally have fairly wide access rights under typical leases in any event). If interpreted this way, it is only a very limited form of co-operation that is required. Whether it would mean that a tenant has to allow a landlord to install better equipment for monitoring energy use is unclear. Alternatively, it could carry a wider meaning, referring to assistance that is *desirable* for the landlord to fulfil its obligations as a participant. Of course, one of the core requirements is for the landlord to purchase sufficient carbon allowances for that reporting year. In order to calculate as precisely as possible how much energy is likely to be used, and therefore how many allowances have to be purchased, the landlord will need to understand the tenant’s business and plans for the coming year. As the CRC Guide for Landlords and Tenants issued by the property industry working party points out, this might include knowing how the premises are occupied (how many occupiers, hours of occupation on weekdays and at weekends, occupier density), what the space is used for (for example, how much space is for storage, data centres etc), whether there is on-site catering, and whether the occupation and use of any space is likely to change (Property Industry Working Party, June 2009, 12). This kind of information is much more intrusive for the tenant, and difficult to provide in larger organisations. An unwilling tenant may argue that it is more than required by the statutory obligation. How it is construed by the judiciary will depend upon the extent to which they adopt a purposive approach to interpretation of the wording.

The widest interpretation of the statutory duty is to read it as providing assistance not simply with the registration, reporting and purchasing requirements placed on the CRC participant, but as including assistance with delivery of the goals of the CRC scheme, that is, carbon reduction. It is this meaning that delivers the ‘good practice’ referred to in the government consultation and goes furthest towards the overall aim of reducing the amount of energy consumed. But it is unlikely that the wording used in the CRC Scheme Order imposes such a wider duty on the tenant. Whilst flagged as a ‘duty to co-operate’, the level of co-operation required will probably turn out to be quite minimal.

Some Concluding Remarks

Although many within the property industry are critical of the fact that the CRC scheme has not been designed well for the tenanted commercial property sector - and it is undoubtedly the case that mapping it onto both existing and new leases is intensely challenging - it is likely to contribute to significant carbon emission reduction in the commercial built environment.

There are many uncertainties. The future price of carbon is one of the most significant. The degree of effort required to promote and sustain behavioural change in shared property is unknown. The extent to which the property industry may be willing to embrace change to the net lease is unknown. The challenges involved with drafting and using CRC provisions in leases are known, but the industry is divided as to whether to go down the route of not providing for the CRC scheme in leasehold drafting, providing for it within general service charge clauses, or designing tailored CRC provisions. The impact of CRC provisions on capital values and rent is unknown.

Notwithstanding all of this uncertainty the implementation of the CRC scheme will, however, have an impact and will lead to energy reduction. The fact that the number of CRC allowances available for purchase each year will be reduced over time will drive down energy use. In addition, whichever model is adopted for dealing with CRC costs – whether it is the landlord or tenant who pays – one of the parties will have at least some incentive to reduce its energy use. The full abatement potential can be realised only if there are incentives to pursue both the technical and the social abatement opportunities. This (combined with the ‘relative’ ease of adoption and operation) is why some of the major landlords are planning to carry the CRC costs themselves whilst engaging actively with their occupiers, and providing incentives for such co-operative working. Amidst all of the uncertainty and market adjustment, however, it is clear that the CRC scheme has grabbed the attention of the property industry. There are many new conversations now happening, about technical solutions, about legal solutions, and about behavioural change. Grabbing ‘senior management attention’ has, from the early days of policy on carbon trading, been an important aim. The CRC scheme has certainly proved a success from this perspective by putting carbon reduction on the management radar.

Biggart, N. W. and Lutzenhiser, L. (2007), "Economic Sociology and the Social Problem of Energy Inefficiency", *The American Behavioral Scientist*, Vol. 50 No. 8, pp. 1070-1087

Brown, P. (2010), "forthcoming", *Estates Gazette*,

Burges Salmon and Ecofys (2008), "Carbon Reduction Commitment - Organisational Structures" Department for Environment, Food and Rural Affairs, London

Department for Environment, Food and Rural Affairs (2008), "Implementation proposals for the Carbon Reduction Commitment; Report on the public consultation (June 2007) and Government's policy decisions" London

Department of Energy and Climate Change (2009), "Consultation on the Draft Order to Implement the Carbon Reduction Commitment; Government Response and Policy Decisions" London

Department of Energy and Climate Change (2010), "Final Impact Assessment on the Order to implement the CRC Energy Efficiency Scheme" London

Dowden, M. (2009), "Foot(print)ing the bill", *New Law Journal*, Vol. 159 No. 7386, p. 1314

Edwards, C. and Krendel, P. (2007), *Institutional Leases in the 21st Century*, EG Books, London

Investment Property Forum (2005), "The Size and Structure of the UK Property Market"

- Janda, K. B. (2009), "Buildings Don't Use Energy: People Do" 26th Conference on Passive and Low Energy Architecture, Quebec City, Canada
- Jayson, P. and Redman, N. (2010), "Market forces should be your guide", *Estates Gazette*, No. 1 May 2010 p. 88
- Lovell, J. (30 April 2010), "CRC refreshes lease structures other regulations cannot reach" *Property Week*
- Lutzenhiser, L. (1993), "Social and Behavioral Aspects of Energy Use", *Annual Review of Energy and the Environment*, Vol. 18 pp. 247-289
- NERA Economic Consulting and Enviro Consulting (2006), "Energy Efficiency and Trading Part II: Options for the Implementation of a New Mandatory UK Emissions Trading Scheme" London
- Property Industry Working Party (June 2009), "The Carbon Reduction Commitment: A guide for landlords and tenants"
- Property Industry Working Party (12 June 2009), "Press release 'Carbon Reduction Commitment Guide Launched'"
- Property Industry Working Party (2009), "Press release accompanying launch of Guide"
- Punshon, M. and King, V. (2010), "A thorough clean-up can save money", *Estates Gazette*, No. 29 May 2010 p. 83
- RICS (4 June 2009), "Response to Department of Energy and Climate Change Carbon Reduction Commitment Consultation"
- RICS (2010), "RICS response to consultation on the treatment of the CRC in the context of Landlord and Tenant Relationships, attached to letter dated 5 February 2010 addressed to Patrick Brown of the CRC Working Party, available from www.rics.org"
- The Carbon Trust (2009), "Building the future, today"
- The Carbon Trust (2005), "The UK Climate Change Programme: Potential evolution for business and the public sector"
- UNEP Finance Initiative Property Working Group (November 2009), "Owner-Tenant Engagement in Responsible Property Investing"
- Williams, P. and Eddington, C. (2009), "A lack of direction", *Estates Gazette*, No. 7 November 2009, pp. 200-201
- Williams, P. and Highmore, S. (2010), "forthcoming", *Estates Gazette*,

A Commercial Right to Buy?

Jill Morgan

School of Law,
Swansea University, Singleton Park, Swansea SA2 8PP,
United Kingdom

Email: J.Morgan@Swansea.ac.uk

Abstract

Leasehold enfranchisement was put in place by the Leasehold Reform Act 1967. The Act entitles long leaseholders of houses at any time during the original terms of their leases to enfranchise, i.e. to 'acquire on fair terms' the freehold reversions of their properties, or to extend their leasehold terms for a further fifty years. Given that the right to enfranchise applies only to houses, it appeared initially to be of only limited application. Recent years, however, have witnessed a relaxation of some of the more stringent qualifying requirements (including a low rent test and rateable value limits) which restricted the right to enfranchise to certain types of property and certain leaseholders. For the purposes of this paper, the most significant change has been the removal of the residence qualification. Not only has this resulted in the extension of the right to enfranchise to corporate leaseholders (and not just individuals as was previously the case) but it also means that the definition of a house can now include certain properties used for commercial purposes.

To appreciate the significance of these developments, it is useful first to consider why the residence requirement was removed. The main part of the paper – which focuses on the impact that its removal has had on the meaning of the word 'house' - considers the decisions of the House of Lords in *Boss Holdings Ltd v Grosvenor West End Properties* (2008), and the Court of Appeal in *Grosvenor Estates Ltd v Prospect Estates Ltd* (2009) and *Hosebay Ltd v Day & Anor* (2009) (a County Court case). The penultimate section considers the extension of the right to enfranchise might be challenged on human rights grounds. The paper concludes that enfranchisement should have been limited to allowing 'occupiers' to move their 'genuine' homes into freehold tenure.

Keywords: landlord and tenant law, leasehold enfranchisement, meaning of 'house', residence qualification.

1 Introduction

The Leasehold Reform Act 1967 gives leaseholders of houses whose leases were granted for a term exceeding 21 years¹ the right to enfranchise, i.e. to acquire the freehold title. Until 2002, a residence qualification - which made enfranchisement conditional upon occupation of the house as the leaseholder's 'only or main' residence' for the requisite period - meant that non-resident leaseholders were unable to enfranchise. However, the Commonhold and Leasehold Reform Act 2002 abolished the residence qualification, replacing it with an ownership test, i.e. a requirement that at the time when leaseholders give notice of their intention to enfranchise, they have been tenants of the house for the last two years.² Not only has the removal of the residence qualification resulted in the extension of the right to enfranchise to company tenants but it also means that, as indicated by recent cases, the definition of a house can now include certain types of commercial property.

The focus of this paper is the impact which the removal of the residence qualification has had on what is meant by a 'house'. This will involve consideration of a number of cases, including the decision of the House of Lords in *Boss Holdings Ltd v Grosvenor West End Properties*,³ the Court of Appeal's decision in *Prospect Estates Ltd v Grosvenor Estate Belgravia*⁴ and *Hosebay Ltd v Day & Anor*⁵ (a County Court case). The paper concludes that enfranchisement should have been limited to allowing *occupiers* to move their houses into a freehold tenure and in any event should not extend beyond 'genuine' homes.

2 The residence qualification

The 1967 Act provided (in its original guise) that a tenant would have the right to enfranchise only if, at the time he served his notice of intention to enfranchise, he had been 'occupying the house as his only or main residence' for the last five years or for periods amounting to five years in the last ten. The imposition of what was seen by some to be a 'lengthy' residential qualification was intended to prevent exploitation of the right to enfranchise by purchasers of 'fag end' leases, i.e. long leases with only a few years left to run. There was little evidence of such a practice outside of London, and the Housing Act 1980 reduced the qualifying period to three years.

By 2002, however, concerns were being voiced that 'deserving' categories of leaseholders who had bought their houses primarily 'for personal rather than investment reasons' were being prevented from exercising their rights simply because they could not meet the residence qualification.⁶ The examples given included those who had sub-let

¹ Leasehold Reform Act 1967, s. 3(1).

² Leasehold Reform Act 1967, s.1(1)(b).

³ [2008] 1 WLR 289.

⁴ [2009] 2 EG 86.

⁵ [2009] PLSCS 318.

⁶ Hansard HC Debs, vol 377, col 434 (8 Jan 2002).

the whole of the premises, or who occupied their houses as second homes (e.g. their 'main home' was 'a house in the country' but they also had 'a pied-a-terre in town'), the 'large numbers of expatriate owners' with properties in central London or, more generally, those leaseholders who had bought or retained their houses with the intention of living in them when they returned from working abroad or elsewhere in the UK. Also excluded were leaseholders who were living in the house as their main home but had leased it through a company because of the landlord's policy (specifically designed to block the right to enfranchise), of letting only to companies.⁷ In the Government's view, justification for the removal of the test – which was said to be 'open to manipulation and abuse' – stemmed from the impossibility of devising 'a fair, workable and unambiguous qualifying test that relies on such a slippery concept as residence'.

At the same time, there was a recognition that 'a lease is inherently unsuited as a form of tenure for the sale of homes (whether as flats or houses)' (Clarke, 2006). In 1967, it was accepted that if freeholders were to be deprived of their assets, this should only be by a resident occupier and the house should be the occupier's home. By 2002, there had been a shift. Now, the long leasehold was considered an unsuitable vehicle for home ownership; the only options should be freehold, commonhold or terms of 21 years or less at a rack rent. It was felt that all residential long leaseholders should be able to enfranchise, but at a market price.

That Parliament did not intend corporate leaseholders to be entitled to enfranchise is discernible from the retention of residence qualification where the tenancy of the house is a business tenancy to which Part 2 of the Landlord and Tenant Act 1954 Act applies.⁸ Unfortunately it overlooked the possibility of corporate leaseholders sub-letting their houses or assigning their business to an associated company so that they were no longer occupying the premises as required by the 1954 Act).

Fears that the abolition of the residence requirement could 'open the door to investors rather than individual homeowners' and allow the 'possibility of short-term windfall gains by speculators'⁹ prompted the replacement of the residence qualification with a requirement that, before the right to enfranchise or to extend the lease can be exercised, the leaseholder must have held the lease for a period of at least two years at the relevant time. The purpose behind this measure was to deter investors or speculators from cashing in on a right which was intended to protect natural persons from loss or exploitation over the reducing value of their homes as a wasting asset. However, the abolition of the residence qualification has allowed non-resident leaseholders, including limited companies, to benefit from enfranchisement. Admittedly, this will also have impeded the opportunities for 'endless arguments over interpretation' which the Government wished to avoid.¹⁰ Yet, what it seems to have done instead is open the door to 'endless arguments over interpretation' as to the subject-matter of the right to enfranchise, namely what is meant by a 'house'.

⁷ Hansard HL Debs, vol 626, cols 920-1 (5 Jul 2001).

⁸ Leasehold Reform Act 1967, s.1(1ZC) and (1B).

⁹ Hansard HC Debs vol 377, col 434 (8 Jan 2002).

¹⁰ Hansard HL Debs, vol ???, col 462 (29 Jan 2001).

3 The meaning of ‘house’

As indicated above, the right to enfranchise is conferred on the long leaseholder of a house.¹¹ By s.2(1) of the 1967 Act a ‘house’ is defined as including ‘any building designed or adapted for living in and reasonably so called...’. In this section of the paper, further consideration is given to the two limbs of the definition, i.e. the meaning of the terms ‘designed or adapted for living in’ and ‘reasonably so called’.

Until the amendments brought about by the Commonhold and Leasehold Reform Act 2002, the existence of the residence requirement meant that the first limb - that a building will only be a ‘house’ if it was ‘designed or adapted for living in’ at the date when the tenant’s notice is served – presented little difficulty ‘because the tenant had to be living in it’ (Radevsky, 2006). In *Malekshad v Howard de Walden Estates Ltd*¹² Lord Millett went so far as to say that the requirement was ‘self-explanatory’. The removal of the residence qualification has muddled the waters, however.

The meaning of ‘designed or adapted for living in’ recently came under the scrutiny of the House of Lords in *Boss Holdings Ltd v Grosvenor West End Properties Ltd*.¹³ *Boss* was one of two appeals heard simultaneously in the Court of Appeal, the other being *Mallet & Sons (Antiques) Ltd v Grosvenor*.¹⁴ Both cases concerned claims to enfranchise made by companies. In *Boss*, the House had to consider the status of a six-storey terraced house which had been built in the 1740s as a single residence and used as such until around 1942 but later fitted out for separate residential and commercial use. From 1990, the three lower floors (which since 1946 had been occupied by a dress-making business) were left vacant. The residential use of the three upper floors ended in the late 1990s and by 2003 the accommodation had been stripped back to the basic structure with many of the floor boards and ceilings and much of the plaster having been removed. The property as a whole had become gravely dilapidated and was not physically fit for immediate residential occupation. The property in *Mallet* had also been built as a residence in the early eighteenth century but since the 1950s, all of it (except for the caretaker’s flat) had been given over to commercial use.

At first instance and in the Court of Appeal, it was held that the properties were not ‘designed or adapted for living in’ for the purpose of s.2(1). The Court of Appeal held that the test was a simple one. The true question was ‘whether the premises (or at least a substantial part) viewed as at the moment when the notice is given, were designed or adapted for living in?’ Is actual and immediate residence the purpose of the design or adaptation at the time of the notice?¹⁵ The House of Lords disagreed. Lord Neuberger,

¹¹ Leasehold Reform Act 1967, s.1(1) .

¹² [2003] 1 All ER 193.

¹³ [2008] 1 WLR 289.

¹⁴ [2006] 1 WLR 2848.

¹⁵ *ibid*, para 18 (Laws LJ).

who gave the only substantive speech, stated that the word ‘designed’, being a past participle, was concerned with the past while ‘adapted’ referred to the present. Accordingly, the phrase ‘designed or adapted for living in’ required a consideration of the property as initially built and the purpose for which it was originally designed, and then whether work had subsequently been done to the property so that the original design had been changed. It was common ground in the present case that, when it had been built in the 1700s, the building had been designed for living in and nothing which had happened subsequently – including its current state of internal dilapidation and unsuitability for immediate occupation – changed that fact. The upper three floors were and remained designed to be lived in, and the lower three floors appeared to be structurally laid out substantially as they were when the property was in single residential occupation. There was no requirement that the property actually be in such a physical state that it could now be lived in. That the property in question had not been occupied for a number of years and had become internally dilapidated and incapable of beneficial occupation did not detract from the fact that it had been designed for living in when it was first built.

The appeal in *Mallett* was settled before it came to be heard in the House of Lords and although it was unnecessary to decide the point, Lord Neuberger then flagged up the possibility that premises which had been designed for living in but subsequently adapted wholly for some other non-residential purpose (e.g. office accommodation) might remain enfranchiseable. He explained:

‘As a matter of literal language, such a property would be a house, because “designed” and “adapted” appear to be alternative qualifying requirements. At least at first sight, such a conclusion seems surprising, so there is obvious attraction in implying a qualification that, if a property has been, and remains adapted for a purpose other than living in, the tenant cannot rely upon the fact that it was originally designed for living in. However, a term is not easily implied into a statute, and further reflection suggests that the literal meaning of the words is not as surprising as it may first appear, particularly bearing in mind the existence of the [original] residence requirement in section 1(1) of the original Act.’

Turning to the second limb of the definition, even if it is ‘designed or adapted for living in’, a building is not a ‘house’ unless it can be ‘reasonably so called’. In the past, this has been the most difficult part of the definition to dissect, depending as it does on a subjective judgment of what a house comprises. The term was no doubt intended to limit the definition of a house which would otherwise apply to ‘many buildings which are designed or adapted for living in but are not houses in the popular sense of the word [such as] hotels, local authority hostels for homeless families, old peoples’ homes, boarding schools, university and colleges’ halls of residence’.¹⁶

It is clearly established that a building can reasonably be called a house even though part of it has been constructed as, or converted into, business premises.¹⁷ In *Tandon v*

¹⁶ *Ashbridge Investments Ltd v Minister of Housing and Local Government* [1965] 1 WLR 1320 (Lord Denning).

¹⁷ *Lake v Bennett* [1970] 1 QB 663.

Trustees of Spurgeons Homes,¹⁸ the House of Lords decided, by a majority of three to two (Lords Wilberforce and Fraser dissenting), that a building which consisted of a purpose-built shop with living accommodation above (comprising 25 per cent residential accommodation and 75 per cent business use) was not excluded from being a house. The factors which would be relevant in determining the character of the building and deciding therefore whether it was a house ‘reasonably so called’ would include its use at the date of the application to enfranchise, its nature and history, the terms of the lease, the proportion of the residential use to the non-residential use and their relative values. Lord Fraser felt that the main element in the character of a building is its appearance and that partial conversion to non-residential use would be relevant only in so far as it affected its present appearance. The setting in which the building is placed might also be significant, so that:

‘A building in a residential street is more likely to be reasonably called a house than the same building would be if it were in a busy street in the commercial area of a town’.¹⁹

Their Lordships held that a building which can reasonably be described as a house will be a house for the purposes of the Act even if it could also reasonably be described as something else. Lord Roskill (who gave the leading speech) stated that if a building is designed or adapted for occupation as a residence, then only in ‘exceptional circumstances’ would a judge be justified in holding that ‘it could not reasonably be called a house’. Those circumstances would have to be such that ‘nobody could reasonably call the building a house’.

‘Exceptional circumstances’ were regarded as existing in *Prospect Estates Ltd v Grosvenor Estate Belgravia*.²⁰ It was accepted that the building – a flat-fronted terraced early Victorian house – had been ‘designed for living in’ and the main question therefore was whether it was a house ‘reasonably so called’. Although some additions had been made since its construction, it still looked like a house from the outside. However, the user of the building accorded with the lease which specified that one floor was to be used as private residential accommodation but the remainder (comprising 88.5% of the floor space) could only be used as business or professional offices. The trial judge viewed the property and came to the conclusion that the essential character of the building was still that of a house. He considered that the

‘overwhelmingly significant factor in determining whether or not the building is ‘a house reasonably so called’ is that the building was designed for living in and that its structure and appearance have (largely) remained unchanged’.

Although only a small amount of the floor space was actually in residential use, he applied the test in *Tandon v Trustees of Spurgeon Homes* that the circumstances would have to be such that ‘nobody could reasonably call the building a house’ for a judge to

¹⁸ [1982] AC 755.

¹⁹ [1982] AC 755, 762.

²⁰ [2009] 1 WLR 1313.

hold that it was not a house. In overturning the decision, the Court of Appeal held that all the circumstances which affected the character of the building must be taken into account and not just physical ones. The judge had not given sufficient weight to the terms of the lease, the actual use of the premises, and the relative proportions of the mixed uses at the date of the claim. Had he done so, he would have had to conclude that it was not then reasonable to call the building a 'house' because of the 'prescribed and predominant office use in accordance with the lease'.²¹

4 The significance of 'living in'

Both limbs of s.2(1) of the 1967 Act fell to be considered in the recent County Court case of *Hosebay Ltd v Day*.²² Three Victorian terraced houses had been converted into individual rooms with self-catering facilities which were used by the leaseholder company for its business of providing short stay accommodation for tourists and visitors to London. It avoided the residence qualification retained by s.1(1B) by setting up another company to which it granted underleases of the properties, thereby creating a situation in which it was not in occupation for the purposes of its business and, therefore, not subject to the 1954 Act. The leases of two of the properties provided that they were to be used as 'high class residential flatlets' and that the appearance of their exteriors was to be kept 'as that of a private dwellinghouse'. The lease of the third property stipulated that it should be used either as 'a single family residence' or as 'a high class furnished property for accommodation not more than twenty persons...' The freeholders challenged the tenant's claim to enfranchise on the basis that the buildings were neither 'designed or adapted for living in' nor houses 'reasonably so called'.

The freeholder argued - and the judge agreed - that contrary to Lord Neuberger's speculation in *Boss*, the test of whether a building is 'designed or adapted for living in' is not determined by the original design purpose of the building. As the judge pointed out, such a reading of the provision would mean that 'this particular qualification would have become permanently fixed in the long distant past' even though 'the focus of all the qualifications under the Act seems to be on the situation at the date of the tenant's notice'.²³ Here, the premises were clearly designed for living in.

The next issue therefore was the effect of the adaptations. In the judge's view, the words 'living in' connote occupation with some degree of permanence, and not merely on a transient basis. Effectively, they entail 'using the property as some kind of a 'home', as contrasted with merely a convenient place to meet an immediate human need for shelter and sleep'.²⁴ She accepted that the buildings were configured in such a way as to make it possible to live in them and that it was only their furnishings which currently made them suitable only for 'staying in' and not for 'living in'.²⁵ But, as she stated, the meaning of

²¹ *ibid*, para 19.

²² [2009] PLSCS 318.

²³ *ibid*, para ???

²⁴ *ibid*, para 76.

²⁵ *ibid*, para 84.

‘designed or adapted for living in’ is concerned with ‘the design or adaptation of the building itself, i.e. its intrinsic physical construction and layout, and the use or purpose at which that is aimed, viewed objectively’ and not ‘the equipping for furnishing of the building’.²⁶

The finding that each building was designed or adapted for living in meant that it would take exceptional circumstances to conclude that each building was not a house reasonably so called. The judge rejected the freeholder’s argument that the decision in *Prospect* now imported a test of user. The crucial point in *Prospect*, she explained, was the lease prohibited the use of almost 90% of the building for residential purposes. In the instant case, not only did the buildings look like houses and contain all the facilities which houses would be expected to have, but none of the leases prohibited their being used for residential use. Indeed, such use was prescribed. She concluded that the removal of the residence qualification has simply removed the requirement for ‘a personal qualifying condition to be met by an *applicant*’; it has not changed or affected ‘the nature of the *property* which qualifies to be the subject of the rights under the Act’. Thus,

‘whilst non-residents, and even investors or speculators may now be able to take advantage of rights conferred by the Act, their investment or speculation still has to be in properties which are ‘residences’ or ‘homes’ in themselves’.²⁷

5 Some issues arising

The cases considered above clearly illustrate the difficulties posed by the fact that

‘...the functions of buildings change. What started as a single dwelling-house 100 years ago may today be an office. What began life as a Victorian warehouse may be converted into a unit comprising modernised flats’ (Yates and Hawkins, 1986).

Although courts at all levels have tried to disentangle the knotty questions as to what is meant by a ‘house’, the situation is far from clear. Lord Neuberger’s suggestion that ‘designed’ and ‘adapted’ appear to be alternative qualifying requirements has an appealing simplicity. It does not mean that the nature of the building would have become fixed at some time in the past because the second limb of the definition must also be fulfilled. Thus, a building designed and built as a house but subsequently adapted as office accommodation cannot reasonably be called a house and cannot therefore be the subject of a successful claim to enfranchise. The same can be said of the examples given by Lord Denning As Law LJ pointed out in *Boss*, ‘Premises designed or adapted for a different use altogether might in various sets of circumstances with a little ingenuity be lived in perfectly readily, but they would not be a house’.²⁸ The nature of any adaptations is also relevant. Unauthorised alterations carried out to make the buildings look more

²⁶ *ibid*, para 85.

²⁷ *ibid*, para 77.

²⁸ (*Boss* CA, para 18)

like houses internally should be disregarded²⁹ as should those for which any necessary planning consents have not been obtained.

As regards the requirement that the building must be capable of being called a house, the Court of Appeal's decision in *Prospect Holdings* suggests there will often be circumstances which militate against the possibility of enfranchisement where 'commercial' buildings are concerned. For example, it may not be reasonable to call a building a 'house' where the lease describes the premises as let for business purposes or prohibits residential user, or clearly contemplates user exclusively for a particular, non-residential purpose. It may be otherwise if – despite the restrictions imposed by the lease – the landlord by an express or implied consent (such as a prolonged acceptance of the rent with knowledge of the change), allows all or part of the building to be put to residential use. If the freeholder permits part of the building to be used for residential purposes, it must then be a matter of degree – as in *Prospect Holdings* – as to whether the element of business use is such that the building can still be regarded as a house 'reasonably so-called'. Questions may arise if the residential use is seasonal (as could be the case with tourist accommodation similar to that in *Hoseday*).

6 A human rights perspective

In *Earl Cadogan v Sportelli*,³⁰ Carnwath LJ noted that the various statutes which, since 1967, have conferred rights of enfranchisement upon owners of long leases have resulted in legislation which is 'unusually complex, reflecting the draftsman's attempts to realise Parliament's not always consistent objectives within the intricacies of landlord and tenant law'. This inconsistency in objectives makes it difficult to formulate a challenge on human rights grounds.

In *James v United Kingdom*,³¹ the trustees of the Grosvenor Estate claimed that the Leasehold Reform Act 1967 infringed their rights under Art 1 of the First Protocol of the European Convention on Human Rights. This provides that 'every natural or legal person is entitled to the peaceful enjoyment of his possessions'. However, it allows for such deprivation where it is 'in the public interest' and it maintains the right of a State 'to enforce such laws as it deems necessary to control the use of property in accordance with the general interest...'

By the time it initiated proceedings, the Estate had been obliged by the 1967 Act to sell the freeholds of 80 houses in Belgravia. The trustees argued that the long leasehold system was unfair and that a long leaseholder had no 'moral entitlement' to ownership merely by reason of occupying a house which had been 'built, repaired or improved by previous [leaseholders] in accordance with the contractual terms of a lease'.³² However,

²⁹ *Colliers CRE plc v The Portman Estate Nominees (One) Ltd* (unreported, 2006, Central London County Court).

³⁰ [2007] EWCA Civ 1042, para 1.

³¹ (1986) 8 EHRR 123.

³² *ibid*, para 49.

the European Court of Human Rights (ECtHR) held that although the effect of the 1967 Act was to deprive the applicants of their possessions, a taking of property in pursuance of legitimate social, economic or other policies may be ‘in the public interest’, even if the community at large has no direct use and enjoyment of the property taken.³³ The ECtHR referred to the ‘margin of appreciation’ which recognises that their direct knowledge of their society and its needs puts national authorities in a better position than international judges to appreciate what is in the public interest. It is for national authorities ‘to make the initial assessment both of the existence of a problem of public concern warranting measures of deprivation of property and of the remedial action to be taken’.³⁴ Given that modern societies regard the ‘housing of the population to be a prime social need, the regulation of which cannot entirely be left to the play of market forces’, the margin of appreciation was wide enough to cover legislation aimed at

‘securing greater social justice in the sphere of people’s homes, even where such legislation interferes with existing contractual relations between private parties and confers no direct benefit on the State or the community at large’.³⁵

The ECtHR noted that leasehold reform in England and Wales had been a matter of public concern for almost a century and that, when the 1967 Act was passed, enfranchisement was accepted as a principle by all the major political parties, although they expressed different views as to how it should be implemented. The applicants’ main criticisms of the substantive provisions of the legislation had been voiced at the time and fully debated in Parliament before being rejected. As the Court explained:

‘The aim of the 1967 Act, as spelt out in the 1966 White Paper, was to right the injustice which was felt to be caused to occupying leaseholders by the operation of the long leasehold system of tenure. The Act was designed to reform the existing law, said to be ‘inequitable to the leaseholder’, and to give effect to what was described as the occupying leaseholder’s ‘moral entitlement’ to ownership of the house’.³⁶

The ECtHR felt that the view taken by Parliament as to the leaseholder’s ‘moral entitlement’ to ownership of the house applied equally to the better off, middle-class leaseholders on the trustees’ estate in Belgravia (who on the whole could not be classified as needy or deserving of protection) as to the leaseholders of far more modest housing who had been the originally intended beneficiaries of the legislation. The policy decision by Parliament - that landlords affected by the legislation should be deprived of the unjust enrichment which they would otherwise enjoy at the end of the term even though a number of ‘undeserving’ leaseholders could make ‘windfall profits’ - was not so unreasonable as to be outside the State’s margin of appreciation.

³³ *ibid*, para 45.

³⁴ *ibid*, para 46.

³⁵ *ibid*, para 47.

³⁶ *ibid*, para 47.

From the ‘moral entitlement’ perspective it may be argued that the removal of the residence requirement exposes the legislation to a fresh challenge under the Human Rights Act 1998. Although Parliament instituted the ownership test to deter ‘investors’ and ‘speculators’, it appears to have overlooked the fact that removal of the residence qualification would open up the right to enfranchise to corporate leaseholders. There was no full debate on the provision such as preceded the 1967 Act. Further, the consequent interpretation of the meaning of ‘house’ to embrace certain types of commercial property may be said to have gone beyond the original aim of the 1967 Act, i.e. ‘to right the injustice which was felt to be caused to occupying leaseholders by the operation of the long leasehold system of tenure’.³⁷ While occupying leaseholders, regardless of their means, may have a moral entitlement to their homes, the same cannot be said of corporate leaseholders.

If, however, the purpose of the 1967 Act is regarded as the prevention of unjust enrichment by the freeholder, a human rights challenge is likely to prove fruitless as a freeholder will, in the absence of enfranchisement, enjoy the same benefits at the end of the leasehold term whether the leaseholder is an individual who has occupied the house as his or her home or a corporate body. This is also the case if the policy underpinning the enfranchisement legislation is perceived as an attempt to address the unsatisfactory nature of long leaseholds and to promote more appropriate vehicles for home ownership. In *Howard de Walden Estates Ltd v. Aggio*,³⁸ the House of Lords referred to the problems posed by the lease as a wasting asset and in *Earl Cadogan v Pitts*,³⁹ it pointed to Parliament’s conclusion that the residence qualification created difficulties in the enfranchisement of large blocks of flats where there was a rapid turnover of some of the flats.

In *Pitts*, Lord Walker described the decision of the ECtHR in *James v United Kingdom* as ‘an insuperable obstacle’ to a claim by the landlord that the compensation provisions of the Leasehold Reform Act 1967 and the Leasehold Reform Housing and Urban Development Act 1993 (which gives long leaseholders of flats a right collectively to enfranchise) were inconsistent with the Human Rights Act. Acknowledging that the scope of leasehold enfranchisement and associated rights had ‘increased greatly’ since the 1967 Act, (‘especially with the removal of the requirement for occupation by the leaseholder’⁴⁰), he nonetheless reiterated the point that it is for Parliament to decide on policies to remedy social injustice and that the ECtHR will accept its conclusions on social policy unless they are manifestly unreasonable.

To summarise, it seems unlikely that the removal of the residence qualification exposes the legislation to a fresh challenge.

³⁷ *ibid*, para 47.

³⁸ [2009] 1 AC 39.

³⁹ [2009] 2 WLR 12.

⁴⁰

Conclusion

The retention of residence qualification where the tenancy of the house is a business tenancy to which Part 2 of the Landlord and Tenant Act 1954 Act applies is evidence that Parliament did not intend corporate leaseholders to be entitled to enfranchise. Had Parliament retained the residence qualification for all leaseholders seeking to enfranchise

The (arguably unnecessary) wholesale removal of the residence qualification was intended to simplify the enfranchisement process and to allow what the Government regarded as ‘deserving’ leaseholders, such as second home-owners, to enfranchise. However, its - seemingly unintended - consequences have also included the extension of the right to enfranchise to corporate leaseholders and the beginning of what may well be a long line of cases centring on the meaning of a ‘house’. In the last edition of *Hague on Enfranchisement*, the authors noted that ‘although the 1967 Act was passed primarily to meet the problems facing ordinary householders in South Wales and other areas where the long leasehold system was widespread’ its principal beneficiaries were

‘relatively wealthy purchasers of short residues of long tenancies of houses in Belgravia, Chelsea, Kensington and similar expensive areas of London, just under the £400 Greater London rateable value limit which then applied’.⁴¹

Now, it appears that further beneficiaries of the right to enfranchise will be companies (some of which may not even be based in the UK) which will - in certain circumstances - be able to buy the freeholds of premises which have a significant commercial element.

References

- Clarke, D., (2006), ‘Long residential leases: future directions’, in S. Bright (ed) *Landlord and Tenant Law: Past present and Future* (Hart Publishing, Oxford).
 Radevsky, A., (2006) ‘What is a “house”? 10(6) *L & T Rev* 181-3.

⁴¹ ????. Following the unsuccessful ECHR challenge by the Grosvenor Estate in 1986, it was reported that ‘Belgravia and Mayfair residents, not best known for their financial privations’ had bought the freeholds of their houses and then sold them for £500,000 or more ‘to realise an overnight profit of as much as £250,000’: *Financial Times*, 22 February 1986.

Judges, child trespassers and occupiers' liability in the built environment

Luke Bennett ¹

¹ Department of the Built Environment,
Sheffield Hallam University, Sheffield, S1 1WB,
United Kingdom

Email: l.e.bennett@shu.ac.uk

Abstract:

This paper analyses how the doctrine of occupiers' liability for the safety of child trespassers has been developed in English Law over the last 100 years by applying Pierre Bourdieu's theorising on the operation of the 'juridical field' (Bourdieu 1987). The analysis traces the evolution across the 'no duty' dictum of *Robert Addie & Sons (Collieries) –v- Dumbreck* (1929)¹, the 'duty of humanity' advanced in *British Railways Board –v- Herrington* (1972)², the qualified duty of care introduced by the Occupiers Liability Act 1984 following the Law Commission, *Report on Liability for Damage or Injury to Trespassers*, Cmnd 6428 (1976) and the subsequent interpretation of that duty in case law such as *Tomlinson –v- Congleton Borough Council* (2002)³ and *Mann v Northern Electric Distribution Ltd* (2010)⁴. The paper analyses the ways in which the judiciary's operation as an 'interpretive community' (Fish, 1980; 1989) has steered this evolution and the ways in which external, supra-juridical, factors have also influenced how occupiers' liability law has been interpreted and applied. The analysis is based upon a doctrinal and narrative analysis of the key cases, policy and legislative instalments. This analysis includes assessment of how child trespassers have been depicted and their motivations interpreted by English judges; evaluation of the impact of the

¹ (1929) AC 358, HL.

² (1972) AC 877, HL.

³ (2002) EWCA Civ 309, HL

⁴ (2010) EWCA Civ 141, CA.

1984 Act; the influence of parallel developments in other common law jurisdictions and the interaction of occupiers' liability law with other 'safety' based legislation. The analysis shows how within the broad policy and legal framework set by legislators, the common law and the 1957 and 1984 Occupiers' Liability Acts, the courts have marked out (and pragmatically adjusted) the limits of the 'reasonable' safety provision required of occupiers for lawful visitors and trespassers.

Keywords:

Occupiers' Liability – Trespass – Safety – Estate Management – Interpretive Communities- Pierre Bourdieu – Sociology of Law

1. Introduction

Does an occupier owe any duty to safeguard child trespassers who may come upon his land or structures? This question is easily stated – but has taxed the courts for over 100 years. This paper considers how the appellate courts in the United Kingdom have sought to find an answer to this question.

This study forms part of a wider project seeking to understand how various stakeholders (e.g. occupiers, land agents, trespassers and regulators) each seek stable and workable interpretations on what the law requires of them in this area of law (Bennett & Crowe 2008; Bennett 2009; Bennett & Gibbeson 2010). For, in this arena, each lay 'interpretive community' (Fish 1980; 1989) starts with the same awareness of the broad principles – that an occupier owes a duty of care towards lawful visitors and something similar to foreseeable trespassers in proximity to foreseeable danger – and then must develop and apply its own pragmatic and approximate interpretation of what that means in practical terms for its community.

To support this wider project this paper seeks to achieve two outcomes. First, to provide a concise summary of the doctrinal evolution of the law in this area. And secondly, to consider how the appellate judiciary have themselves operated as one such 'interpretive community' in developing their own interpretation of how these legal principles should properly be applied in cases before them. This second layer of analysis will draw upon, and seek to empirically illustrate, Pierre Bourdieu's theorising of the operation of the 'juridical field' (Bourdieu 1987). In each case the analysis is confined to the 'problem' of child trespass to make the task

manageable and because the child trespass has proved a particularly active and emotive part of the wider field of occupier liability for trespassers *per se*.

This paper's analysis will specifically consider the following questions:

- Has there been a neat linear evolution in the law of occupiers' liability as regards child trespassers over the last 100 years? – and, if so, has this seen unequivocal humanisation of the doctrines, with an ever greater level of duty arising?;
- What has remained *constant* within this judicial discourse? In considering the operation of the appellate judiciary as an interpretive community we need to assess continuities as well as changes; and
- What does this likely mix of change and constancy tell us about the dynamic interplay of law with wider political, social and economic factors such as anxieties about parenting, safety and change within the built environment?

To achieve this 'dual-purpose' analysis this paper sets out a brief summary of the 'trajectory' of this area of law over the last 100 years before then drilling into the specific milestones of this 'story' in greater detail. Following this, the analysis will consider the cross-cutting themes of 'constancy' and 'context' which endure or otherwise interact with this narrative.

Before commencing this journey, a few words on methodology are appropriate. The author is a lawyer by profession, but one with a strong sociological orientation. This paper's aim is not to discover what the law is now, or was in the past – but rather to use aspects of legal historical and doctrinal analysis in support of a wider consideration of how the evolution of the doctrine has occurred within the juridical field. For Bourdieu, the juridical field is a "social universe" (816) in which law, lawyers and lawyering achieve an enduring discourse in which individual actors both shape, and are shaped by their community of knowledge and practice. But, importantly, the juridical field is not solely inhabited by law and law principles (that is, the conventional realm of jurisprudence) but instead it is constantly interacting with and existing by means of social, economic, psychological and linguistic practices within and beyond the field.

This paper is based upon a study of appellate court reported judgments, legal journal articles, law reform reports and wider policy and stakeholder debates. Inevitably Court of Appeal and

House of Lords decisions play a significant part in this source material, and it should be plainly stated, and following Simpson (1996), that judicial cognition does not only take place within the appellate courts – the law is interpreted by lawyers advising their clients and by judges in unreported first instance decisions. In these other (harder to access) sources lie the day to day application and well trodden path following behaviours that make up the ‘day-to-day’ lived reality of the law (what Bourdieu would style law’s ‘*habitus*’). But this paper seeks to show that *habitus* can also be detected at work within the appellate courts.

2. The story in outline

Text-book accounts of the development of legal principles in English Law concerning occupiers’ liability towards child trespassers have it that the last 100 years have seen a “humanising” trend, starting from a “draconian”⁵ position “derived from the nineteenth century moral judgment that a trespasser was deserving of his plight” (Markesinis & Deakin 1999: 327), a doctrine rooted in an “an over-zealous preoccupation with the sanctity of real property rights, even over that of human life.” (OLRC, 1972: 6-7). The 1929 House of Lords judgement in *Addie*⁶ is vilified as symptomatic of this attitude – a decision finding no duty owed by a mine owner to a four year old child trespasser fatally crushed by a spoil conveyor on its site, because the common law was held in that case to offer no duty of care to a trespasser. Their Lordships had thus upheld a view that liability (if any) was dependent upon a “technical question of the capacity of the injured: invitee, licensee, trespasser”⁷.

In 1957, at the recommendation of the Law Review Committee, the principles of occupiers’ liability for lawful visitors was codified in the Occupiers Liability Act 1957, removing a differentiation of levels of duty as between various types of invitee and licensee and replaced these with declaration of a “common duty of care” to be owed by occupiers to all lawful visitors – but the common law was left uncoded as regards (any) duty to trespassers.

In 1972 the House of Lords, in *British Railways Board –v- Herrington*⁸ overruled *Addie* and, in a spirit of modernisation, and self-declared “humanisation” of the common law on this

⁵ Per Lord Diplock in (1972) *British Railways Board –v- Herrington* A.C. 877, HL at 931.

⁶ *Robert Addie and Sons (Collieries) Limited –v- Dumbreck* (1929) A.C. 358, HL.

⁷ Per Lord Hailsham in (1972) *British Railways Board –v- Herrington* A.C. 877, HL at 924.

⁸ A.C. 877, HL.

question, held that circumstances could arise in which an occupier would owe a “common duty of humanity” to a trespasser. However, due to concerns about how the five Law Lords had variously defined that possibility, the Law Commission was asked to examine the law and submitted its recommendation for codification of the principles of occupiers’ liability to trespassers in 1976. In 1984 a modified version of the Law Commission’s proposals was enacted, as the Occupiers’ Liability Act 1984 setting the current basis for occupier liability towards trespassers and other uninvited visitors (and whether adult or child).

The aim of this paper is not to fundamentally challenge the accuracy of this “story” – but rather to show the relevance of some of the subtleties, continuities and changes that lie beneath the surface.

3. The law at the start of the 20th century

Prior to *Donoghue –v- Stevenson* (1932)⁹ there was no notion of a singular, generic, duty of care applicable to all circumstances. Duties of care were identified on a case by case basis. Therefore, unsurprisingly, questions of whether a duty existed as regards various types of injury suffered by persons visiting premises were regarded as self-contained areas of jurisprudence. In the case of visitors to land, these incidents involved real property and the primary question appeared a property law one: whether an owner should have unfettered entitlement to enjoy his property without interference or duty to others. In this context contemplation of liability to trespassers was indeed viewed from a “property rights” perspective.

In 1902 the harshness of the doctrine – that a trespasser must take the land as they find it and cannot complain about it to the owner – was reaffirmed in *Devlin –v- Jeffray’s Trustee*¹⁰. In this case a six year old child fell an unfenced clay pit 25 yards from a public road and was drowned. The court held there was no liability for the occupier – as presence of a danger was not, in itself, sufficient – there had to be an invitation (or allurement) to come onto the land and suffer that fate.

The doctrine of ‘allurement’ had developed during the nineteenth century as an attempt to provide some protection for child trespassers, expressing a judicial view of what motivates

⁹ AC 562, HL.

¹⁰ (1902) 5 F. 130

children to trespass which we will consider further later in this paper. The allurements doctrine held that some features may beguile a child into straying onto someone's land, and that occupiers must therefore anticipate (in certain circumstances) that the nature of those features may give rise to some form of duty towards a child trespasser (but not to an adult – the allurements doctrine only applied to children).

The first 20 years of the twentieth century saw a plethora of appellate decisions in cases concerning injury or deaths to child trespassers. Many of these cases were Scottish in origin, possibly reflecting the prevalence of jury trials in that jurisdiction in such claims and/or reflecting a perceived need to find opportunity to develop a particularly dour approach to child trespass seen to be characteristic of Scots law at the time, encapsulated (in many aspects) by Hamilton LJ, who said in *Latham –v- R. Johnston & Nephew Ltd* (1913)¹¹:

“Children’s cases are always troublesome. English law has been very ready to find remedies for their injuries. Scotch law has been less indulgent...it is hard to see how infantile temptations can give rise to rights, however much they may excuse peccadilloes...”:

Addie was viewed by the judiciary as an opportunity to address a drift away from a clear statement of principle. The House of Lords’ decision in the Irish case, *Cooke –v- Midland Great Western Railway of Ireland* (1909)¹² had caused concern by appearing to impose a liability upon occupiers for children known to play upon (or in the vicinity of) industrial machinery (in this case a railway turntable). Liability had been imposed on the Railway company on the basis that it had granted an implied licence to the boy injured whilst playing on that equipment because it had failed to effectively exclude the boy and his friends from that location. Subsequent cases (like *Latham*) had sought to distinguish *Cooke* on its facts. The Law Lords needed to be strong in their ruling if the feared effect of *Cooke* was to be resisted. Accordingly, in *Addie* Lord Hailsham declared (at 365):

“Towards the trespasser the occupier has no duty to take reasonable care for his protection or even to protect him from concealed danger.”

¹¹ 1 KBD 398, CA at 413

¹² A.C. 229, HL.

Addie was not, however, followed in *Excelsior Wire Rope Co Ltd -v- Callan* (1930)¹³, in which the House of Lords found the operator of a haulage conveyor liable for fatal injury caused to a five year old child trespasser playing upon that equipment. Subsequent commentators have struggled to see the factual differences that their Lordships claimed justified departure from *Addie*. However, *Excelsior* did not overrule *Addie*, and it was *Addie* that came to be regarded as encapsulating the general rule. Accordingly cases in the inter-war years tended to deny liability – finding the cause of danger to be evident, not an allurement and thereby something that a trespasser should take “at his own risk”¹⁴.

4. The move towards ‘humanisation’

In 1952 the Law Reform Committee (a precursor to the Law Commission) was asked to examine occupiers’ liability law, and to propose codification if it considered that appropriate, which it did. However, as regards liability towards trespassers the Committee declared themselves happy with the existing common law rule, although, as Heuston (1955: 271) noted, a minority of the Committee appear to have favoured a relaxation of the doctrine re: child trespassers but did not hold out for this.

Whilst the resulting (for England & Wales) Occupiers’ Liability Act 1957 therefore applied codification only to lawful visitors, the codification undertaken in Scotland (the Occupiers’ Liability (Scotland) Act 1960) made no distinction between invited and uninvited visitors, instead imposing a duty towards all persons entering premises, that duty being “such care as in all the circumstances as is reasonable” (Section 2(1)). Whilst Section 1(2) appeared to have intended (unlike the 1957 Act) only to have codified the *standard* of care – rather than altered common law rules for the identification of who had the benefit of these duties, the 1960 Act was subsequently called in aid by judges pressing for reform of the liability situation regarding trespassers in England, as we will see below. It appears that the 1960 Act did not trouble the House of Lords’ traditionalist approach to a trespasser injury claim in *McGlone – v- British Railways Board* (1966)¹⁵ in which Lord Reid – despite the wide duty imposed by the 1960 Act the circumstances of the case showed that the occupier had fulfilled his

¹³ AC 404, HL.

¹⁴ Per Murphy LJ in *McLaughlin –v- Antrim Electricity Supply Co.* (1941) NI 23, CA; see also *Morley –v- Staffordshire County Council* (1939) 4 All ER 92, CA.

¹⁵ SC (HL) 1, 15

statutory duty to act reasonably “if he erects an obstacle which a boy must take some trouble to overcome before he can reach the dangerous apparatus.”

Elsewhere, attempts were made during the 1960s to revisit these basic principles, for example Denning MR’s attempts to restate the law around principles of foreseeability in *Videan –v- British Transport Commission* (1963)¹⁶: such that if the presence of the trespasser was foreseeable then a duty of care would be owed. Attempts had also been made (rather unsuccessfully) to develop a distinction between ‘active’ and ‘static’ liability – that a difference in duty or standard of care might lie in examining whether the occupier was ‘simply’ occupying land or ‘actively’ carrying out some operation upon it, an early example being *Davis –v- St Mary’s Demolition & Excavation Co* (1954)¹⁷.

But it was *Herrington* that finally saw *Addie* assailed. We will discuss in more detail below how the judiciary (in both the Court of Appeal¹⁸ and the House of Lords¹⁹) mobilised rhetoric and other juristic devices to achieve this effect). Suffice to state here that the appellate judges in *Herrington* declared the *Addie* principles to be inhumane and that occupiers’ liability needed to be brought into line with general negligence, to align it with Lord Atkin’s axiomatic formulation in *Donoghue –v- Stevenson* (1932)²⁰ that:

“You must take reasonable care to avoid acts or omissions which you can reasonably foresee would be likely to injure your neighbour.”

and that whilst trespassers should not be owed a full duty of care, a “duty of common humanity” should apply, commensurate with the needs of a civilised society.

5. The problems with humanisation

In April 1972 the Law Commission were asked to consider the occupiers’ liability to trespassers in the light of the House of Lord’s decision in *Herrington*, reflecting concern that

¹⁶ 2 QB 650, CA, 666.

¹⁷ 1 WLR 592, QBD per Ormerod J.

¹⁸ *Herrington –v- British Railways Board* (1971) 2 Q.B. 107, CA.

¹⁹ *British Railways Board –v- Herrington* (1972) A.C. 877, HL.

²⁰ AC 562, HL at 580

that judgement had failed to set down “a clear principle applicable to the generality of cases” (Law Commission 1976: 3). As the Law Commission noted, the five Law Lords had each formulated different tests by which the duty could arise towards trespassers. Whilst each broadly contemplated the need for some actual or constructive awareness of the trespasser and their proximity to danger *Herrington* did not give a clear tool by which the “special circumstances” might be spotted in other future cases. Particular concern was focused upon the operability of the “humanity” test, a test which appeared, at least in Lord Morris of Borth-y-Gest’s formulation to equate to taking “such steps as common sense or common humanity would dictate”²¹. The Law Commission noted that Lord Denning MR the Court of Appeal in *Pannett –v- McGuinness & Co* (1972)²² had truncated the Lordship’s ruminations with characteristic short shrift: “*The long and the short of it is that you have to take into account all the circumstances of the case and see then whether the occupier ought to have done more than he did*”. Yet in *Herrington*, four of the Law Lords had said (each in their own way) that the duty towards trespassers needed to be calibrated so as to achieve:

“a compromise between the demands of humanity and the necessity to avoid placing undue burdens on occupiers”.

In short, to be readily applicable (perhaps with a lay audience in mind) the Law Commission considered that a codification – a formula – was needed.

Thus the Law Commission considered that *Herrington* would leave uncertainty – because the variant rationalisations and the early signs of abbreviation and selective invocation failed to enunciate a clear guide to what appeared to have been intended as two separate questions, namely:

- 1) what are the special circumstances in which a duty towards trespassers should be said to arise - and specifically what level of knowledge and/or foreseeability of (i) danger and (ii) trespass is required?; and

²¹ *British Railways Board –v- Herrington* (1972) AC 877 at 899.

²² 2 QB 599 at 606-7

2) how should the degree of care (i.e. “common humanity”) be determined? – specifically how should it be distinguished from the “reasonable care” due to lawful visitors?

In 1984 the Occupiers’ Liability Act was enacted for England & Wales, implementing many (but not all) of the Law Commission’s recommendations on occupiers’ liability for trespassers. The Act declared that the test to determine when a duty arose comprised (Section 1(3)) a three point test, namely that the occupier:

- (1) is aware of the danger or has reasonable grounds to believe that it exists;
- (2) is aware or has reasonable grounds to believe that the trespasser is or may come into the vicinity of the danger; and
- (3) the risk is one against which, in all the circumstances of the case, the occupier may reasonably be expected to offer the trespasser some protection.

And the standard of care was provided for as (Section 1(4)) a duty to:

“take such care as is reasonable in all the circumstances of the case to see that [the trespasser] does not suffer injury on the premises by reason of the danger concerned.”

However, “the danger concerned” is only that defined by Section 1(1)(b), namely:

“danger due to the state of the premises or to things done or omitted to be done on them.”

and it is this element of the codification that has proved to be the main avenue by which a distinct cooling in the judicial mood regarding recovery for injury suffered by trespassers has emerged over the last decade, reaching its zenith in the powerful House of Lords decision *Tomlinson –v- Congleton Borough Council*²³. *Tomlinson* shows the ascendancy of a powerful judicial discourse in which aspects of risk are seen as good (or at least unavoidable), the virtues of ‘personal responsibility’ are extolled and the view expressed that limits must be set by judges to the (perceived) growth of a compensation culture. Furthermore, that resources are finite, and that just because evidence shows that more could have been spent and/or done to increase trespasser’s safety does not mean that there was thus a liability to take those steps.

²³ *Tomlinson –v- Congleton Borough Council* (2002) EWCA Civ 309, HL.

We will consider aspects of this discourse further below, here we will look at how the ambit set by Section 1(1)(b) has been used in order to defeat recent claims.

In *Donoghue –v- Folkestone Properties Ltd* (2003)²⁴, a case involving severe injuries suffered by an adult diving into a dock, the claim was rejected by the Court of Appeal on the basis that the “danger” on the premises arose from the claimant’s actions and misjudgement rather than the state or condition of the premises. Lord Hoffman then adopted this reasoning in *Tomlinson* (another unauthorised diving case).

This line of argument also recurred in *Keown –v- Coventry Healthcare NHS Trust* (2006)²⁵. In this case an 11 year old child suffered brain injury when he fell 30 ft whilst climbing the exterior of a fire escape. It was held that the occupier had no liability for nothing wrong with the fire escape – it was the claimant’s deviant use of it that created the danger – and that danger was foreseeable to an 11 year old.

The same effect was achieved in *Baldaccino –v- West Wittering Estate PLC* (2008)²⁶. Here use of navigation beacon by boys as a diving platform was unauthorized – thus the boys were trespassers. The claimant, a 14 year old boy, had been paralysed by diving from beacon into shallow water after tide had receded. The court held there was nothing inherently dangerous in the beacon or the sandy floor or the seawater, thus nothing dangerous in the premises and even if there had been, a warning notice would have achieved nothing – the danger was obvious.

Likewise in *Simonds –v- Isle of Wight* (2003)²⁷ a 5 year old was injured jumping off a swing whilst running back across sports field from his parent to rejoin a school group. The High Court held that the trial judge had erred in finding there was a duty to fence off swings to prevent their unauthorized use during a school sports day: playing fields could not be made hazard free, and the danger of swings was self evident. Indeed, the hazard was the child’s use of the swing, not the physical existence of the swing.

²⁴ QB 1008, CA.

²⁵ 1 WLR 953, CA.

²⁶ EWHC 3386 (Qb).

²⁷ EWHC 2303 (Qb)

Consistent with Section 1(4) the standard of care is limited to that which is reasonable in the circumstances. Ascertaining what that reasonable standard requires is not an easy task – but recent cases show that statutory requirements and the practical limitations of eliminating risk entirely will come into play, and there is no duty upon occupiers to point out the obvious thus:

In *Mann –v- Northern Electric Distribution Ltd* (2010)²⁸ Wilson LJ gives leading judgment and forensically reviewed the factual evidence about how the claimant, who aged 15, had managed to climb into the substation and reach a place at which he was then electrocuted. The claim was based around breach of statutory duty: a duty to fence substations under the Electricity Supply Regulations 1988. The Court of Appeal found no liability by drawing upon the dictum of Asquith LJ in *Edwards –v- National Coal Board* (1949)²⁹ that the Regulations’ requirement of “reasonable practicability” in its degree of fencing provision involved a balancing of risk against costs and other sacrifice, declared that the claimant’s prowess in overcoming the defences that had been provided was not reasonably foreseeable and noting that “no amount of security measures will keep out a sufficiently determined trespasser” (para 18).

These recent cases also promulgate a notion of the ‘blameless accident’, as in *Bourne Leisure Ltd t/a British Holidays –v- Marsden* (2009)³⁰. In this case a 2 year old child wandered away from parents at holiday park and was later found drowned in an on-site lake. The parents claimed against the holiday park operator (under 1957 Act) that it had failed to adequately warn the parents and or prevent a drowning in the lake. Moses LJ (who gave the leading judgment) displayed the modern approach by looking for any guidance identifying the park operators as being at fault (and finding none), noted the obvious nature of the hazard posed by the on-site lakes and likelihood that further warning notices or briefings would have been unlikely to have had any effect.

²⁸ EWCA Civ 141, CA.

²⁹ 1 K.B. 704, CA at 712.

³⁰ EWCA Civ 671, CA.

6. Context: interpreting the changes

What can this succession of cases tell us about the influence of physical changes in the built environment and changes in social attitudes upon the evolution of this area of law over the last 100 years?

6.1 Changes in the built environment

As noted by Lord Wilberforce in *Herrington*³¹ the law in this area tends to develop around issues – often physical things - that emerge and need addressing. Viewed in this light, and for an example, a distinct jurisprudence on railway-line fencing can be identified. The railways industrialised long narrow tracts of the countryside from the mid-nineteenth century onwards, later the hazards (electrical and mechanical) of tram systems had to be adjudicated on and so forth. The occupiers' liability cases of the industrial era (of which *Addie* is a classic example) were wrestling with the spatial consequences of industrialisation. New machines now sat in fields, exposed to the curious. Standard patterns of housing development in mining and other activity-specific location were huddled for obvious convenience adjacent to the mills and railways to which they provided worker accommodation (as was the case in *Addie*). The child trespassers lived locally and played locally.

However from the 1920s onwards, as planning legislation started to promote zoning of development industry started to be separated from residential settlements. The new factories were increasingly located in homogenous 'industrial estates' (Darley 2003; Stratton & Trinder 2000: 5). This did not remove the problems of child trespass to industrial premises – but steadily it reduced its relative importance, as reflected in the nature of the premises trespassed upon in the cases surveyed above. In the post 1945 cases we also see the rise of bomb sites³², urban clearance³³ and landfill sites³⁴. In the early twenty first century the child trespasser appears to mainly haunt hospitals, schools, beaches and lakes rather than factories

³¹ *British Railways Board –v- Herrington* (1972) A.C. 877 at 930

³² *Davis –v- St. Mary's Demolition & Excavation Co* (1954) 1 W.L.R. 592, QBD

³³ *Harris –v- Birkenhead Corporation* (1975) 1 W.L.R. 1345, CA.

³⁴ *Penny –v- Northampton Borough Council* (1974) 20 July, The Times, CA..

and mines. Trespass to transport and power in infrastructure can, however, be regarded as a constant across the whole study period though.

The point here is that child trespass is now less about machinery and locality, and play patterns have changed. The places where (judges at least) would expect to find children has changed since Lord Atkinson's assertion that:

“...every person must be taken to know...that public streets, roads and public places may not unlikely be frequented by children of tender years and boys of this character” in *Cooke* (1909)³⁵.

Clearly it is easy to make sweeping generalisations about such socio-economic changes: but the judges have appeared happy to do so. Take for example Lord Pears on's call to arms in *Herrington* invoking a particular image of change in the built environment to justify the repositioning of the law thus:

“With the increase of the population and the larger proportion living in cities and towns and the extensive substitution of blocks of flats for rows of houses with gardens or back yards and quiet streets, there is less playing space for children and so a greater temptation to trespass. Also with the progress of technology there are more and greater dangers for them to encounter by reason of the increased use of, for instance, electricity, gas, fast-moving vehicle, heavy machinery and poisonous chemicals. There is considerably more need than there used to be for occupiers to take reasonable steps with a view to deterring persons, especially children from trespassing in places that are dangerous for them.”³⁶

Lord Pears on's anxieties appear characteristic of the ‘*population bomb*’ (Eurlich 1968) concerns that emerged in the 1960s, a similar anxiety apparently being abroad in relation to the crowding out of the countryside, here the Duke of Edinburgh speaking at the National Playing Fields Association's annual conference in 1963:

“We are on the threshold of the age of leisure, and the problem is no longer academic. There are stark and immediate problems facing us. The queues for playing fields are

³⁵ A.C. 229, HL at 237

³⁶ Lord Pearson, *British Railways Board –v- Herrington* (1972) A.C. 877, HL at 927.

getting longer; the pressure on general sports and recreation clubs of all sorts is increasing. Swimming pools and sailing clubs are getting crowded, and recreational users of water are beginning to run into each other.” (quoted in Dower 2001: 35)

From the 1960s onwards the focus of lawmaking in creation and control of the built environment also changed its emphasis. The Welfare State and the consumer society (we somehow never quite reached the “age of leisure”) brought with them new premises, services and activities – each to be embodied within new structures, and this was reflected in the 1963 Offices, Shops and Railways Premises Act and its imposition of safety standards upon these places, extending the remit of health & safety legislation beyond the realm of the Factories Acts. Then the framework legislation created by the Robens Committee³⁷ and embodied in the Health & Safety at Work etc Act 1974, imposed both a generic statutory duty of safety managers of all business premises to all visitors (Section 4) but also an array of generic and activity specific secondary legislation requiring and controlling the assessment and management of place safety.

6.2 Changes in the role of the judiciary

The 1960s also saw a change in the way that the judiciary came to see their role. In 1966 the House of Lords had declared that it could, if changed circumstances warranted it, overrule its prior decisions – and the *Herrington* decision was an early example of the use of that power. The role of the judiciary as active law makers (rather than just law-diviners) was becoming increasingly accepted – both by academic commentators on the *Politics of the Judiciary* (Griffiths 1977) but also the senior judges themselves, such as Lord Reid (1972):

“Those with a taste for fairy tales seem to have thought that in some Aladdin’s cave there is hidden the Commission on Law in all its splendour and that on a judge’s appointment there descends on him knowledge of the magic words ‘open sesame’. Bad decisions are given when the judge has muddled the password and the wrong door opens. But we do not believe in fairy tales anymore”.

³⁷ *Safety and Health at Work: Report of the Committee 1970-72*, London: HMSO Cmnd 5034 (Chairman: Lord Robens).

In the industrial unrest of the 1970s, the energy crises, political turmoil, and economic uncertainty the judiciary started to become more attuned to the public policy implications of their appellate findings of law. We will consider below how those concerns were reflected in the judicial rhetoric underlying the ‘leading cases’ summarised above.

6.3 Changing attitudes to parental responsibility

The case reports give us insight into how the judges have regarded parental responsibility at various instances across the last 100 years. These insights are available in passing comments, no doubt intended as statements of what seemed obvious at the time. Tracking the evolution of these statements appears to show the retreat from the widespread practice of leaving children (even those of what would now be pre-school age) to wander and play alone in their locality.

For example in *Cummings –v- Darngavil Coal Co Ltd* (1903)³⁸ Lord Young tells us that parents think the risk low to let their children wander unaccompanied, whilst in the Court of Appeal in *Addie* the Lord President (at 556) tells us that letting four year old child roam alone in poor neighbourhoods is not contributory negligence on their or their parents part, for it is normal in such areas (but in the same judgement, Lord Blackburn’s dissenting judgement asserts that the child’s parent should have supervised his child better).

The apparent belief in the prevalence of the ‘wandering poor child’ appears to have persisted into the early 1950s, for Lewis wrote in 1954 that:

“the law tacitly takes account of the fact that among the poorer classes even very young children are usually under the charge of nobody of responsible years when they go abroad, and that this, taken by itself, will not deprive them of a remedy” (Lewis: 716)

The judicial mood does appear to have changed soon after though, perhaps in reflection of widespread post-war anxieties about ‘maternal deprivation’ and so-called ‘latch-key children’ caused by the parental distraction of the war (Kynaston 2008: 113) and a wave of juvenile

³⁸ 5 F. 513.

delinquency attributed to this in the early 1950's. By the time of *Bates –v- Stone Parish Council* (1954)³⁹ Birkett LJ could write:

“I am bound to express my own deep regret that the infant plaintiff should have been allowed to go to the recreation ground at all, accompanied only by another small boy of some six years of age. A boy of six could not be expected to be a proper or a competent guardian, and to a child of three and a half most things in this world take on an aspect of danger if he is not protected against them. To reach the recreation ground from the plaintiffs' home required the two little boys to cross two public highways with all the dangers of modern traffic, and that fact by itself, I should have thought, would have been enough for a careful and thoughtful parent to forbid a small boy to go unless he were in watchful and careful hands. Yet the fact remains that the infant plaintiff was allowed to go in the care of a boy of six, and it was the alleged negligence of the defendants that was alone in question.”

Devlin J in *Phipps –v- Rochester Corp* (1955)⁴⁰ – in which a boy of five, out blackberry picking with his sister aged seven, fell into a sewer pipeline trench on a site awaiting housing development - conducts a thorough review of the ‘parental responsibility’ cases. In his judgement he moves a way from the fiction of liability not arising because of an implicit condition requiring the child only to be present if supervised and instead shifts to a foreseeability basis and tries to account for the ‘older’ view of the wandering child, by suggesting that in determining whether he has a duty the occupier “may have to take into account the social habits of the neighbourhood”⁴¹. However, he expresses general conclusion in favour of parental responsibility being greater than that falling upon occupiers as regards child trespassers:

“But the responsibility for the safety of little children must rest primarily upon the parents [either to ensure their children don’t wander, or to check out the suitability of where they might be to wander] It would not be socially desirable if parents were, as a

³⁹ 1 W.L.R. 1249, CA at 43.

⁴⁰ QBD 450, QBD.

⁴¹ QBD 450, QBD, at 472.

matter of course, able to shift the burden of looking after their children from their own shoulders to those who happen to have accessible land”⁴²

Devlin’s rationalization of why it wasn’t foreseeable that a child of five would be sent out to play alone at the site, shows us how in considering liability (at this time on the basis that an occupier would owe a duty to a “child of tender years” if there was an allurement or an hidden danger there) each of an occupier, a parent and ultimately a judge must “read” the locality – and make judgements about whether it was overcrowded, the site was the only local green space, whether the child’s house had a garden in which it could play and what perils existed in the locality.

In 2009, in *Bourne Leisure*⁴³, Moses LJ drew on Devlin J’s analysis in *Phipps* – reasserting the view duty to public is discharged if dangers are obvious to an adult – because little children can be assumed to be supervised. Moses (who gave the leading judgment) sought to move the debate away from a binary blame game of parent versus occupier, and regretted (at paragraph 16) that child accident cases sometimes become “bedevilled with the quest for attaching blame to either the parent or the occupier”. In his opinion the ‘innocence’ of the occupier is not, dependent upon proving the ‘fault’ of the parents. Stanley Burnton LJ concurred, with the contemporary juridical view that “Accidents may and do happen to young children without anyone being at fault.”(paragraph 25).

7. Constancies

7.1 Why constancies matter

Bourdieu’s work advocates the study of the subtle structural generative processes that underlie fields of practice like the law, scientific and religious interpretive communities. Therefore understanding how their meaning making practices – their discourse – adapt over time requires a study of constancy as much as a description of change. Here Braudel’s (2001: xiv) dictum that historians should look beyond the “surface disturbances” of events can be applied to our doctrinal analysis – as a need to enquire into what was remaining constant beneath these doctrinal changes.

⁴² QBD 450, QBD, at 472.

⁴³ *Bourne Leisure Ltd t/a British Holidays –v- Marsden* (2009) EWCA Civ 671, CA

7.2 The balancing act

As Lord Reid noted in *Herrington*, managing child trespassers is a thorny issue, particularly if presented as a straight choice between parents and occupiers. The law is required to achieve a balancing act – and:

“Legal principles cannot solve the problem. How far occupiers are required by law to take steps to safeguard such children must be a matter of public policy”⁴⁴.

As Lord Reid noted, alongside his increasingly pessimistic view on the prospects of parents being able to control their children⁴⁵, putting responsibility for protecting them upon premises occupiers – would “in most cases impose on occupiers an impossible financial burden”⁴⁶. This anxiety about costs or other impositions upon occupiers has haunted the cases throughout the study period. That anxiety can be found in concerns about the potential ramifications of the *Cooke* judgement in 1909⁴⁷, and in *Edwards-v-Railway Executive* (1952)⁴⁸ where Lord Goddard feared:

“had they to provide watchmen to guard every place on the railways of the Southern Region where children may and do get on to embankments and lines, railway fares would be a great deal higher than they are already.”

In response advocates of reform have over the years pointed back to those fearing such consequences – pointing out that the feared consequences have never materialised. Here Edmund Davies LJ in the Court of Appeal ruling in *Herrington*⁴⁹:

⁴⁴ *British Railways Board –v- Herrington* (1972) A.C. 877, HL at 897.

⁴⁵ At 897 Lord Reid had expressed the view that supervision of children was only “practicable at one time for a limited number of well-to-do parents but that number is now small.”

⁴⁶ *British Railways Board –v- Herrington* (1972) A.C. 877, HL at 897.

⁴⁷ E.g. Hamilton LJ in *Latham* (1913) CoA at 421; Bankes LJ in *Hardy –v- Central London Railway Co* (1920) 3 KB 459 at 467.

⁴⁸ AC 737 HL at 747.

⁴⁹ *Herrington –v- British Railways Board* (1971) 2 Q.B. 107, CA at 127.

“the reports are scattered with expressions of fear as to the probable outcome of imposing *any* sort of duty of care in relation to trespassers – that to do so would unjustifiably restrict occupiers’ enjoyment of their property; that the parents of child trespassers would in consequence be encouraged to neglect performing the duty of control which is properly theirs; that even burglars would be invested with a cause of action if they came to harm in their criminal invasion of another’s property: in short, that the occupation of property would be rendered intolerably burdensome. The alarmist has indeed been busy in the law. But he has not deterred the Scots from devising a rule which meets the situation admirably. Ten years have passed since their Act of Parliament reached the statute book, yet there has been no plague of writs on behalf of injured trespassers, nor has there been any reason for suspecting that the activities of railway or other authorities have been gravely disrupted in Scotland”

But a concern to avoid wider consequences or the diversion of resources remains a key motivating factor, for example Longmore LJ in *Keown*⁵⁰ (2006) considers the policy implications (expressed as a “tentative (obiter) view”):

“...it would not be reasonable to expect an NHS Trust to offer protection from such a risk. If it had to offer protection from the risk of falling from a normal fire escape, it would presumably have to offer the same protection from falling down drain pipes, balconies, roofs...windows and even trees in the grounds. This seems to me to be going too far.”

In defence of his view Longmore LJ gives two justifications. First, that NHS resources are better used in treating patients, and secondly, that there will be a fortification effect (which Longmore LJ notes had indeed happened already at the hospital in question) and local people will lose the amenity benefit of short cuts and harmless play in the grounds:

“It is not reasonable to expect that this should happen to avoid the occasional injury, however sad when injury occurs.”⁵¹

⁵⁰ *Keown –v- Coventry Healthcare NHS Trust* (2006) 1 W.L.R. 953, CA at 959.

⁵¹ *Keown –v- Coventry Healthcare NHS Trust* (2006) 1 W.L.R. 953, CA at 959.

Lewison J (in support of Longmore LJ) picks up this point: “If the trial judge was right, then occupiers of buildings up and down the country will have to ‘child proof’ their buildings in case children try to climb them.”⁵²

A specific concern about the effect upon access has also underlain these cases across the study period, thus this concern was raised by Lord Collins in *Cooke* (1909)⁵³: that liability risk may deter landowners from admitting the public “to enjoy the amenities of their private domains” and Farwell LJ in *Latham* (1912)⁵⁴:

“It is impossible to hold the defendants liable unless we are prepared to say that they are bound to employ a groundkeeper to look after the safety of their licensees, and the result of such a finding would be disastrous, for it would drive all landowners to discontinue the kindly treatment so largely extended to children and others all over the country.”

Farwell LJ’s concern was reiterated in *Bates –v- Stone Parish Council* (1954)⁵⁵ (a case involving a playground injury) where Romer LJ adding that provoking the closure of playgrounds would propel children back in to playing in the streets, places now considerably more perilous than playground chutes and swings.

Gross J in *Simonds* (2003)⁵⁶ – also a playground safety case - provides a contemporary echo of this specific concern, stating that there is:

“a danger with decisions such as this. The upshot would not be that swings are fenced off, it is far more likely that sports days and other simple pleasurable sporting events would not be held if word got round that a school could be liable in a case such as this. Such events would become uninsurable or only insurable at prohibitive cost.”

⁵² *Keown –v- Coventry Healthcare NHS Trust* (2006) 1 W.L.R. 953, CA at 960.

⁵³ *Cooke –v- Midland Great Western Railway of Ireland* (1909) A.C. 229, HL at 241.

⁵⁴ *Latham –v- R. Johnston & Nephew* (1913) 1 K.B. 398 CA at 408.

⁵⁵ 1 W.L.R. 1249, CA at 49.

⁵⁶ EWHC 2303 (QB) at Paragraph 13.

The language of risk assessment is also not an entirely recent invention in this jurisprudence for, in addition to the implicit cost/benefit reasoning of the access and resource implication concerns of the early twentieth century cases mentioned above, reference to risk calculation of sorts is present in Lord Young's description of parental reasoning about allowing their child to wander: *Cummings –v- Darngavil Coal Co Ltd* (1903)⁵⁷ 5 F. 513:

“We know quite well that children get off public roads, and that their parents do not trouble themselves, and simply let the children go, knowing, no doubt, that the risk which the children incur is so slight as to make it hardly worthwhile to interfere with the children.”

However, it is in the recent cases (influenced no-doubt by risk assessment concepts derived from occupational health and safety, the economic analysis of law and human rights discourse) we find, for example Lord Hoffman in *Tomlinson*, talking of: “The balance of risk, gravity of injury, cost and social value” (paras 34 – 37) and human rights (his heading “free will”: paras 44 – 50). We see also in Hoffman's judgement an advocacy of personal responsibility and the positive power of a measure of risk in daily life. This accords with a powerful lobby active within the senior judiciary, ‘captains of industry’, government and pressure groups seeking to advocate a pro-risk, pro-entrepreneur and anti-over regulation and anti-safety obsessed position, as detailed by Bennett & Crowe (2008) and expressed in a number of formative policy papers in recent years.

7.3 Images of child trespassers

One might expect the judiciary to have demonised child trespassers. But the depiction of the child trespasser is restrained, and at times quite benign, throughout the study period. On the whole the child trespasser is not presented as a sinister ‘other’, a threat to social order. This in part may derive from the fact that these reports concern situations in which such children have suffered injury, but still there is something notable in the judge's apparent keenness to differentiate child trespassers as “meritorious”⁵⁸ or “trivial and amiable.”⁵⁹ Care is often

⁵⁷ 5 F. 513.

⁵⁸ Per Diplock in *British Railways Board –v- Herrington* (1972) A.C. 877 HL at 936.

taken by the judges to emphasise that there are many types of trespasser and a child trespasser one of the least blameworthy, thus in *Herrington* Lord Morris:

“The term “trespasser” is a comprehensive word: it covers the wicked and the innocent: the burglar, the arrogant invader of another’s land, the walker blithely unaware that he is stepping where he has no right to walk, or the wandering child – all may be dubbed trespassers.”

This differentiation assists those in *Herrington* who wish to argue in favour of a duty of care towards child trespassers without invoking another recurrent fear in these cases: that to soften the line would lead to householders having a duty of care towards burglars⁶⁰.

The judicial characterisation of child trespassers is also worthy of study in terms of what it reveals of judges attitudes to the behaviour of children, and the source of the urge to trespass as a child. Without exception those judges who reveal a view on this matter subscribe to what might be called the ‘boys will be boys’ school of thought. Implicit within this rule of thumb theorising is a belief that something innate drives boys to trespass, it is not something that can be reasoned out of a child. It is a variant of the doctrine of allurement, that children will be drawn, automaton-like, towards certain structures. Whilst the language used to express this theory changes somewhat across the study period, the sentiment remains throughout.

The first traced outing of this view is found in Lord Denman’s judgment in *Lynch -v- Nurdin* (1841)⁶¹ where the child “merely indulged the natural instinct of a child amusing himself with the empty cart and deserted horse.” In *Latham* (1912) we see the view that “a child can get into mischief and hurt itself with anything if it is young enough”⁶² and in *Devlin* the view of Lord M’Laren:

⁵⁹ Per Sands in *Dumbreck -v- Robert Addie & Sons (Collieries) Ltd* (1928) S.C. 547, Court of Session at 557.

⁶⁰ For example, Salmon LJ in *Herrington -v- British Railways Board* (1971) 2 Q.B. 107 CA at 120.

⁶¹ 1 QB 29.

⁶² Hamilton LJ in *Latham -v- R. Johnston & Nephew Ltd* (1912) 1 KBD 398 at 413.

“...that every person must be taken to know that young children and boys are of a very inquisitive and frequently mischievous disposition, and are likely to meddle with whatever happens to come within their reach.”⁶³

Then there are cases in which fencing or other protective measures are pictured as having some magical and irresistible challenge and that:

“it is perhaps fortunate that there is no rule of law imposing on proprietors the impossible task of fencing their lands so as to keep out schoolboys, for the more difficult the fence was made the greater would be the temptation to overcome the obstacle.”⁶⁴

But these are not Edwardian vestiges - such views linger in to present day jurisprudence, for in the Court of Appeal hearing of *Herrington v Salford City Council* Lord Hale expressed the view that “an exceptionally agile and resourceful little boy may be able to get over or through almost any fence, if determined to do so”⁶⁵ and an equivalent view was expressed by Lord Hale in *Mann v British Railways Board* (2009): “No amount of security measures will keep out a sufficiently determined trespasser” (para 18).

The role of such constructions of the child trespasser and his motives appear to operate to desensitise – to take the (injured or deceased) trespasser out of the blame calculation – that is not to say that blame may not sometimes be attached (as it was, for example in *Young v Kent County Council* (2005)⁶⁶, but it removes the need for the judges to dwell on motive and the chosen aspect of the trespass.

7.4 Rhetorical techniques

We can now turn to consider the use of recurrent rhetorical devices within child trespasser cases. Once again, the aim is to show how constancy runs across the seemingly divergent cases (and thus reveal the operation of certain fixed points of reference within the judicial

⁶³ *Devlin v Jeffray's Trustees* (1902) 5 F. 130 at 134.

⁶⁴ Lord Atkinson in *Cooke* (1909) HL at 237.

⁶⁵ *Herrington v British Railways Board* (1971) 2 Q.B. 107, CA at 123.

⁶⁶ EWHC 1342 (Qb).

interpretive community). Reference here is to rhetoric in the neutral sense used by Aristotle – to identify forms and augmentations of argumentation – rather than in the perjorative sense of the term inherited via Plato (Postner 2009: 331-332).

A time honoured judicial rhetorical technique is to invoke ‘common sense’ when arguing for a departure from the line given by precedent. Twining & Miers have shown how Lord Atkin’s formulation of the duty of care in *Donoghue v Stevenson* (1932)⁶⁷ is based only upon two cases (with many in opposition) and rhetorical appeals to ‘common sense’, reasoning that is “permissible, inconclusive and not very cogent.” (1999: 365 note) yet Lord Atkin’s formulation (not least because of its simplicity and explicit moral linkage to Judeo-Christian precepts) has provided the foundational conceptual hub around which the tort of negligence has revolved ever since.

Equivalent appeals to common sense are invoked in *Herrington*⁶⁸ – for example Lord Wilberforce’s assertion that – “no greater than we ought to expect from the common law, which always leaves a residue to be completed by common sense” and Lord Morris’ assertion⁶⁹ that “any well-disposed but fair-minded member of the public” would think the defendant owed a duty towards the child in *Herrington*. He further supports his emotive line of argumentation by characterising the injured child as “a legal outcast” in the defendant’s characterisation of the applicable law⁷⁰.

Another rhetorical strategy at work in *Herrington* is what we might call ‘summoning a sense of alien-ness’ about the existing law. Thus in the Court of Appeal, Salmon LJ⁷¹ speaking of *Addie* – “a case which I confess I am unable to read, 40 years on, *without a sense of shock*.” (emphasis added). Salmon LJ then follows through with the following damning assertion that the doctrine of trespassers taking all risk of their entry:

“...may have been all very well when rights of property, particularly in land, were regarded as perhaps more sacrosanct than any other human right. This view was

⁶⁷ AC 562.

⁶⁸ *British Railways Board –v- Herrington* (1972) A.C. 877 HL at 919.

⁶⁹ *British Railways Board –v- Herrington* (1972) A.C. 877 HL at 901.

⁷⁰ *British Railways Board –v- Herrington* (1972) A.C. 877 HL at 902.

⁷¹ *Herrington –v- British Railways Board* (1971) 2 Q.B. 107, CA at 118.

widely held in the nineteenth century and, perhaps, even at the beginning of the present century, *influenced by the minds of those who were then no longer young*” (emphasis added)⁷²

Finally Salmon LJ⁷³ heaps further invective upon the status quo:

“It is difficult to see why today this doctrine should not be *buried* with the *artificial doctrine* based on *fictitious* permission springing from a *supposed* allurement which was *invented* by the courts in an *attempt* to mitigate the harshness and injustice caused by the *ancient legal dogma* relating to infant trespassers” (emphasis added)

But the modernisers in *Herrington* were also happy to draw support from long standing authority where that suited their purpose, thus Lord Morris in the House of Lords judgement⁷⁴ deftly resurrects the early nineteenth century spring-gun cases to show there has ‘always’ been a duty of common humanity underlying occupiers’ responsibilities towards trespassers.

In *Tomlinson*, Lord Hutton’s judgement in the House of Lords is keen to show parallels back to the early 20th century cases. To achieve this he has to perform opposite of the ‘alien-ness’ effect. Thus he refers to these old cases cherishingly, making light of their aged appearance, for even though:

“couched in old fashioned language...they still express a principle that is still valid today, namely that it is contrary to common sense, and therefore not sound law, to expect an occupier to provide protection against an obvious danger on his land arising from a natural feature...”⁷⁵

⁷² *Herrington –v- British Railways Board* (1971) 2 Q.B. 107, CA at 120. There is reasonable ground to speculate that Salmon LJ was actually no younger than the nineteenth century judges he was criticising: according to <http://thepeerage.com/p23173.htm> (accessed 1 July 2010) he was born in December 1903, making him 66 years of age when the Court of Appeal heard *Herrington* in 1970.

⁷³ *Herrington –v- British Railways Board* (1971) 2 Q.B. 107, CA at 120.

⁷⁴ *British Railways Board –v- Herrington* (1972) A.C. 877, HL at 904.

⁷⁵ *Tomlinson –v- Congleton Borough Council* (2002) 1 A.C. 46, HL at 88.

Furthermore, in stating⁷⁶ that “that there are certain risks which the law, in accordance with the dictates of common sense, does not give protection – such risks are ‘*just one of the results of the world as we find it*’⁷⁷”, Hutton specifically resurrects and incorporates Lord Dunedin’s fatalistic view, in *Hastie –v Edinburgh Magistrates* (1907)⁷⁸.

Personal biography is deployed by Lord Diplock in his *Herrington* judgement to add to the ‘alien-ness’ of *Addie*. Diplock claims in his judgement that *Addie*’s case was decided in the year he started reading for the bar and that “even at that time it offended against what Lord Atkin, only three years later, was to call “a general public sentiment of moral wrongdoing for which the offender must pay”⁷⁹, he then goes on to align himself personally with a that sense of disappointment:

“I well recall the disappointment with which it was received by those who thought that previous cases in this house had shown the common law as moving towards a less draconian treatment of those who trespassed innocently upon other people’s land”⁸⁰

But this reminiscence appears to be hollow rhetoric, for Diplock had prior opportunity to express such enduring disagreement with the *Addie* decision and, in particular, revealed none of this apparent dissatisfaction with the common law position in 1954 when he provided a minority report dissenting from the Law Reform Committee’s Third Report⁸¹ recommendation in favour of statutory consolidation of occupiers’ liability principles towards lawful visitors. In his minority report Diplock stated that he feared that codification would lack the subtlety and predictability of a case by case approach. Furthermore he said nothing to challenge the Committee’s conclusion that the common law position on occupiers’ liability towards trespassers did not need statutory reform (and at the time the only way in which *Addie* could have been circumvented).

⁷⁶ *Tomlinson –v- Congleton Borough Council* (2002) 1 A.C. 46, HL at 88. (emphasis added to show Lord Dunedin’s quote.

⁷⁷ Lord Dunedin *Hastie –v Edinburgh Magistrates* (1907) SC 1102 at 1106

⁷⁸ SC 1102.

⁷⁹ *Donoghue –v- Stevenson* (1932) A.C. 562 HL at 580.

⁸⁰ *British Railways Board –v- Herrington* (1972) A.C. 877, HL at 931.

⁸¹ 1953-54 [Cmd. 9305] *Law Reform Committee. Third Report (occupiers’ liability to invitees, licensees and trespassers)*.

In considering judicial discourse we need also to consider how metaphor and other figurative devices may give us insight into the operation of the interpretive community here. The extent to which meaning can be imputed to selection of figures of speech is perhaps open to debate – however the incongruous eruption of rural and/or classical imagery within judicial reasoning about child trespassers appear to be something other than random occurrence. In the study cases we find instances of elaborate deduction involving variously whether or not firing a gun in a field might hit a tram p⁸², imaginary children teasing a peaceful donkey or cow into kicking⁸³, “apple trees, streams, and other infantile joy”⁸⁴ and a hypothetical farmer afflicted by trespassers:

“...tearing up his prim roses and bluebells, or picking his mushroom s or stealing his turkeys, or f or the purpose of taking country walks in the course of which they will tread down his grass and leave gates open a nd watch their dogs chasing the farm er’s cattle and sheep.”⁸⁵

And, once again, use of these tropes follows thro ugh to the present day jurisprudence, for i n *Tomlinson* Lord Hoffm an appears m omentarily to fall into reverie when describing the Bereton Heath Country Park’s:

“paths now [running] through woods of s ilver birch and in summer bright yellow brimstone butterflies flutter in grassy meadows”⁸⁶.

Interestingly this r esort to th e poetic does not appear to be in tended as an aid to ruling in favour of the claim ant (i.e. by painting a picture of the si te of the accident as an allurement), for Hoffman then goes on to *deny* liability.

Hoffman then appears to be slip into an Ar cadian moment as a display of words mithery, for he then sum mons a Dionysian im age describing the persistence of m ore ad hoc recreational uses at this now restored quarry site, thus: “B ut the trad itions established in the previous

⁸² Cross LJ in *Herrington –v- British Railways Board* (1971) 2 Q.B. 107, at 138.

⁸³ Farwell LJ in *Latham –v- R.Johnston & Nephew Ltd* (1912) 1 KBD 398, CA at 407.

⁸⁴ Scrutton LJ in *Hardy –v- Central London Railway Co.* (1920) 3 KB 459, CA.

⁸⁵ Lord Pearson in *British Railways Board –v- Herrington* (1972) A.C. 877 HL at 925.

⁸⁶ *Tomlinson –v- Congleton Borough Council* (2002) 1 A.C. 46, HL at 72.

anarchic state of nature were hard to eradicate.”⁸⁷ He then completes the descent into a wanton display of classical learning by⁸⁸ attacking Ward LJ’s use of metaphor in the Court of Appeal’s adverse ruling, a metaphor likening the Mere’s water as “a Siren’s call strong enough to turn men’s minds”. Hoffman regards this as ‘gross hyperbole’ but can’t himself resist the temptation to show his (greater) familiarity with Greek mythology by seeking to demonstrate in some detail why the Siren analogy is unworkable. Hoffman’s metaphoric reverie eventually redeems his digressions by revealing his wider point – that the danger lies in how the claimant came to use the water rather than that the Mere was in itself in a dangerous condition (and thus a “danger due to state of the premises” in the meaning of the 1984 Act).

All that can perhaps be said of this digression and imagery is that it was display behaviour, aimed at showcasing Lord Hoffman’s erudition and classical education (and thus for Bourdieu 1987: 812) an example of the creation of “symbolic capital” – through the display of a traditional symbol of erudition.

7.5 The spiral towards rigid rationalism

Finally, we can consider what the trajectory of the law in this area over the study period reveals about the operation of the ‘juridical field’.

In 1955 Heuston described occupiers’ liability as:

“...a field of ‘lawyers’ law”, which although untouched by statute, has acquired a bad reputation. The broad and simple principles established by our great Victorian judges have been obscured by over-refinement and subtlety...when English law makes a mistake it is usually due to an excess rather than to a lack of rationality” (1955: 271).

Two points can be extracted here. First, that Heuston appears to be taking some apparent pride in the fact that that this area of the common law had avoided statutory intervention, for he carries on to describe the prospect of statutory codification as both drastic and “likely to

⁸⁷ *Tomlinson –v- Congleton Borough Council* (2002) 1 A.C. 46, HL at 77

⁸⁸ *Tomlinson –v- Congleton Borough Council* (2002) 1 A.C. 46, HL at 80

make many people cringe” (271). This suggests a purism amongst the judiciary in their decision taking (and law making) abilities – something echoed in the tone and content of Lord Diplock’s minority report appended to the Law Reform Committee’s 1953 report.

The second point is that Heuston is suggesting that the very processes that make the common law successful can, if left to run too long or too deep, make the law *too* rational. This may sound a strange assertion, but it finds echo in diverse places, for Law Commission’s Working Paper No. 52 (1973)⁸⁹ recommended statutory consolidation because the common law principles had via *Herrington* become “over-refined”, whilst the Law Commission’s rebuff to the judicial assertion that they are the best people to resolve the post *Herrington* doctrinal confusion was that statements of the law to give a “reasonably certain guide to the law *before* cases are actually decided”⁹⁰ (emphasis added). Here the Law Commission was focussing on the need for the law to be ‘available’, in a practical sense, for easy interpretation and application by those whose behaviour it may seek to influence – in short, occupiers.

Bourdieu is very relevant here, for he observes that the rationalising tendencies of academic and appellate level jurisprudence may at times become too rarefied, and that then the survival of the authority of the juridical field requires that judges (or other law makers) adapt law and keep it (in a sense) unperfected: In Bourdieu’s words, what has to happen is “adaptation to reality of a system which would risk closing itself into rigid rationalism if it were left to the theorists alone” (824). This adjustment is necessary because law *has* to remain workable at the pragmatic level – it needs to be seen as ‘useful’ to the “requirements and the *urgency* of practice” (824) (emphasis in the original) – otherwise the field may collapse (or at least be weakened). Thus such changes, such adaptation, along with the apparent signs of conflict between various groupings within the legal professions when such adjustment takes place, should not be seen as a sign of weakness or pathology within the law. Instead such processes ultimately preserve the power and practical efficacy of the law.

⁸⁹ “*Liability for damage or injury to trespassers and related questions of occupiers’ liability*” – at page 26.

⁹⁰ The Law Commission (LAW COM. No. 75) (Cmnd. 6428) (1976) *Report on liability for damage or injury to trespassers and related questions of occupiers’ liability*, page 6.

8. Conclusions

The aim of this paper has been to examine the operation of meaning making, change and constancy processes within the appellate judiciary in relation to the sensitive issue of the occupiers' liability towards child trespassers in the United Kingdom over the last 100 years.

A process of doctrinal change has been mapped out – but undercurrents of constancy have also been identified. Recurrent tropes and rhetorical devices have been illustrated embodied within the succession of 'leading cases'. An attempt has also been made to situate this understanding of change and constancy within the "shaping influence of the social, economic, psychological and linguistic practices which, while never being explicitly recorded or acknowledged, underlie the law's explicit functioning" (Terdiman in his introduction to Bourdieu 1987: 2).

Bourdieu's conceptualisation of the 'juridical field' is complex (because that 'thing' itself is complex). In applying Bourdieu's theories the author's aim has not been to unmask a n ideological conspiracy at the heart of the doctrinal evolution of the law in this area – but rather to *illustrate* Bourdieu's abstract theoretical ideas by showing how a mix of historical, spatial, psychological and linguistic interactions underlie what might, at conventional glance, appear a fairly clear cut linear jurisprudential progression between different and distinct doctrinal paradigms and cultural epochs. In lingering tropes and forms of argumentation, a sense of what is and is not legally *pertinent*, tendencies towards formalisation and codification and the ability of the field to moderate its' own tendencies where this becomes necessary, we glimpse the deep structures of this field. For, as Terdiman (Bourdieu 1987: 807) notes, it is these deep and enduring structures within the juridical field that determine *what* and *how* the law will decide in any specific instance.

But, that's not necessarily to suggest that the law, in this or any other area, thereby becomes any easier for lay users to interpret and apply.

References

- Bennett, L. & Crowe, L. (2008), *Landowners' Liability? Is perception of the risk of liability for visitor accidents a barrier to countryside access?*, Countryside Recreation Network: Sheffield. Available at: <http://shura.shu.ac.uk/678/> (last accessed 6 July 2009).
- Bennett, L. (2009), "Why, What, And How? a case study on law, risk, and decision making as necessary themes in built environment teaching" *Journal of Legal Affairs and Dispute Resolution in Engineering and Construction*, Vol. 1, No. 2, May 2009, 105-113 . Available at: <http://shura.shu.ac.uk/694/> (last accessed 2 February 2010).
- Bennett, L. & Gibbons, C. (2010), "Perceptions of occupiers' liability risk by estate managers: a case study of memorial safety in English cemeteries", *The International Journal of Law in the Built Environment*, Vol.2, No.1. Available at: <http://shura.shu.ac.uk/1737/> (last accessed 2 July 2010).
- Bourdieu, P. (1987) "The force of law: toward a sociology of the juridical field" 38 *Hastings Law Journal* pp. 805 – 853 (translated and introduced by Richard Terdiman)
- Braudel, F. (2001) *The Mediterranean in the ancient world*, Allen Lane, London (translated by Sian Reynolds).
- Darley, G. (2003) *Factory Reaktion*, London.
- Dower, M. (2001) "New pressures on the landscape: recreation" in Smout, T.C. (ed.) *Nature, landscape and people since the second world war*, Tuckwell Press, Edinburgh
- Eurlich, P. R. (1968) *The Population Bomb*, Ballantine, New York.
- Fish, S. (1980), *Is there a text in this class? The Authority of interpretive communities*, Harvard University Press, London.
- Fish, S. (1989) *Doing What Comes Naturally: Change, Rhetoric, and the Practice of Theory in Literary and Legal Studies*. Clarendon, Oxford.
- Griffiths, J.A.G. (1977), *The politics of the judiciary*, Fontana, London.
- Heuston, R.F.V. (1955) "Occupier's Liability" in 18 *Modern Law Review* 3, pp 271 – 274.
- Kynaston, D. (2008) *Austerity Britain 1945-48: A world to build*, Bloomsbury, London.

Law Commission (1976) *Report on Liability for Damage or Injury to Trespassers*, Cmnd 6428.

Law Review Committee (1954) *Law Reform Committee. Third Report (occupiers' liability to invitees, licensees and trespassers)*. 1953-54, Cmd. 9305.

Lewis, T.E. (1954), *Winfield on Tort* (6th ed) Sweet & Maxwell, London

Markesinis, B.S. and Deakin, S.F. (1999) *Tort Law* (4th ed.) Oxford University Press, Oxford.

Ontario Law Reform Commission (OLRC), (1972), *Report on Occupiers' Liability*, Toronto.

Posner, R.A. (2009) *Law & Literature* (3rd ed.) Harvard University Press, London.

Reid, Lord (1972) "The Judge as law maker" (1972) 12 *JSPTL* 22.

Simpson, A.W .B. (1996), *Leading cases in the common law*, Oxford University Press, Oxford.

Stratton, M. and Trinder, B. (1999) *Twentieth century industrial archaeology*, E & FN Spon, London.

Twining, W. and Miers, D. (1999) *How to do things with rules*, Butterworths, London.

Maps, Land and Land Title – are they synonymous?

Carl Calvert
University College London (Hon Visiting Lecturer)
carl.calvert@wolfson.oxon.org

Pierre Clergeot
Nicolas Chauvin
l'Ecole Supérieure des Géomètres et Topographes, CNAM
pierre.clergeot@wanadoo.fr

Nicolas Chauvin
l'Ecole Supérieure des Géomètres et Topographes, CNAM
nicolas.chauvin@esgt.cnam.fr

1. Preamble

Certainty of land tenure is a prerequisite of land economy. Therefore the descriptions of the extent of the land need to be unambiguously defined to obtain sufficient certainty of title. The paper examines the similarities and differences in both requirements and actualities of land title mapping for land registration in England & Wales and France.

2. The INSPIRE Directive

Section 3 of the preamble to the DIRECTIVE 2007/2/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 14 March 2007 establishing an Infrastructure for Spatial Information in the European Community (INSPIRE) reads as:-

(3) The problems regarding the availability, quality, organisation, accessibility and sharing of spatial information are common to a large number of policy and information themes and are experienced across the various levels of public authority. Solving these problems requires measures that address exchange, sharing, access and use of interoperable spatial data and spatial data services across the various levels of public authority and across different sectors. An infrastructure for spatial information in the Community should therefore be established.'

At s2 of Article 3 defines spatial data:-

'spatial data' means any data with a direct or indirect reference to a specific location or geographical area;'

So it seems that appropriate definition and description of land is important to a multitude of public authorities within the EU, and certainly definition of land parcels must be one of those considered. If we look back at historical records of land parcels, their ownership and other attributes it will become clearer as to why such a Directive has become necessary.

Of course the INSPIRE Directive applies to public information. This begs the question; ‘Are Land Titles public or private information?’ In France the concept of a *titre foncier* (land title) does not exist and the nearest document to it is the *acte notaire*, but this is a private, not public document. Therefore the affect of INSPIRE may be more limited in France as regards definition of land on a map.

3. Appropriate definition of the land

A good definition of land is usually with the use of a topographical map. In The British Isles there are the Ordnance Surveys (OSGB, OSNI and OSI) whilst in France there is the *Institut Géographique National* (IGN) who produce the *carte IGN* at 1:25000 scale, and previously the *carte d’état major* (1:80000). The cadastre is at a larger scale but its objective is for taxation purposes and not to define legal extents of ownership.

The OS provide mapping at 1:25000 (the same as the *carte IGN*) but importantly the OS provide mapping at 1:1250 in urban areas, 1:2500 in rural areas and 1:10000 in mountain and moorland areas. The Land Registry in England and Wales (LR) use these large scale OS topographical maps as a basis for the LR Title Plans.

The IGN map is not used for defining extent of legal ownership for three reasons: a) it is insufficiently accurate; b) has no legal value, just like the cadastral plan in that respect; and c) it is a topographical map.

In the UK a chartered surveyor may be instructed to prepare a deed plan at a large enough scale to bind with the transfer deed; in which case the words and picture (map!) will enable both LR to register the land unambiguously as to extent and subsequent proprietors will be in no doubt as to the physical holding and hopefully the extent of easements such as rights of way. In France the *Géometre-expert* can provide a plan giving details of the boundary and also the textual description (*process-verbal de bornage*). However, this process when a boundary point (*borne*) is determined on the ground must be done in the presence of the two adjacent owners.

The advantage of the French system is that there is no need for the owners to try and interpret either a map alone or words alone whereas in England and Wales a Determined Boundary plan is more prescriptive in accuracy and precision and requires only a signature and not the presence on the ground of the parties.

4. Land or Legal Estate?

In England and Wales land is usually split into two parts, the physical land and the rights over that land. As Blackstone said (Gray, 1993) ‘land’ is ‘a very extensive signification’. The problem that is highlighted in this paper is the question of the significance of the shape, size and location of the land itself. What was appropriate or suitable to circumstance at one period of time is inappropriate for

current or future times and purposes. As we shall see the ‘Fixed Boundary’ of the Land Registration Act 1925 was introduced to give more certainty than the ‘General Boundary’ provided. The ‘General Boundary’ is defined in r278 of the Land Registration Rules 1925 (LRR 1925) as:

“... the exact line of the boundary will be left undetermined – as, for instance, whether it includes a hedge, fence or ditch, or runs along the centre of a wall or fence, or its inner or outer face, or how far it runs within or beyond it; or whether or not the land registered includes the whole or any portion of an adjoining road or stream.”

Unfortunately the costs of ‘fixing the boundary’ was such that very few boundaries were fixed and the introduction of ‘Determined Boundaries’ in s60 of the Land Registration Rules 2003 with the interpretation of Land Registry Practice Guide 40 are in danger of falling into the same trap; that of cost.

In Horn v Phillips [2003] EWCA Civ 1877 Lord Justice Jacobs stated that:

‘12. This document [transfer deed] is intended to transfer title to land. It gives a precise measurement -- I pass over "or thereabouts", which everybody agreed did not make much difference. To my mind the clear intention is to indicate exactly where that line is supposed to be. So relevantly, this map is not ambiguous. True it is that other parts of the map may be not very satisfactory because the Ordinance (sic) Survey is not very satisfactory, but in its relevant respect one can see exactly where the beginning and the end of the line is, and if you go to the land you can see exactly where they are.’

A good example that the map good, bad or indifferent had little or no bearing on the matter: a dimension had been given and was all that was sufficient and necessary to fix the other end of the line. So it seems that maps which have errors or limited by the scale of survey are adequate provided more information is available to locate the boundary exactly. This is, I believe, the principle behind the general boundary rule. The determined boundary relies on both accurate and precise measurements and mapping.

5. The Entail Improvement Act 1770 (c.51) – an Enclosure Act

This act sought to enable the identification of land parcels with property rights and to that end the land was surveyed and maps were drawn.

The Preamble to the act starts:

“Whereas by the Entail Act 1685 (of Scotland)—all His Majesty’s subjects are empowered to taillie their lands and estates in Scotland with such provisions and conditions as they shall think fit, and with such irritant and resolute clauses as to them shall seem proper; and which taillies, when completed and published in the manner directed by the said Act, are declared to be real and effectual against

purchasers, creditors, and others whatsoever: Wherefore, to prevent a mischief and inconveniency so hurtful to the publick.”

In other words this is a statute which relates real property and real property right to a surveyed position and defined topographical extent. This map shows physical features and the names of occupier of the land.

Figure 1: An 18th Century Enclosure map of part of Lincolnshire, England



In France the *acte terrier* of the 18th century produced similar maps – for similar reasons; enclosures and not property boundaries.

6. Examples of ambiguous definitions of the land

Before looking at the details of any legislation for land tenure in any country it may be instructive to consider some recent legal cases in England and Wales where problems have arisen. It is the Ordnance Surveys' (OS) maps which the UK Land Registries use as the basis for Land Title Plans. In England & Wales the Land Registry (LR) examine the plan, which hopefully is attached to the document which conveys the land parcel from one party to the other, and try to show the parcel's extent on the OS map.

There are often ambiguities and difficulties in this process and the cases below are examples of some of these problems.

The first case is that of Willsher v Scott & Ors [2007] EWCA Civ 195 in which it was held that “The quality of the [transfer] plan is very poor, it is not to scale and the colours are indistinguishable from each other. It is based upon a less than perfect trace of an OS map (itself unreliable).” Laws LJ continued by saying that:

“In all [these] circumstances the Ordnance Survey maps offer, in my judgment, an uncertain guide as to the precise boundary line. It was submitted to Arden and Dyson LJ that while (as I have said) the Ordnance Survey maps do not purport to fix private boundaries, yet if parties to a conveyance choose to use such a map to mark the boundary they will be held to what the map shows: *Fisher v Winch* [1939] 1 KB 666 in this court.”

So now the proposition is that if parties to a contract decide that the transfer plan is the OS plan then what the OS plan shows is critical. However, it is well known – I hope – that OS do not show on their published paper plans the nature of the physical feature surveyed (except where an administrative boundary is linked to it with an annotation such as 0.9m Root of Hedge etc). I understand that metadata for the latest OS digital spatial information (Master Map) does have this information but unfortunately it is not available on the LR Title Plan, thus the nature of the boundary feature is, within a short space of time, unknowable from the map alone. Laws LJ then quoted Sir Wilfrid Greene MR:

"Of course, the fact that the boundary is shown in a particular place on an Ordnance (sic) map is in itself no evidence of what the true boundary is as between the parties, but where the party's title is derived from a document which refers to the Ordnance map, it is necessary to look at the Ordnance map and ascertain where the boundary shown on that map is truly positioned."

It is true that there was evidence in Winch, to which the Master of the Rolls refers (at 672), including testimony from a senior employee of the Ordnance Survey, that the relevant boundary in that case was the centre line of a hedge and fence. However to my mind it is of the first importance that the executive words of the 1986 conveyance in the present case describe the transferred land as "adjoining" the road. The deed plan has to be read conformably with this text. The highest Mr Douglas puts his case (skeleton argument paragraph 28(vi)) is that there was a "presumption that the boundary lay along the mid point of the said wall and/or hedge" (my emphasis).

So it seems that the construction of a transfer deed is a word and pictures job. In Sefton v Halliwell [2007] EWCA Civ 473 the Appeal Judge, Lady Justice Hallett, said of the parcels clause of the transfer from the appellant to the respondent.

“This clause may have identified the property conveyed with reference to the 1982 plan, but even the appellant was forced to concede that the 1982 plan was of poor quality. It was fuzzy and virtually illegible. It lacked precision. It was not to scale. In fact, to say it was not to scale is to understate the position. Any attempt by the expert at producing a scale produced huge variations. This is particularly important when attempts are made to fix the boundary by reference to the finger or tongue of the land.”

The point is quite clear here; a poorly drawn map is of little use and is a time bomb waiting to engulf the unwitting purchaser in the haze and murk of such legal wrangling as could have been avoided.

“The appellant adopted the expert's conclusion that "the Land Registry have used the end of the verge as the point to which to interpret the deed line and, due to the error in the position of the verge on the Ordnance Survey mapping, have drawn the red line closer to Woodfield House than as shown on the 1982 plan". That conclusion may or may not be a reasonable one, but it was not a conclusion for the expert to draw; it was a conclusion for the judge to draw if he thought appropriate.”

So here it seems the expert was on thin ice because he gave an opinion on how the LR would have done something without knowing the facts, which even if he knew them may not have been sufficient reason for him to have said what he said. Lady Justice Hallett then continued:-

“In any event, in my view, the alleged error on the filed plan paled into insignificance when compared with the paucity of information and the deficiencies in the 1982 plan.”

So, what is the conclusion? Firstly; that the conveyancers did not understand the need for an accurate, precise and contemporaneous plan to attach to the transfer deed: secondly that the OS plan, regardless of any positional inaccuracies due to limitation of scale or otherwise, did not convey sufficient information to LR to determine what the feature surveyed was. The solution would be to produce the Title Plan with such information available from metadata to determine what was surveyed, when it was surveyed and to what accuracy.

7. Boundaries under the Land Registration Act 2002 (LRA 2002)

In English law, boundaries are dealt with in section 60 of the Land Registration Act 2002 (LRA 2002)

s60 Boundaries

(1) The boundary of a registered estate as shown for the purposes of the register is a general boundary, unless shown as determined under this section.

(2) A general boundary does not determine the exact line of the boundary.

(3) Rules may make provision enabling or requiring the exact line of the boundary of a registered estate to be determined and may, in particular, make provision about—

(a) the circumstances in which the exact line of a boundary may or must be determined,

(b) how the exact line of a boundary may be determined,

(c) procedure in relation to applications for determination, and

(d) the recording of the fact of determination in the register or the index maintained under section 68.

(4) Rules under this section must provide for applications for determination to be made to the registrar.

Note that the accuracy and precision of ‘determined boundary’ is undefined and it is in the Land Registration Rules 2003 (LRA 2003) PART 10 to add more detail. In fact the only matter which relates to the survey specification is in section 118.

s118. - (1) A proprietor of a registered estate may apply to the registrar for the exact line of the boundary of that registered estate to be determined.

(2) An application under paragraph (1) must be made in Form DB and be accompanied by -

(a) a plan, or a plan and a verbal description, identifying the exact line of the boundary claimed and showing sufficient surrounding physical features to allow the general position of the boundary to be drawn on the Ordnance Survey map, and

(b) evidence to establish the exact line of the boundary.

It is LR themselves who specify the survey specifications in Practice Guide 40, which at section 3.3.2 states that the plan must, amongst other things:

- show sufficient surrounding physical features to allow the position of the boundary to be drawn on the Ordnance Survey map. If the plan does not allow the boundary to be identified on the Ordnance Survey map the application will be rejected
- be drawn accurately to a stated recognised scale. We advise that this is no smaller than 1/200
- Another bullet point
- Measurements shown on the plan need to be both precise and accurate to 10mm and should be taken from at least two defined points on surrounding permanent features. In this context 'permanent features' are taken to be physical features, which it is reasonable to assume will remain in position for at least 10 years, taking into account the nature and construction or character. The measurements must be taken horizontally, that is not along a slope.

The accuracy of any map is determined by the accuracy of physical definition of the ground itself as well as both survey and plotting. The RICS themselves publish Client Specification Guidelines for surveys of 1/500 and larger and there is no definitive accuracy given for detail.

8. What if the boundary shown on the LR Title Plan is 'subject to survey'

The case in question is Derbyshire County Council (DCC) v Fallon & Fallon [2007] EWCH 1326 (Ch) which is an appeal to the High Court from the Adjudicator to Her Majesty's Land Registry (the Adjudicator).

The facts were these; DCC owned unregistered land next to the Fallons and the Title Plan for the Fallons showed the boundary between DCC's land and the Fallons' land in the 'wrong place' so as to include land owned by DCC.

The issues were: 1) where was the boundary between the two plots? 2) was there an Adverse Possession issue? 3) if DCC won on both these issues then was the dispute a boundary dispute or a property dispute, in other words was it a General Boundaries issue? There were other issues but these are not discussed here. At the hearing before the Adjudicator, it was held that DCC had won on the three issues numbered 1 to 3 above save for 'a comparatively minor point on the Paper Title Issue' but the Adjudicator declined to amend the register. DCC appealed against the fact that the Adjudicator refused to amend the Title Plan.

Stepping back from the legal issues for a moment, the case seems to be that there are physical features on the ground which were mapped by OS and as such used by LR in their Title Plan. The boundary in this case is shown on the Title Plan as a dotted line with the rubric that such lines have been plotted from the deed plans and the title plan may be updated from later survey information so the question of map and ground disagreeing is not an issue. It is the claim that the map does not show the legal

boundary which is DCC's problem. On the face of it this does not seem a problem as the General Boundaries Rule which is quoted from ss60(2) of the Land Registration Act 2002 as

(2) A general boundary does not determine the exact line of the boundary.

The High Court held, following Lee v Barrey [1957] Ch 251, that the dotted line on the original Title Plan created when the Fallons bought the land could not be relied upon to indicate the precise extent of their land, 'That could therefore only be determined by looking at where the boundary lay by reference to the pre-registration documents (at paragraph 17).

This satisfies subsection 2 of section 60 of LRA 2002 quoted above in that the general boundary does not determine the exact line of the boundary.

However, all is not as it seems for the Adjudicator held that the boundary shown on the Title Plan was in the wrong place and was therefore a mistake which could be corrected and the High Court upheld this. The High Court then asked itself a series of questions; if the alteration was made would it affect the Fallons' title? Would it affect land in the possession of the Fallons? Would it be just to make the alteration? And were there **exceptional circumstances which justified not making the alteration** (emphasis added)? The Adjudicator had ruled that altering the Title Plan would not affect the paper title but would produce "another general boundary in a more accurate position than the current general boundary" (at paragraph 68).

There are two points which I take from this. The first is rather obvious in that a general boundary is a general boundary unless it is a determined boundary (subsection 3 of section 60 LRA 2002 above). The second point is that there has to be evidence to judge that there is a more accurate general boundary than was shown on the Title Plan. In this case the evidence was in the form of a survey by DCC prior to the sale of land to the Fallons by a third party; the same third party which sold the land to DCC. In other words there was an accurate survey available and it had been retained. A crucial matter for if that evidence had not been available then there was nothing to gainsay the dotted line on the Title Plan as being in the wrong position and there was nothing in the register for the Fallons other than a planning application plan which the Adjudicator held was not as likely to be correct as the pegging out by DCC following their survey. In my opinion there is often a presumption by clients that a planning application plan shows the correct legal boundary, unfortunately this is not always the case and this itself often leads to boundary disputes.

The crunch comes at paragraph 37 of the High Court's judgement when in answer to the question, 'What then would be the purpose of altering the register?'

The Fallons were in possession of a strip of land belonging to DCC and altering the register would not be 'more accurate' except in the limited sense that it would accord with the paper title but not accord with what was now on the ground and according to the Adjudicator 'would not lead to greater clarity but only confusion'. This was because a garage had been built on the disputed strip and so any topographical survey would be at odds with a more accurate plotting of the Paper Title on the filed plan.

The exceptional circumstances which the Adjudicator found and which were upheld by the High Court were that 'altering the register would only redraw the general boundary in a place which more accurately reflected the paper title but might not more accurately reflect the de facto right to the land.'

Paper Title depends on the pre-registration deed plan and not the Title Plan and that a more accurate general boundary may not add anything to the matter if it does not accord with what is on the ground and reflect the parties' rights which exist at that time on the ground. Or, put another way, just making a more accurate map or plan may not lead to more precise determination of rights in the land.

9. Common Law – Legislation and Case Law

Paper Title depends on the pre-registration deed plan and not the Title Plan and that a more accurate general boundary may not add anything to the matter if it does not accord with what is on the ground and reflect the parties' rights which exist at that time on the ground (Calvert 2010). Or, put another way, just making a more accurate map or plan.

Whether the INSPIRE directive enables more effective and efficient transfer of mapping details to those parties involved in defining, buying or selling land will be for time to tell.

10. And now?

The LRA 2002, LRR 2003 and also the continuing dialogue between OS and LR is leading to more consistency in the interpretation of mapping as well as a greater requirement for conveyancers to get the 'picture people' – the Chartered land surveyors involved. In some of the cases I have dealt with the best description of the deed plan was a 'toffee paper'; crumpled, hard to read, bits missing and food particles attached in places. In short changes are afoot which I expect will alter the perception of deed plans by lawyers.

In France the two national databases, IGN's and that of the cadastre are unsuitable for a land registration system and *GeoFoncier* is a project designed and promoted by the *Ordre des Géomètres-Experts* (OGE) in partnership with IGN and the cadastral administration in order to create and then

maintain the correspondence between the mapping systems. This will be the first and unique initiative in France towards the mapping of legal boundaries of land.

But of course, land use changes and physical boundaries change, a hedge is replaced by a fence or a large house is demolished and three or four new houses are built within those grounds – and land is sold ‘off plan’ (see Chadwick & Ors v Abbotswood Properties & Ors [2004] EWHC 1058 (Ch) for a glimpse of how bad things can be).

It is quite clear that the words of Lord Hoffman in Alan Wibberly Building Ltd v Insley [1999] 1 WLR 894:

‘Boundary Disputes are a particularly painful form of litigation. Feelings run high and disproportionate amounts of money are spent. Claims to small and valueless pieces of land are presented with the zeal of Fortinbras’s army¹.’

Let us hope that, in quoting Gustave Flaubert, ‘*Earth has its boundaries but human stupidity is limitless*’ we are not being hostages to fortune.

References

Calvert, C (2010) Best laid plans, New Law Journal, 26 February 2010, p315

Gray, K (1993) Elements of Land Law, Butterworths, London.

¹ In Shakespeare’s play Hamlet, Fortinbras, seeks to access a piece of land in the middle of Denmark, land worth nothing but a few coins.

Constructing the notion of market value of a property in the context of liability of valuer.

Prof.dr hab. Stanisława Kalus
 State Council of Immovable Property
 At the Ministry of Infrastructure,
 University of Silesia in Katowice
 e-mail: stasia@onet.pl

Abstract:

Problems with defining market value are especially significant in the countries which operate in the conditions of market economy only since recently, but concerns also other states which have not sufficiently regulated the notion. In the context of cross-border transactions and the EU freedom of capital transfer the problem of a uniform definition should be addressed as well. In Poland the 1997 Act on property management introduced the definition of market value. After a number of amendments the Act speaks of the most probable price possible to obtain on the market in consideration of actual transaction prices between willing parties who act at an arm's length, once the time necessary to expose the property on the market has lapsed, which definition conforms with sector EU solutions. Such an approach is based on the property's current use.

Quite conversely, International Valuation Standards promote assessment based on the highest and best value, which might in practice lead to the situation in which the buyer pays a much higher price than it would be the case if the valuation was based on the current use of the property, and then incurs further costs in order to achieve the optimal use. In most cases such proposals seem unacceptable and cannot be implemented in many national systems. As a result, the universal scope of the problem and the necessity to assure free transfer of capital in the Community definitely calls for a legal regulation on the level of the EU, preferably in the form of Directive which should set uniform standards of valuation.

Outline :

1. The notion of market value of immovable property in Polish law in the light of its definition in professional standards applied in European countries and the problem of applicability of foreign standards by Polish valuers.
2. Market value and other types of property value in the situation of economic crisis and the influence of the crisis on the approaches towards market value.
3. *Current use* and *the highest and best use* premises and the purpose of appraisal.

4. Legal status of a valuer in Polish law and the scope this profession.

Definition-related problems connected with market value are extremely crucial especially in the countries which have only since recently functioned in the conditions of market economy. Also in other states which have not experienced a period of socialist economy, but failed to legally regulate this concept, the question of definition raises many doubts. This seems to be a result of uneven development of property markets in such places, co-existence of a number of legal systems which regulate various rights to immovable property within one environment, provide for rights of the *in rem* and obligational type. The situation may be different in the systems of statutory or common law.

In Polish conditions the issue at play has a specific character, as in the Polish context private ownership of immovable property has never been eliminated, even within the framework of the so called communist economy. Nevertheless, it has to be admitted that free market of immovable properties did not function here, and questions of assessing value of immovable properties were most frequently connected with the problems of expropriation of private entities in favor of the State Treasury. Nevertheless, valuation of properties has always been necessary, for instance for the purposes of establishing the value of inheritance where properties made a part of the patrimony, in disputes over the ownership of immovable property or co-ownership cases.

Before the entry into force of the Immovable Property Management Act of 1997¹ value of properties was assessed according to the cost approach. As a result, replacement value of the property was used, since the notion of market value – for the lack of the market itself – could not function. Even then certain references to prices in actual transactions were taken into consideration, which however was not tantamount to the application of market value. The Act of 29 April 1985 on the management of immovable property and expropriation of immovable properties² included art.38(1) providing that proper persons assess the value of undeveloped and developed land, as well as buildings and other devices or apartments situated on that soil. The value of land was to be determined in reference to current prices in land transfer transactions, expenditures made by the possessor, the plot's function provided in the respective plan of spatial development, location of the land, access to communal devices, electricity and gas, as well as development of the land. In the same manner buildings, other devices and apartments were appraised.

The switch into the system of market economy took place at the turn of 1989 and 1990. One should trace the emergence of Polish property market back to this transition. Transformations led to the increase in property value and growth of resources active on the market as a result of privatization of state-run enterprises, emergence of municipal property and alienation of communal immovable assets to natural and juridical persons. This situation called for the necessity of introduction of new

¹ Act of 21 August 1997 on the management of immovable property /Dz.U. (Journal of Acts) No.115, pos.741/ as amended.

² Act of 29 April 1985 on the management of land and expropriation of immovable property /Consolidated text. Dz.U. of 1991 No.30, pos.127.

principles of immovable property valuation to the Polish system, based on completely novel terms, taking into account the level of market development and intense activeness of many private and public entities in the sphere of property transfer. It was necessary to apply the principle of autonomy of will of the parties to a contract, introduced to the Polish Civil Code in art. 353¹. Parties may currently frame their contractual relationship freely, which includes the decision about the price made at the parties own discretion, as long as the content and purpose of the contract do not stand in opposition with the nature of the legal relationship, mandatory provisions of law or principles of ‘social intercourse’.

The first inclusion of market value in legal documents, its definition and classification of particular (sub)types appeared in the Professional Standards for Valuers published by the Polish Federation of Valuers’ Associations in 1994. The Standards were based on similar British solutions (The Red Book), American uniform standards of professional practice of property assessment (The White Book), and the Directives on appraisal of fixed assets TEGOVOFA, known as “The Blue Guide”³. The first edition of the Polish Standards defined market value of a property as “predictable price which may be obtained on the market on the following terms:

- a) parties to the transaction are at an arm’s length, act reasonably and are not motivated by any special reasons,
- b) the parties are decided to conclude the contract,
- c) they are aware of the co-existing conditions which affect the value of the immovable property,
- d) no compulsory situation,
- e) the period necessary to expose the property on the market and negotiate proper conditions of the contract has lapsed, and the parties took into account the character of the property and its market situation” (Standard III.1.1.)

These Standards enumerated a number of (sub)types of market value, namely: for the current use, for alternative use, for optimal use (*highest and best use*), for a forced sale, and finally for future sale. Obviously, the most broadly promoted and popular of all subtypes was the first one, based on the current use of the immovable property at the moment of preparation of the valuation report (*current use*).

The development of standards, especially those of international character, led to the elimination of majority of these subtypes of market value. In consequence, at present one can only differentiate between market value and so called non-market values. The latter category includes the so called fair value, investment value, loan mortgage value, insurance value, value for alternative use, value for forced sale, and finally value reflecting replacement-cost depreciation (which should be considered a non-market but replacement type).

³ Kalus S. (2002) Status prawny osób zajmujących się zawodowo rynkiem nieruchomości, Lexis Nexis, Warszawa, Kalus S. (2009) Pozycja prawna uczestników rynku nieruchomości, Lexis Nexis, Warszawa,

The Immovable Property Management Act in its first version, which entered into force on 1 January 1998, defined market value of immovable property in art. 151 as **the property's predictable price, possible to obtain on the market, determined in consideration of market transaction prices obtained while the following conditions are met:**

1) parties to the contract were at an arm's length, did not act in a compulsory situation and had a definitive intention to conclude the contract,

2) the time has expired necessary to expose the property on the market and negotiate proper contractual terms.

Defining the value as "predictable price" was not accurate, because such an approach blurred the undoubtedly observable distinction between the value and price. While it is true that value serves as basis for determining the price, the resulting price after the negotiations between the parties of a transaction may be eventually different. The actual price may be higher than the estimated value if the seller finds a buyer who is especially interested in purchasing this particular property. It may prove lower, if the buyer manages to find arguments convincing to the seller in favor of lowering the price below the estimated value.⁴

The Polish Professional Standards for Valuers amended in 2002 included a definition of market value analogous to their first version of 1996 and provided that this type of value is determined by **the most probable price of the immovable property possible to obtain on the market while the following conditions are met: the parties act at an arm's length, without any special motivation and have a definitive intention to conclude the contract** (which eliminates offer prices – annotation by S.K), **they are aware of the co-existing circumstances which affect the value of the immovable property and do not act in a compulsory situation** (which eliminates sale in the course of execution proceedings – annotation by S.K.), **and the period of necessary exposition of the property on the market after proper advertisement, as well as the period necessary to negotiate contractual terms, taking into account the character of the property and condition of the market, have expired.**

In doctrinal writings it is presumed that the above extension of the definition of market value is undoubtedly correct, even though all of its elements could be deduced from the text of the Immovable Property Management Act, already in force, and the Ordinance delegated by this Act.⁵

In the following amendments of the Immovable Property Management Act in the years 2002 – 2010 the provision referring to the definition of market value was framed in a different way: Art.151 .1.: **Market value of an immovable property makes its most probable price possible to obtain on the market, determined in consideration of transaction prices, making the following presumptions: 1) parties to the contract were in an arm's length position and did not act in a**

⁴ Panek G. (2009) Wartość rynkowa nieruchomości w warunkach polskich oraz w świetle europejskich i międzynarodowych standardów wyceny, Nieruchomość nr 4/2009, Katowice

⁵ Panek G.[4], oraz Rozporządzenie Rady Ministrów z dnia 21 września 2004 w sprawie wyceny nieruchomości i sporządzania operatu szacunkowego /Dz.U. Nr 207, poz.2109 ze zm./ dalej cytowane jako RWN

compulsory situation, they had a definitive intention to conclude the contract; 2) the time necessary to expose the property on the market and negotiate the terms of the contract has lapsed.

This definition is approved of in legal doctrine⁶. Scholars stress its conformity with the definition contained in Directive 2006/48/EC regulating the functioning of credit institutions, as well as with previous Directives of the Council 91/674/EEC and the Directive on annual consolidated financial reports by insurance companies⁷. According to these pieces of Community legislation, **market value denotes an estimated amount for which an interested buyer should purchase a given immovable property from the owner interested in selling it on market terms, where the seller made proper efforts in order to find the best buyer and each of the parties acted reasonably and without duress. This value should be clearly and transparently documented.** However, there are a couple of EU sector directives covering relevant areas. Each is limited because of its scope of regulation. As a result, there arises a question if and to what extent, as well as which definitions should be considered binding for the national law-maker, since the national legislator is obliged to implement EU directives into the national legal system.

This definition of market value was received in the European Standards of Valuation of 2009⁸ prepared by the European Group of Valuers' Associations which document reflects the tendency to frame legal provisions under the influence of EU legislation. It must, however, be stressed in advance that TEGoVA has no pretences to influence directly the legal system of the state where valuation is made. This is expressed transparently in Standard 3.1 pursuant to which the definition included in the standards should be only applied **unless "provisions of law of a particular state provide otherwise."** In consequence, it is beyond any doubt that the Standards can impose neither the definition nor the way of estimating market value on anyone.

The above is linked to more general issue of the relation between international standards and national standards accepted in particular countries. There arises the question what standards should be binding for Polish valuers with the view to the process of integration and Poland's accession to the European Union. In particular, there is the problem whether one can assume that the binding standards are the ones mentioned above, that is the European Standards of Valuation elaborated by TEGoVA in 2009, or the International Standards of Valuation by the International Valuation Standards Committee, recently published in Poland (IVSC).

Moreover, this is not the sole source of doubt as to the possibility and necessity to apply these standards in Polish conditions. Another obvious question to ask would be about the legal status of International Standards, especially whether or not they belong to European, or Community law, and

⁶ Kalus S.(2009) [[3], Wilkowska-Kořakowska,(2010) Wycena nieruchomości. Operat szacunkowy .Rzeczoznawstwo majątkowe. Lexis Nexis, Warszawa.

⁷ Panek G. [4], (31)Bienieć G.Kalus S., Marmaj Z.Mzyk E.(2008) Ustawa o gospodarce nieruchomościami, Lexis Nexis, Warszawa.

⁸ European Valuation Standards(2009) , sixth edition. TEGoVA Gills

what are the consequences of such classification, if there is any official need to implement these into the Polish legal system or apply them directly. The Accession Treaty of Poland into the European Community is a source of the obligation to implement respective Community legislation in our national legal system, or to apply the EU law directly, accepting its precedence over national enactments. In earnest, neither TEGOVA European Standards of Valuation nor IVSC Standards would qualify as Community law, since they were not enacted by the Community institutions and were not adopted in the form of EU directives or, what is more, regulations. One can only treat these as norms which constitute a part of the so called European law *sensu largo* which has little direct influence on the system of law in the EU Member States. The obligation of implementation in national law concerns only Community legislation enacted by the Community legislative bodies. Among such pieces of legislation one can enumerate regulations, directives, recommendations and opinions. It is beyond any doubt that neither the International Valuation Standards Committee (IVSC) nor TEGOVA can be viewed as Community institutions, even though they are organizations gathering representatives of many countries, including non-EU members.

In the literature of the subject it has been noticed that harmonization of the principles of valuation in the UE Member States is an absolute necessity. It follows especially from the Treaty of Rome which provides for the free transfer of people within the Community, as well as free transfer of capital.⁹ If firms are willing to invest in other EU Member States, they must be aware on what terms and for what price they can acquire immovable properties necessary for their business activity. As a result, the appraisal of immovable property for the purposes of business activity in various EU Member States should follow uniform rules.

Although none of the International Standards can be recognized as EU directives, they often invoke such directives and in that matter one may assume that the documents themselves should indicate to the EU Member States the need of their application as an implementation of the quoted directives. Such standards are also significant for national legislators who enact on certain areas relating to the performance of the profession of valuer in a particular state with regard to property appraisal. However, one should bear in mind that EU directives are directly applicable in such countries only where respective authorities failed to implement in due time the EU Directive into the national legal system, and if the concrete provision of EU legislation is suitable for direct application.

As a result, international standards may set a certain pattern for valuers in particular EU Member States, still this pattern needs a lot of adjustments because of the specific conditions of national legal systems. In conclusion, the goals of international standards are as follows: to offer aid for valuers at the elaboration of understandable valuation reports by means of providing clear hints and recommendations; to promote consequent application of the definition of market value included in the standards in all Member States of the EU, as well as standardization of approaches towards

⁹ Mankowski, P. (2010) Der Immobilienwertungsvertrag in europäischen Internationalen Privat- und Prizzessrecht, Festschrift für Stanislaw Kalus, Ius est ars boni et aequi, Peter Lang Verlag, Frankfurt am Main.

appraisal; to improve professional skills of valuers and promote their **best practice**); to provide precise grounds for economic analyses and effectiveness of use of the limited property resources, consideration of the client needs and requirements of due appraisal in valuation reports; to promote the rank of the profession of valuer and its role; to implement procedures aimed at precise estimation of value which conform with legal systems of **particular states and the EU legislation, general as well as sector solutions.**

2. Market value and other types of value referring to immovable property in the conditions of economic crisis and the impact of the crisis on the approach to market value.

The phenomenon of economic crisis has been well known and on many occasions described in the literature of the subject.¹⁰ Also the influence of previous financial crises of global range has become known in Poland, e.g. Asian, Scandinavian, American crisis of savings and crediting, the Tequila crisis in South America. However, we have not thus far experienced a crisis of such scale on property market in connection with the troubles of financial institutions and a natural decline in the business cycle. None of the so far recognized and well described mechanisms of crisis emergence, consisting in the dependency between crediting immovable properties in the conditions of wide credit accessibility and disturbances in the financial system, has caused such devastation as the present crisis. As a consequence of previous crises, certain minimal standards have been elaborated as regards measures adapted by banks in the area of immovable property financing, however, observations of the banking sector lead to the conclusion that such policies are not and were not observed in 2008, in the period directly preceding the crisis.

Undoubtedly, provisions of Polish banking law establish a certain normative framework preventing conclusion of this type of crediting contracts which already at the stage of negotiation create a danger of unsatisfied receivables of the bank stemming from the allowed credit.¹¹ It follows from the provisions of the Act on banking law that the procedure of allowing credit involves two stages and requires in the first place the examination of credit capacity of the borrower (even though banks may allow credit to persons without borrowing power), and then the conclusion of a credit contract in which a measure of securing the allowed credit must be indicated. Most frequently, especially in relation to natural person borrowers, the collateral is mortgage on immovable property

¹⁰ Handorf W.(1998) Regulacja i nadzór nad finansowaniem hipotecznym, USAID , Warszawa, Kalus S. (1996) Znaczenie tzw Raportu Mallinsona dla zawodu rzeczoznawcy majątkowego, „Rzeczoznawca Majątkowy”, Warszawa 1997/1.Kalus S.(2002) Der Zusammenhang zwischen der Wirksamkeit der Sicherung und dem Wert des sicherungsstandes aus der Hinsicht des polnischen und europäischen REchts, Sicherung von Forderungen und deren Befriedigung, Slovakischen Akademie der Wissenschaften, Smolenice 2002, Manatshefte für Osteuropäisches Recht, 2003,nr.1., Kalus S.(2009) Rzeczoznawcy a banki, Śląskie Stowarzyszenie Rzeczoznawców Majątkowych, Katowice, Nieruchomość, nr.4/2009.

¹¹ Kalus S. (2008), Umowy zabezpieczające wykonanie umowy kredytowej, [in:] Europeizacja prawa prywatnego, t. I., Wolters Kluwer business, Warszawa, Narożny T. (2000) Czynności kredytowe banku. Zagadnienia prawne, Warszawa, Strzelecki A. (2005) Zarys prawa bankowego, Włocławek.

which belongs either to the borrower himself or another person.¹² As regards this means of security, the basic problem is to examine whether the property charged with mortgage has a sufficient value to enable satisfaction on the part of the bank if the borrower should not comply with the contract and fail to pay back the credit sum plus the cost of its allowance. The decision of the bank in that matter requires that a valuer should determine in advance whether the property is apt as security for the receivable of the bank, and appraise the immovable property.

Valuations carried out for banks are an object of disputes in Poland, both in the doctrine of private law and among valuers confronted, or rather collaborating, with banks and the Banking Supervision Authority (currently – after the liquidation of the latter institution – with the Polish Financial Supervision Authority).¹³ This problem is not resolved in a uniform way in the banking practice, which is a consequence of the fact that many miscellaneous banks operate in Poland. These banks apply varying solutions when it comes to the practice of valuation. The most controversial part was the question regarding mortgage bank actions on the basis of the Act on letters of charge and mortgage banks.¹⁴ When it comes to assessment of loan mortgage value of an immovable property which secures a receivable of the mortgage bank which issues letters of charge (a type of security), scholars emphasize the need to carry out especially conservative estimates which are to take into account a high level of risk of the bank connected with the property charged as collateral. The Act stipulates precisely that by **loan mortgage value of an immovable property one should understand the value which in the view of the mortgage bank** (and not the valuer – annotation by S.K.) **reflects the level of risk connected with the property charged as collateral** (art.2). In consequence, **estimation of this value should be carried out with due diligence and caution, in consideration of only those features of the property, and incomes connected with the property, which may be durable in character and can be obtained by every possessor of the property** (art.22).

Consequently, it was obvious that in the light of this enactment loan mortgage value had nothing in common with market value of an immovable property, or replacement or catastral values. In the light of the above there was no possibility of its assessment by a valuer whose scope of competence was laid down in art. 150 of the Immovable Property Management Act. Loan mortgage value fell outside that range. However, it turned out shortly that without collaboration with valuers the estimation of this value was extremely difficult. As a result, the amendment of the Immovable Property Management Act of 2000 added pt. 4 to art. 150(1) according to which valuers became

¹² Other means of securing credit are in particular collateral guarantee and aval, blank hand bill, assignment for collateral purposes, access to debt, assumption of a debt, irrevocable power of attorney, banker's letter of credit, bank guaranty, blocked funds, bank caution, registered pledge and ordinary lien, transfer of ownership for collateral purposes – Kalus S. [10]

¹³ Blicharz R. (2009), Nadzór Komisji Nadzoru Finansowego nad rynkiem kapitałowym w Polsce, Oficyna Wydawnicza Branta, Bydgoszcz-Katowice 2009

¹⁴ Act of 20 August 1997 on letters of charge and mortgage banks /Dz.U. of 2003, No.99, pos.919 as amended./, Drewicz-Tułodziecka A. Łosiak – Szewczyk M., Obradović Z., Orzechowska J. (2001) Bankowo – hipoteczna wartość nieruchomości, C.H.Beck, Warszawa, Bieniek G. Kalus S.Marmaj Z., Mzyk J., (2008) Ustawa o gospodarce nieruchomościami, Komentarz, Lexis Nexis Warszawa, Wyd.2.

entitled to give a **proposal of the loan mortgage value of a property**. Specific rules of calculating this value on the part of banks followed as well from Recommendation F issued by the Bank Supervision Authority. The regulation set procedures which had to be observed by mortgage banks while estimating this type of value. Currently, the principles of valuer participation in the appraisal of that value have been modified. Pt. 4 of art.150(1) of the Immovable Property Management Act has been repealed. Instead, the legislator introduced pt. 5 in the discussed article providing that **valuers may assess other types of property value stipulated in separate enactments**, and added pt. 3a to art. 174, which entitles the valuer to **elaborate analyses and expert opinions other than valuation reports, concerning for instance the loan mortgage value of immovable property**.¹⁵

This shows that caution rules directly affect the safety of transactions and protection of the interest of the credit grantor, which might have contributed in Polish conditions to the relatively moderate forms of presence of the worldwide crisis which originated in the American context from the crisis of many renown large banks.

3. Current use and highest and best use premises and the purpose of valuation.

As already mentioned, the EU Directive 2006/49/EC¹⁶ introduced very cautious rules concerning the assessment of market value of immovable property, which seem to be in conformity with Polish provisions covering this issue. The Directive is significant not only when it comes to securing credit liabilities, but in general in relation to all transactions connected with the transfer of ownership of immovable property.

The Immovable Property Management Act and a respective Ordinance delegated by this statute stipulate that the assessment of market value of a property can take place only by means of elaboration of a valuation report on the part of a valuer (art. 7 of the Immovable Property Management Act), and that the valuer autonomously and independently chooses the proper approach, method and technique of assessment, taking into consideration **“in particular the purpose of valuation, type and location of the property, function provided in the local spatial management plan, development of technical infrastructure, condition of the property’s development and available data on incomes and prices of similar immovable properties. For the lack of the local spatial management plan ... the factual use of the property is taken into account.”** This demonstrates undoubtedly that the

¹⁵ Kalus S. [13]

¹⁶ The Directive was amended by the Directives of the European Parliament and Council of 14 June 2006 on capital adequacy of investment companies and credit investments, the Directive 2009/111/EC of 16 September 2009 amending Directives 2006/48/EC and 2007/48/EC in relation to banks connected with central institutions, certain investment funds, large exposures, agreements on supervision and management in crisis situations, and finally the Directive of the Commission 2009/83/EC of 27 July 2009 amending Appendices to the Directive 2006/48/EC of the European Parliament and the Council in relation to technical provisions on risk management – Prystupa, M., (2010) Standardy wartości w procesie zabezpieczenia wierzytelności kredytodawcy, “Rzeczoznawca Majątkowy”, Polska Federacja Stowarzyszeń Rzeczoznawców Majątkowych, Warszawa No. 1/2010.

Polish statute prescribes that market value should be determined where the purpose of valuation is the alienation of property according to its use on the date of valuation. Obviously, the above does not eliminate the assumption of another basis of the estimation of value if the purpose of valuation differs, for instance in reference to developer properties where the purpose of valuation may be to determine the highest and best use. However, here the value of the property will be of a different type, other than market value referring to the possibility of sale of a particular immovable property on the market. This shall especially cover *investment value (worth)*, and may also be associated with the notion of *special value*, in the case of which one may speak of a surplus as compared to market value.¹⁷

The Polish approach conforms with the 2006/48/EC Directive which provides that market value stands for the estimated value for which an interested buyer should purchase a given property from the seller interested in the selling it on accepted market terms, having previously made proper efforts to find the best buyer, where each of the parties acts reasonably and without duress. Moreover, one of the following provisions of the Directive provides that the value should be calculated in a satisfactory way on the basis of **safety standards of valuation**, established in pursuance of statutory, delegated or administrative rules (Art. 115). If this is the case, there is no reason to doubt that only a valuation taking into account the current use of the property meets this safety condition.

It follows even more apparently from the International Standards of Valuation,¹⁸ according to which market value is an **estimated value which can be obtained on the valuation day for an asset**, assuming that the parties have a definitive intention to conclude the contract, act in an arm's length, reasonably and with an insight of the situation, they are not in a compulsory position, and the period has lapsed necessary to expose the property on the market. This is another provision which refers not to a use alternative to the current (one on the date of valuation), but expressly points to the current use.

However, in the last two years there has emerged, and become frequently applied, a concept of interpretation of market value **according to the highest and best use**. The international Standards of Valuation stipulate that market value should be regarded as the value of a property for its optimal use. The Standards prescribe in particular that **the unique features of the land determine its optimal use**. This would however entail that it is necessary, in order to carry out valuation in accordance with this method, to find certain specific features of the property which justify such a procedure. While it is argued by scholars¹⁹ that every immovable property (not only land, but also a building or apartment if they make an object of ownership separate from the land) is different, and that every property has potentially a different optimal use, the acceptance of such an assumption could lead to the absolute lack of clarity as to what use, even in relation to similar plots, is optimal. Scholars emphasize as well that the value of a property is affected not only by its use, but also **physical and environmental**

¹⁷ Kucharska-Stasiak, E. (2006) Warszawa, Nieruchomość w gospodarce rynkowej, Wydawnictwo Naukowe PWN.

¹⁸ Międzynarodowe Standardy wyceny (ISV), PFSRM (Polish Federation of Valuers' Associations) Warszawa 2007

¹⁹ Kucharska-Stasiak, E., (2006) [16]

factors, such as location, **economic factors**, such as condition of economic development, phase in the business cycle, level of unemployment and prosperity of the population, diversity and access to various forms of credit, interest rates of the credit, prices on the market, earning capacity in the light of profitability of other investments. Also **legal factors** are of crucial importance, such as the freedom of contract, taxation of the transaction, system of tax reliefs, spatial planning, enactments on environmental protection. Similarly, **demographic and social factors** and **political** ones are at play.²⁰

The International Standards clearly promote the value based on the optimal use as market value. The optimal use is understood as **the most probable and physically possible use of a property, properly justified, legally admissible, feasible in economic terms and accounting for the highest value of the assessed property.**

This definition seems rather imprecise and easily questionable in a dispute concerning valuation. It would be extremely difficult to make pronouncements as to the probability of a given use of the property if it is subject to market transactions and no steps have been taken to change the current use. Also the premise of “physical possibility” seems highly indeterminate. In general, every use of a property is possible where the owner has considerable financial means. I cannot agree as well with the requirement of due justification of the use. Who would be entitled to evaluate such a justification? What factors should he consider? How could his opinion be better grounded than valuer’s opinion? These are all questions which are not general in nature, but the answer to them results in valuer liability for a defective valuation report, both professional responsibility (in the sphere of valuer’s professional entitlements) and civil law (compensatory) liability.

Beyond any doubt, the **requirements of optimal use of a property stemming from the definition are variable.** What happens if the presumptions referring to the optimal use change entirely, for instance as a result of economic or political changes? This is a very concrete question if we bear in mind the British crisis on the property market at the end of the 80’s and Mallinson’s Report which depicted the consequences of this crisis for the practice of valuers.²¹ The report expressly unveils that assumptions as to the future use of certain immovable properties (e.g. the City of London) may prove totally failed and the results of such prognoses can affect the owners who adjust their own properties (developers) to the expected, but impossible to achieve as a result of political changes, incomes from this real estate.

It is also difficult to accept the thesis presented in the Standards that the “**concept of highest and best use is** (or rather should be – annotation by S.K.) **the fundamental and integral part of the assessment of market value.**” One has to realize that in certain countries, including Poland, valuers are asked to prepare a valuation report most often in the course of tender proceedings where the most frequent criterion of selection is the very price of the valuation report. Consequently, if the valuer is to

²⁰ Ibidem

²¹ Kalus, S. (1996) N Znaczenie tzw. Raportu Mallisona dla zawodu rzeczoznawcy majątkowego, „Rzeczoznawca majątkowy” Warszawa 1996, Nr 3.

appraise a number of apartments of similar surface and location within the same building, it seems peculiar to ask of him that he should carry out a separate analysis for each of these apartments. The fee for such valuation reports would have to be incommensurately high.

The decisive argument against the statement presented in the Standards is the fact that in order to achieve the optimal use the buyer has to invest serious financial means. It seems obvious that each investment in property consumes capital. **Why would the buyer have to pay first a price disproportionately high in relation to the current use of the property and then pay additional sums in order to achieve the optimal value of the property?**

European Valuation Standards elaborated by TEGoVA take a much more moderate stance in this respect. As for the market Value, European Standards assume that it is the **best price to reasonably obtain by a seller and the most favorable price reasonably achieved by the buyer**. When it comes to the value for alternative use, which is treated as a non-market value, the Standards define it as **the value assessed without assuming continuity of the current use**.²²

Bearing in mind the above dilemma regarding the basis of assessment of market value, I am a strong supporter of the previous approaches as better suited to the position of immovable property in modern property management and the property market as such, functioning as a part of the general market. Additionally, the previous approaches did not provide grounds for excessive expectations towards valuers and valuation reports.

4. Legal status of a valuer in Polish law, scope of this profession and valuer's liability.

Property valuation is a profession in Poland regulated by the provisions of the Immovable Property Management Act of 21 August 1997 /Dz.U. (Journal of Acts) No. 115, pos. 741 of 1998; No. 106, pos. 668 of 1999; No. 49, pos. 484 and No. 86, pos 966 of 2000; No. 6, pos. 70/. There are also many doctrinal pronouncements concerning the legal position of this profession and professional responsibility of valuers.²³

The main function of valuers, following from the quoted statute, consists in assessment, estimation of value of immovable properties, which corresponds to the fundamental role of professionals called **valuers, or chartered surveyors** in the English legal system.²⁴ By the notion of property valuation the statute understands procedures in the course of which the value of property becomes estimated together with installations and devices permanently attached to the real estate. The result of valuation

²² Prystupa, M. [15]

²³ Cf. S.Kalus [in:] Bieniek G.Kalus S., Marmaj Z. Mzyk E. : Ustawa o gospodarce nieruchomościami, Komentarz, Lexis Nexis, Warszawa 2008, p. 558 ff., S.Kalus, Status prawny osób zajmujących się zawodowe rynkiem nieruchomości, Lexis Nexis, Warszawa 2002, pp. 58-133, S.Kalus, Pozycja prawna uczestników rynku nieruchomości, Lexis Nexis, Warszawa 2009, p. 59-152, D.Wilkowska- Kołakowska, Wycena nieruchomości. Operat szacunkowy. Rzeczoznawstwo majątkowe, Lexis Nexis, Warszawa 2010, p.130.

²⁴ Cf. J.W. Hendrikse, L.Hendrikse (2009), The Valuations Handbook. Concepts and Cases, Lexis Nexis, s.39,

is a valuation report prepared by the valuer, which also corresponds to the solutions accepted in the English context.²⁵

Pursuant to art. 156 of the Immovable Property Management Act, valuation report is a written opinion of the valuer concerning the value of property. It is a document, since every object which may serve as proof of legally relevant facts is regarded as document. Valuation report may for instance be a proof that the assessed property has a given value, and this value in turn serves as a basis for determining its price. In consequence it must be accepted that a valuation report is a document.²⁶ It is a private document whose content is regulated by the delegated Ordinance of the Council of Ministers.²⁷ It follows from the provisions of the Ordinance (§ 56) that each valuation report should mention the technique of property appraisal, which comprises: determination of the object, scope and purpose of valuation, its formal grounds and sources of the substantive data, dates important for the estimation of value of the property, description of the property condition, including its legal status and practical use, suitability for use, analysis and characterization of the property market as regards the purpose and method of valuation, the type of estimated value, choice of approach and technique of assessment, and eventually a presentation of calculations of the property value and the final result of the appraisal with motives.

Polish valuers enjoy a broader scope of professional competences than those referring to appraisal of immovable property. They are also entitled to prepare elaborations and expert opinions other than valuation reports. Such documents concern the property market and professional counseling as regards this market, effectiveness of investment in property and its development, financial consequences of enacting local property management plans, determination of objects to separate apartment ownership, loan mortgage value of immovable properties, property valuation for the purposes of an individual investor, appraisal of properties included in investments, or valuation of properties as fixed assets of entities in the understanding of the Act on accountancy.²⁸

As a result, the range of activities of professional valuers is delimited by statute in an exhaustive manner, which entails that apart from the enumerated entitlements, other competences cannot be legally acknowledged. When it comes to confronting the profession with problems discussed above, I would like to stress that valuers are not competent to assess the loan mortgage value in the form of valuation reports, but only in an opinion or expertise. Determination of loan mortgage value is a task of a bank trustee. As far as evaluation of effectiveness of property investment and developing is concerned, the participation of a valuer does not consist in elaborating a valuation report, but only opinion or expertise which is usually going to contain an analysis based on the highest and best use of

²⁵ Cf. European Valuation Standards, sixth edition Gilles, Standard ESW 5, oraz Międzynarodowe Standardy Wyceny (International Standards of Valuation) Warszawa 2005,

²⁶ Kalus, S. [22]

²⁷ Of 21 September 2004 on valuation of immovable property and elaboration of valuation reports, /Dz.U. No. 207, pos. 2109 as amended/

²⁸ Kalus, S. [22]

the property. This use is to be optimal from the perspective of a particular investor, hence it would be impossible to apply directly the interpretation of the International Standards in reference to this notion.

The activity of a valuer, both in the area of valuation reports and expert opinions is reviewed from the point of view of the valuer's responsibility. As has been mentioned above, this responsibility assumes the shape of both professional and civil law liability. The differences between these two boil down to the fact that the former consists in administering professional sanctions, namely disciplinary penalties for defective performance of the valuer's professional activities, which task does not encumber courts of law but a licence body, the Minister of Infrastructure. We may speak of professional liability regardless of whether the client of a valuer has suffered any damage. Sanctions which can be imposed on the valuer are: admonishment, rebuke, suspension of professional competences or deprivation of these entitlements. This means that the sanctions are very severe, they may lead to a ban on performing the profession.

Should the client of the valuer suffer any damage as a result of defective contractual performance on the part of the valuer, the latter becomes subject to compensatory liability based on general provisions regulating contractual liability expressed in art. 471 and the following of the Civil Code.²⁹ While it is true that Polish valuers are obliged to insure their professional activity, their insurance, regulated by the Ordinance of the Minister of Finance³⁰ is not sufficient, because the Ordinance prescribes a minimal guaranteed sum in relation to a single event, and other all events the results of which are covered by the contract of civil liability insurance up to the equivalent of the amount of € 25,000 expressed in PLN. Taking into consideration actual prices of immovable property it should be stressed that the guaranteed sum is not particularly high. At the same time the requirements towards the valuer are intense, which creates a situation in which performance of this profession becomes difficult and demands constant self-cultivation and professional improvement of the valuer, also when it comes to novel approaches, methods, techniques, as well as other factors which fall outside the scope of activity of valuers in other legal systems.

Conclusions:

1. The problems of defining and settling the notion of market value of immovable property and the methods of its assessment is of primary significance for the harmonization of property transfer transactions and securing credit liabilities with immovable property in the EU Member States. Undoubtedly, the need to **harmonize this notion** is an absolute necessity and should take place in the form of a Community Directive. Existing International Standards are elaborated by various panels of experts and neglect statutory provisions of the EU Member States covering these issues. The standards contain different provisions referring to specific questions, which makes their application impossible in particular countries, especially in the

²⁹ Act of 23 April 1964 /Dz.U. No. 207, pos. 2114/

³⁰ Of 21 September 2004 on mandatowy insurance of the civil liability of a valuer.

countries of statutory law tradition. **The uniformity of practice of the assessment of market value may be attained only by means of introducing an EU Directive providing a consolidated definition of this value and determining approaches, methods and techniques of its assessment for the purposes of universal transactions, as well as establishing principles of proper conduct of the valuer in the course of valuation.**

2. In the situation of crisis on property markets **one should promote a very careful approach towards the assessment of market value**, because in the situation of economic crisis the demand for immovable properties is seriously impaired. As a result the predictable price possible to obtain in the market in such circumstances cannot be significant. It is connected with lesser chances of being allowed a bank credit secured by mortgage on the purchased immovable property, which reflects the level of risk linked to the particular collateral property. Especially in the countries where the creditor's satisfaction from the property does not consist in its seizure by the bank, but a compulsory enforcement sale and satisfaction from the obtained sum the estimation of value must take into account very unfavorable terms of sale of such an immovable property.
3. **It is impossible to accept**, especially in a situation of economic crisis, **the way of determining market value proposed by the International Standards of Valuation which is not based on the property condition on the date of valuation (current use), but according to its optimal, highest and best use, or the proposal that the concept of optimal use should be the fundamental and integral part of the valuation report.** If there is a person willing to buy the property on the market, he generally tends to satisfy his current needs related to the particular property the person intends to buy, and is not interested in its theoretically most beneficial use. Moreover, one has to bear in mind that the circumstances at the moment of valuation may be completely different from the situation in future. Hence, what has been determined **in the valuation report as the highest and best use** might be no longer valid once the buyer purchases the property.
4. If the assessment in every valuation report was based on the highest and best use, the seller could demand an incommensurately higher price for the immovable property than it is the case in estimation according to the property's current use. This could lead to the situation in which **the buyer would be obliged first to pay a much higher price in relation to the current use of the property, and then incur further costs, invest in the property so as to achieve the highest and best use.**
5. I consider the application of this type of basis of valuation sufficiently justified in relation to the **needs of a developer (as far as the investment value is concerned)** who in principle intends such an investment in the property which generates highest possible proceeds. As a result, the developer's intention is to eventually maximize the property's market value in comparison to the property's value on the date of its purchase.

6. Since such basis of valuation requires investigations with regard to the optimal use of immovable property – and changes of economic conditions as such are not easily predictable – the universal **application of this method could result in valuer's liability** as provided in particular national legislations and national insurance provisions. In such a situation **valuers would be excessively encumbered with damages** for undue anticipation of the future value of the assessed property.

References :

1. Bieniek g., Kalus S., Marmaj Z., Mzyk E., (2008) *Ustawa o gospodarce nieruchomościami. Komentarz*. Lexis Nexis, Warszawa 2008,
2. Kalus S. (2002) *Status prawny osób zajmujących się zawodowo rynkiem nieruchomości*, Lexis Nexis, Warszawa.
3. Kalus S. (2009) *Pozycja prawna uczestników rynku nieruchomości*, Lexis Nexis, Warszawa.
4. Panek G. (2009) : *Wartość rynkowa nieruchomości w warunkach polskich oraz w świetle europejskich i międzynarodowych standardów wyceny*, „Nieruchomość” , Śląskie Stowarzyszenie Rzecznawców Majątkowych, Katowice, Nr.4/2009,
5. Wilkowska-Kołąkowska D.(2010) : *Wycena nieruchomości. Operat szacunkowy. Rzecznawstwo majątkowe*, Lexis Nexis , Warszawa.
6. Mankowski P. (2010) *Der Immobilienwertungsvertrag im europäischen Internationalen Privat- und Prozessrecht*, Festschrift für Stanisława Kalus. Ius est ars boni et aequi, Peter Lang GmbH, Frankfurt am Main.
7. Handorf W. (1998) *Regulacja i nadzór nad finansowaniem hipotecznym*, USAID, Warszawa,
8. Kalus S. :(1997) *Znaczenie tzw. Raportu Mallinsona dla zawodu rzeczoznawcy majątkowego* ,”Rzecznawca Majątkowy”, Polska Federacja Stowarzyszeń Rzecznawców Majątkowych . Warszawa 1997/1.
9. Kalus S. (2002) *Der Zusammenhang zwischen der Wirksamkeit der Sicherung und dem Wert des Sicherungsstandes aus der Hinsicht des polnischen und europäischen Rechts*, Sicherung von Forderungen und deren Befriedigung, Slovakischen Akademie der Wissenschaften, Smolenice 2002, Monatshefte für osteuropäisches Recht, 2003/1.
10. Kalus S., (2009) *Rzecznawcy a banki*, „Nieruchomość” , Śląskie Stowarzyszenie Rzecznawców Majątkowych, Katowice Nr4/2009,
11. Kalus S., (2008) *Umowy zabezpieczające wykonanie umowy kredytowej*, Europeizacja prawa prywatnego, t.I , Wolters Kluwer business, Warszawa.
12. Nrożny T. (2000) *Czynności kredytowe banku. Zagadnienia prawne*. Warszawa.
13. Strzelecki A. (2005) *Zarys prawa bankowego*, Włocławek.
14. Blicharz R. (2009) *Nadzór Komisji Nadzoru Finansowego nad rynkiem kapitałowym w Polsce*. Oficyna Wydawnicza Branta, Bydgoszcz – Katowice
15. Drewicz-Tułodziecka A., Łosiak- Szewczyk M., Obradović Z., Orzechowska J. (2001) *Bankowo hipoteczna wartość nieruchomości*, C.H.Beck, Warszawa.
16. Prystupa M. (2010): *Standardy wartości w procesie zabezpieczania wierzytelności kredytodawcy*,”Rzecznawca Majątkowy”, Polska Federacja Stowarzyszeń Rzecznawców Majątkowych, Warszawa Nr.1/2010.
17. Kucharska – Stasiak E.(2006) :*Nieruchomość w gospodarce rynkowej*, Wydawnictwo Naukowe PWN,

18. Hendrikse J.W., Hendrikse L. (2009) *The Valuations Handbook, Concepts and cases*, Lexis Nexis
19. *Europejskie Standardy Wyceny, European valuation Standards, sixth edition* Gilles,
20. *Międzynarodowe Standardy Wyceny*, Warszawa 2005,

Property loss due to coastal erosion: judicial, legislative and policy interventions

Mick Strack

School of Surveying

University of Otago

New Zealand

Email: mick.strack@otago.ac.nz

Abstract:

Ownership and management of the coastal margin are contested. Climate change is inducing sea level rise and increased frequency and severity of storm events. These processes are likely to put more coastal property at risk from erosion. Case law, legislation and central and local government policy in New Zealand support the general concepts of promoting the natural character of coastal land; recognising that land is not permanent and that the sea may take land; avoiding attempts to build structures against the sea; and preventing inappropriate development. However, demand for coastal property and the increasing value attaching to that property suggest that these concepts are contrary to the expectations of land owners. Local authorities are pressured to protect the property interests of their citizens, but also expected to make cost effective decisions about spending income wisely. There is central government support for implementing a policy of managed retreat, and there is strong economic logic to resist the costs of coastal defence. But inevitably, local communities are demanding continued protection, and preferably at government expense. This paper introduces some background legal issues and then discusses how law and policy is being implemented to manage the competing demands on private and public coastal land. The conflict is being played out in many coastal communities, and has not been satisfactorily resolved.

Keywords:

Coastal defences, erosion, managed retreat,

1 Introduction

Access to and protection of private property are strongly held principles in New Zealand, and legislation and case law have regularly provided protection of private property as a basic individual right. This has reinforced the emphasis on the rights attaching to land and property rather than on the responsibilities we have towards the land. The developing principles of sustainability, however, suggest that humans need to adapt to the forces of nature rather than to attempt to control or modify them for human needs.

Many areas of the coastline are now subject to significant erosion, much of it prompted by physical intervention in coastal processes (e.g. intervention in sediment replenishment from dammed rivers, modification of coastal landforms and clearance of vegetation cover), but increasingly recognised as arising from climate change and sea-level rise.

The coastal process relating to the power of the sea to change coastal topography is an inexorable and natural occurrence. A coastal hazard arises when developed and privatised coastal property is adversely affected by the sea's erosive effects.

This paper describes the judicial, legislative and policy interventions that direct how New Zealand land managers will deal with anticipated increasing encroachment of the sea over coastal property. A useful case study is presented of the Wainui Beach area in the Gisborne District on the east coast of New Zealand, where property is increasingly susceptible to loss by coastal erosion and the local authority attempts a withdrawal from the provision of coastal protection works.

2 Common law

The common law doctrine of accretion and erosion is strongly accepted within New Zealand law. Many property parcels have been created with the sea, lake or river defined as a boundary. This type of boundary is a natural riparian (or littoral) boundary, and as such, it is ambulatory; it moves with the movement of the land/water interface, and is not fixed by measurements on the ground or dimensions on a plan. One of the defining characteristics of such a boundary is that, as it moves, it may increase or decrease the area of land incorporated in the parcel, and a doctrine of mutuality applies; if land can be lost in this process then it can also be gained.¹ The process of gaining more land by the withdrawal of the water, or the building up of dry land by natural processes² is accretion, while the opposite effect (the removal of dry land or the encroachment of

¹ The North Carolina Court of Appeals in the case *Shell Island* (cited in Kalo 2005; 1489) states: "The courts of this state have considered natural occurrences such as erosion and migration of waters to be, in fact natural occurrences, a consequence of being a riparian or littoral landowner, which consequence at times operates to divest landowners of their property."

² See the case *Southern Centre of Theosophy v State of South Australia* [1982] 1 All ER 283 for some clarifying statements about the legal process of accretion; that it affects a natural riparian boundary; that it must be slow, gradual and imperceptible; that it

water over the land) is erosion. One of the significant legal tests to determine whether the legal boundary moves is whether the movement of the water/land interface is slow, gradual and imperceptible.³ While this rule is widely applied in common law jurisdictions, it is not clear why it should be so,⁴ as it produces a variety of anomalous situations.⁵ If the legal boundary does not move, either the foreshore and seabed becomes subject to private title and therefore the right of the public for freedom of access is questionable, or the title becomes remote from the sea and loses those natural riparian rights of direct access and use.⁶

In New Zealand, therefore, much emphasis is placed on where the property boundary is at any particular time; whether the boundary has moved and how it has moved; what private or public rights exist; and where those rights can be asserted. This common law scenario appears to represent the legal position as it has been decided in New Zealand. However, the common law can always be superseded by legislative intervention, indeed, the purpose of legislation is to change, or clarify any uncertainty in interpretation of the common law.⁷

3 Coastal Legislation

The Crown Grants Act 1908 may provide such a statutory override of the common law doctrines of accretion and erosion as they apply to littoral boundaries⁸ in New Zealand. The Act states (s35):

Where in any grant the ocean, sea, or any sound, bay, or creek, or any part thereof affected by the ebb or flow of the tide, is described as forming the whole or part of the boundary of the land granted, such boundary or part thereof shall be deemed and taken to be the line of high-water mark at ordinary tides.

As Kalo (2005) has discussed in respect to exactly this situation in North Carolina, “An appropriate reading of the statute is that the mean high water mark remains the seaward

must be as a result of natural processes; that it can be as a result of windblown sand; that it affects leasehold title just as it affects a fee simple title; and that it applies to inland lakes as to the sea.

³ See *Humphrey v Burrell* [1951] NZLR 262 at p267 per Gresson J in the Supreme Court trial summarizing several Privy Council decisions “It is necessary that [the accretion] should have been brought about by a process so slow and gradual as to be in a practical sense imperceptible, ...” The justification of this rule appears to rest on the maxim: *de minimus non curat lex* – the law takes no account of trifling matters.

⁴ The alternative, is that the water/land interface moves by avulsion (by a sudden change such as by storm action), in which case the boundary does not move, and either the boundary encroaches into the seabed (in the case of physical erosion), or it becomes remote from the sea (in the case of actual build up of land).

⁵ “Logically, the consequences of avulsion should be no different than those of the natural processes of erosion and accretion.” (Kalo 2005;1440)

⁶ The reason for the ambulatory nature of the boundary is “to preserve the fundamental riparian right-on which all others depend, and which often constitutes the principal value of the land-of access to the water” (Kalo 2005;1440 quoting *Lamphrey v. Metcalf* 1893)

⁷ The common law remains as the fall back position if there is no overriding legislation. Bennion (1992;112) notes: “To describe the way Acts operate on existing law one can use the image of a floor upon which rugs are spread. The floor consists of unwritten law or *lex non scripta*, in other words common law, rules of equity, and customary rules. An Act is like a rug laid down on this floor. The Act conceals, for the area it covers, the texture underneath. That texture becomes visible again if the rug is later removed, that is that Act is repealed.”

⁸ littoral – adjoining the sea or tidal waters

boundary irrespective of changes in the contours of the shoreline and regardless of whether the changes are the product of the processes of erosion and accretion or the result of avulsion” (Kalo 2005;1442). The purpose of s35 of the Crown Grants Act is to clarify the definition of the littoral boundary and should logically to be read to mean that the boundary remains the mean high water mark, no matter how it has moved in the past.

However, case law, particularly *AG v Findlay*,⁹ relating to a claim for land accreting to a natural seaward boundary, has continued to apply the common law tests of the slow, gradual and imperceptible accretion in a claim for title to accreted land adjoining a tidal boundary. Kearns *et al* (1997) cite Coulson & Forbes as authority for the law as applied in New Zealand: “Where the sea, or an arm of the sea, by gradual and imperceptible progress encroaches on the land of a subject, the land thereby covered belongs to the Crown¹⁰ or the Crown’s grantee of the foreshore; but where land is suddenly overflowed, and any marks remain by which its limit can be recognised, it remains to the original owner, and may be regained by art or industry; or if the sea retires again it is his as before” (cited in Kearns *et al*. 1997;34-35). I respectfully suggest that in applying a correct interpretation of the Crown Grants Act, but also looking at current values in land and assertions of public rights over the foreshore, this should not be the case. It is very unlikely that a littoral owner, having lost land to the sea, could have any success in asserting proprietary rights over that sea space.

The individual rights attaching to property ownership are not absolute and must always be subject to legitimate planning rules that now, more than ever, focus on environmental protection and public rights. The Resource Management Act 1991 (RMA) has as its purpose “to promote the sustainable management of natural and physical resources” (s5). This implies “the ongoing ability of communities and people to respond and adapt to change in a way that avoids or limits adverse consequences” (MfE 2008;43). The Act establishes a regime that devolves authority to regional councils to manage, *inter alia*, the coastal environment, and implements a consent process for activities and land use. It also requires the preparation of national and regional coastal policy statements to guide landuse decisions.

The RMA defines matters of national importance (s6) which include: “the preservation of the natural character of the coastal environment ... and protection from inappropriate subdivision, use, and development” and “the maintenance and enhancement of public access to and along the coastal marine area.” Also authorities must have particular regard to “the effects of climate change” (s7).

In relation to coastal development, the activities specifically prohibited include reclamation of any foreshore, or any erection of any structure on any foreshore or seabed¹¹ unless expressly

⁹ *Attorney General v Findlay* (1919) NZLR 513 GLR 207

¹⁰ Implicit in this statement is the assumption that the Crown is the owner of the foreshore and seabed. In New Zealand this is an assumption that has been questioned in the recent Court of Appeal ruling of *Ngati Apa v AG* [2003] 3 NZLR 643, and that is still being negotiated by legislative changes of the Foreshore and Seabed Act 2004 in 2010.

¹¹ RMA part III Section 12. Section 9 covers similar conditions in regards to use of any other land.

allowed by a regional plan or a resource consent. In other words the regional council is charged with complete powers about what can be placed on the land – in this case, what structures can be placed to defend the land against the invasion of the sea.

The administration of the coastal environment, while required to be consistent with national policy, is delegated to local authorities. Regional authorities have jurisdiction covering land, air, water and over the coastal marine area which includes the foreshore (from Mean High Water Springs to mean low water springs) and the seabed out to the 12 nautical mile mark. The territorial authorities have jurisdiction over land use upland from MHWS. Furthermore, land below MHWS is now generally assumed to be owned by the Crown,¹² while upland is owned privately or by the local authority. This has potential to cause disputes about, for example, the ownership of shoreline protection structures relative to which side of MHWS¹³ they are located.¹⁴ This fragmentation of jurisdiction and ownership is contrary to the goals of integrated coastal management and must be solved.¹⁵

4 Coastal Policy

The proposed New Zealand Coastal Policy Statement 2008¹⁶ was prepared “to strike the right balance between use, development and protection of our precious coastal environment” (DoC. 2008;5). It is based on the implementation of sustainability goals, including applying a precautionary approach.¹⁷ The policy statement supports the preservation of the natural character of the coastal environment including natural processes, and allows for the management of subdivision and use, while enhancing the public utility of the coast as accessible and public open space (DoC 2008;13-14).

“Coastal Hazard risks are managed increasingly by locating or relocating development away from risk areas, protecting or restoring natural defences and discouraging recourse to hard protection structures” (DoC 2008;7 Objective 8). The policies addressing Coastal Hazards (DoC 2008;25-26) are more specific. Coastal hazard risk assessment must have regard to “the potential for natural coastal features and areas of coastal hazard risk to migrate as a result of dynamic coastal processes including sea level rise” and to “the effects of climate change ... on storm frequency, intensity and surges, and coastal sediment dynamics” (DoC 2008;25 Policy 51(d) and (e)). Local authorities shall avoid new subdivision and development on land at risk from coastal

¹² This is asserted in the controversial Foreshore and Seabed Act 2004, but this Act is about to be repealed by a new Act that will vest the foreshore and seabed as public domain rather than Crown owned.

¹³ A position not easily identified on the ground.

¹⁴ A problem that Brookes (2001;10) describes as jurisdictional apartheid.

¹⁵ In fact, in the case of Gisborne and Wainui Beach discussed herein below, the territorial authority; the Gisborne District Council, is actually, and rather unusually, a unitary authority and holds the powers of a regional council, so that conflict is alleviated.

¹⁶ The current operative Coastal Policy Statement 1994 (DoC 1994) is long overdue a review.

¹⁷ “A precautionary approach shall be adopted towards proposed activities whose effects on the coastal environment are uncertain, unknown or little understood, but whose effects are potentially significantly adverse to that environment” (DoC 2008;10). It is apparent that hard coastal defense structures must fall into this category of activity.

hazards and encourage managed retreat, by relocation, removal or abandonment of existing structures (DoC 2008;25 Policy 52).

Pursuant to this, the Ministry for the Environment is active in planning for climate change effects on coastal margins. A report to the Ministry (Bell et al. 2001;viii) warns that “coastal development and global warming are on an eventual collision course” and that “managed retreat and adaptation are the only reasonable long-term options.” Regional councils have also prepared documentation (Turbott 2008) to assist the planning for coastal erosion and particularly the issues arising from implementing managed retreat – that is, abandoning property or relocating structures threatened by coastal processes. But in terms of putting any effective limits on coastal development, the proposed New Zealand Coastal Policy Statement is far from helpful.¹⁸

The point of the foregoing discussion is that owners of coastal property must be aware that while the nature of their natural ambulatory boundary allows them to gain by the seaward movement of the mean high water boundary, it also requires them to relinquish any claims to land lost by a landward movement. The question about what action such owners may take to defend their land from any loss to the sea is at issue here.

5 Case Law

Not surprisingly, given the significant financial stake they hold in their land, coastal property owners usually expect either to be able to defend their land themselves, or perhaps ideally, demand the government defend it for them. The reference above, that eroded land “may be regained by art or industry,” supports the claim that there is a common law right for property owners to not just protect their land but also to reclaim it from the sea. It is doubtful whether such a right would have ever succeeded in a New Zealand court but certainly would not now.¹⁹

The right to protect ones land from the sea was argued in the case *Falkner v Gisborne District Council* 1995.²⁰ Since the 1920s the near foreshore of Wainui Beach has been occupied by housing. The local authority undertook various beach protection works, including sheet pile groynes, rock walls, and post and rail walls, in an attempt to retain and support the beach fore-dunes as a line of defence for the inland properties and houses. Consequent upon occasional severe storm damage to the defensive works, a high level of pressure was exerted on the local authority to continue to enhance these defences to protect private property. By the 1990s the Department of Conservation and the Gisborne District Council reassessed the feasibility of them providing this engineered solution to holding back the sea, and proposed to allow the sea naturally to take its toll – to implement a policy of managed retreat.

¹⁸ Peart (2009;237) describes the document as “a damp squib.”

¹⁹ under the regulatory regime of the RMA

²⁰ *Falkner v Gisborne District Council* [1995] NZRMA 462

Adjoining landowners expected that the protection work was a continuing local government responsibility. The expectations that property should not be lost without just compensation, led residents to object to the new policy by lodging a claim to the Planning Tribunal and subsequently an appeal to the High Court.

Evidence was presented about the likely increased threat of coastal erosion due to sea level rise and also about the desirability of maintaining a fight against the forces of nature. The question of whose responsibility it was to protect property was raised at a philosophical level but also at a practical level regarding ownership of land where coastal protection works were built, and the location of MHWS.²¹

The judicial argument focused around 3 main issues, namely the common law duty on the Crown to protect the land from invasion; the common law right of individual owners to protect their own property; and whether the statutory intervention of the Resource Management Act 1991 has abrogated or modified those duties and rights (at p466).

In examining the arguments for the rights and duty to protect property, the court quoted Halsbury (at p471): “It is the royal duty of the Crown to preserve the realm from the inroads of the sea by appropriate defences; and every subject has a corresponding right.” However, the court observed that the private right was not absolute: “... no doubt the ordinary rights of property must give way to that which is done for the protection and safety of the public” (at 471).²²

The court, in response to the claim that every landowner has a right to protect his or her land from the inroads of the sea, stated that such an approach “manifests a narrow 19th century preoccupation with proprietary rights, out of keeping with the more holistic policy concerns of sustainability and environmentalism popular today” (at p475). The court decided that the policy of managed retreat was a justifiable response to coastal erosion and that this process was not an evil that takes away property but an expected result of natural processes.

In essence, therefore, the court did not deny that there may have been those rights and duty to protect property, but that the statutory intervention of the RMA and its purpose of the sustainable management of natural and physical resources, establishes a regulatory regime that overrides the common law rules of protection. In other words, the local authorities may refuse resource consents for protective structures that are contrary to their vision of the natural character of coastlines, and that effectively destroy the public amenity values in favour of protecting private property. Property purchasers and owners must be well warned that they carry all liability against loss.

²¹ Under the Resource Management Act 1991 the coastal marine area was defined by Mean High Water Springs. The fact that the property boundaries were defined as Mean High Water Mark appeared to be ignored in these arguments. Expert evidence of surveyors was required to define the actual location of the MHWS line on the ground, and it was not a straightforward decision. It is no surprise, however, that the tide regularly reached the foot of the hard defensive structure; as it always will on an eroding beach.

²² quoting Coulson & Forbes.

6 Risk, Responsibility, Property Value and Compensation

The assurances and protections that attach to a fee simple title in New Zealand suggest that title to land is forever, but this “ignores the reality that the underlying land may itself be impermanent” (Turbott 2006;44). By common law, the ownership of property with a riparian or coastal boundary allows for a gain of land by accretion, matched by the loss of land by erosion: the possibility of a benefit is matched by the threat of a corresponding loss. This principle has been lost sight of often by those owners suffering from drastic loss of property, but local authorities and the courts could more forcefully reassert that the loss is a natural consequence of coastal property ownership.

The responsibility for property loss is primarily the concern of the affected property owner (ARC 2003;20), as are any effects of structures built to protect one property but have nuisance and aggravating effects on neighbouring properties.

Evidence suggests that the market value of coastal property is largely unaffected by the risk of coastal erosion (Turbott 2006;12).²³ Similarly, American research (cited in Dahm 2002;24) has found that coastal property owners find “the risk [of coastal erosion] acceptable because they ‘want to be there’, the amenities of an oceanfront location (e.g. view, easy access to the beach, water recreation, peace and quiet) appearing to meet deeply felt emotional needs of the people who owned property there.”

Economic arguments could also be raised to support the loss of property. Coastal proprietors have almost invariably gained by the disproportionate increase in property values as a result of continuing demands for coastal land. The normal risks of investment in property, such as the windfall profits reaped by taking advantage of social demands, should be balanced by exposure to the risks of loss to the erosive forces of the sea. It should not be for the wider public, through taxes or property rates, to compensate for such property loss.

In the past, local authorities have built coastal defences to protect their communities, and the political and economic power deriving from the value of coastal property has imposed a strong imperative to maintain such structures. Furthermore, the more recent style of coastal development²⁴ suggests that “we have ‘given up’ on preservation and protection of natural character of much of our coast” (Brookes 2001;8).²⁵

Property insurance is the normally expected fallback position for property owners affected by loss. Property damage from flooding is insurable, but usually such property can be reoccupied.

²³ Even when hazard warning notices are entered on the title document, values continue to increase (Turbott 2006;22)

²⁴ Recent development typically involves high value holiday homes dominating coastal landscapes, built to maximize view and proximity to the beach, and several examples of the creation of artificial waterways and canal-centred development on totally manmade shorelines.

²⁵ Brookes continues (2001;9) “A cohesive and informed community concerned about the health of their local coastal system becomes replaced (if it ever existed) by individuals concerned with the protection of private property and the effective removal of now hazardous coastal processes from their proximity.”

Eroded land is gone for good, there is a total loss. Given the inevitability and expected worsening of the coastal threat, it would appear reasonable to expect insurance companies to withdraw or deny cover for losses incurred by coastal erosion, especially by the gradual process of sea level rise. Government compensation schemes (such as the Earthquake Commission) have often covered property loss due to natural disaster, but this is only available for severe unpredictable events. While the government has an interest in supporting the property market, it should not take responsibility for the gradual effects of coastal erosion from which all property owners have due notice and time to retreat. The possibility of Crown or local authority buy-out is a useful option, but the expense of compensating for unrealistically inflated coastal property values is a significant impediment. There are a wide variety of financial assistance packages that could be made available, including full compensation for loss, subsidies for relocation, and buyout and lease back arrangements. But ultimately any such financial incentives will artificially prop up the property value, providing individual benefit at public expense, and continue the resistance to voluntary retreat.²⁶

More appropriate legal interventions may include covenants on titles identifying a coastal hazard (including no-complaints covenants preventing claims being made against neighbours, or local authorities, or even the acts of nature), time limited landuse consents, and short term occupation licences. Ambulatory public access or conservation strips, as are allowed for under the Conservation Law Reform Act 1990²⁷ in relation to public strips abutting water, and as have been described in Titus (1998;1313) as rolling easements, could provide appropriate notice to coastal property owners of the vulnerability of coastal titles. It could also be possible to convert existing use rights on the sale of any at risk property, to short term occupation licences.²⁸ Such takings or confiscations²⁹ of otherwise unencumbered title would, however, inevitably initiate compensation applications that would be hard for government to defend.³⁰

Probably the best coastal management strategies of all, although mostly limited to future land use planning on undeveloped land, is to establish regulations requiring mandatory setback distances, setting aside of wide public land margins, requiring relocatable housing, and only allowing the subdivision of land into short timeframe leasehold titles. In addition, public infrastructure, like roads should be designed perpendicular to the coast to avoid future erosion threats to the fixed and vital hard structure.³¹

²⁶ See Defra (2009;7) "Government intervention should not undermine the policy messages in relation to coastal erosion, that: coastal erosion is a natural process and cannot be stopped, and individuals investment decisions should be made reflecting the risks of coastal change and should not rely upon the provision of defences or financial assistance."

²⁷ Part IVa - although only in relation to the disposal of Crown land

²⁸ By placing a memorial such as a caveat, on affected Certificates of Title.

²⁹ See also Stallworthy's arguments about environmental justice and the difficulty of determining how much restriction on land amounts to confiscation and may trigger compensation (2006;367).

³⁰ It is worth emphasizing Turbott's (2006;2) assertion that "Fee simple property tenure is probably the largest barrier to implementation of managed retreat. This is because it involves an expectation of permanent use of the titles. Unfortunately, the underlying physical land may not be permanent. This problem is compounded by the high and increasing value of coastal property."

³¹ Some design possibilities are illustrated in Turbott (2006)

7 Coastal Protection Works

National policy prioritises natural character and processes, legal decisions dismiss the absolute rights of property owners; technical reports confirm the inevitability of sea level rise and erosion; resource consent conditions and encumbrances on titles may warn of land loss; but actual practice illustrates the power base that coastal property owners can exert on local authorities. Brookes (2001) quotes a Ministry of Works report that identified the difficulty of getting political leverage to support integrated coastal management “given the present financial structure and electoral base of most coastal counties.” This may suggest that local coastal planning should be removed from local councils and passed to “some organisation which is less dependent on local opinion and local financial support” (Brookes 2001;7).

In spite of the success the Gisborne District Council had in front of the Court in the *Falkner* case, the council has prepared the Wainui Beach Management Strategy (WBMS) (GDC 2003) to provide for the continued protection of coastal property by the erection of defensive structures to maintain protection of private property – relenting to the demands of local property owners.

A Resource Consent application (GDC 2009) has been lodged by the local council for the construction of a sloping rock revetment to protect the dunes and adjacent coastal property. The consent application implements aspects of the WBMS and asserts that the works will be an improvement on previous structures that have become ineffective and have become a hazard. The application also asserts that the beach is already a highly modified environment and the beach protection projects now ‘form part of the natural character’ (GDC 2009;9) of the beach. It seems an effective policy of managed retreat is too hard, too expensive and too politically unacceptable to implement.

Hard structures such as rail irons, posts and logs, have been exposed as ultimately succumbing to the forces of wind, tide, waves and sea level rise, and they disturb natural processes and ecosystems. Inevitably the beach is removed by erosion up to the base of the hard structure, and storms and the tide will continue to undermine hard defences. Moreover, the structures are unsightly on the coastal landscape, create dangerous obstacles and are impediments to public access. But that is not to say that nothing can be done. Research is being undertaken on geotextiles to hold sand and beach sediment³² as an alternative to the hard structures used in the past, and there are many examples of restoration of dune vegetation successfully protecting beach frontages.³³

³² Beach replenishment has only been applied on a small scale in New Zealand.

³³ See Peart (2009).

8 Conclusion

The coastal margins of our land mark a significant zone of conflict, where not just the land and the sea clash, but environmental, social, economic, legal and engineering issues also are at odds. The extremely strong preference for coastal property exhibited by the population overrides the hazard risk evident by even casual observation. Furthermore, there is a strong expectation that land and property rights should be permanent and protected from loss. The judicial, legislative and policy suggestions to the contrary will require careful negotiation. But just as the physical defences against the power of the sea are failing, inevitably, so too will people and property give way to the sea. We cannot afford, economically or environmentally, to hold back the sea.

9 References

- ARC. Auckland Regional Council. (2003). *Coastal Erosion Management Manual*. Auckland Regional Council.
- Bell, R. G., Hume, T. M. and Hicks, D. M. (2001). *Planning for Climate Change Effects on Coastal Margins*. ME410. Ministry for the Environment, Wellington.
<http://www.mfe.govt.nz/publications/climate/effect-coastal-sep01/> viewed: 20/11/09.
- Bennion F. A. R. (1992). *Statutory Interpretation* 2nd Ed. Butterworths.
- Brookes, H. (c.2001). Chances Beckoning: Integration in New Zealand's Coastal Management. www.rmla.org.nz/publications/h_brookes.doc viewed: 20/11/09.
- Dahm, C. (2002). *Beach User Values and Perceptions of Coastal Erosion*. Environment Waikato. Technical Report 2003/03. 68. <http://www.ew.govt.nz/Publications/Technical-Reports/Beach-user-values-and-perceptions-of-coastal-erosion/> viewed: 20/11/09
- Defra. Department for Environment Food and Rural Affairs. (2009). Coastal erosion assistance package Impact Assessment. <http://www.defra.gov.uk/corporate/consult/coastal-change/> viewed: 20/11/09.
- DoC. Department of Conservation. (1994). *New Zealand Coastal Policy Statement 1994*. Department of Conservation, Wellington.
- DoC. Department of Conservation. (2008). *Proposed New Zealand Coastal Policy Statement 2008*. Policy Group, Department of Conservation: 35, Wellington. www.doc.govt.nz viewed: 26/06/10
- GDC. Gisborne District Council. (2003). *Wainui Beach Management Strategy*. Gisborne: 49 plus appendices. <http://www.gdc.govt.nz/wainui-beach-management-strategy/> viewed: 20/11/09.

- GDC. Gisborne District Council. (2009). Application by GDC Rivers & Land Drainage Department. 2009. For Coastal Permit and Land Use Consent. <http://www.gisborne.govt.nz/assets/Files/Planning-Policy/Wainui-Beach-works/Stage3-ResourceConsentApplicationAEE.pdf> viewed: 20/11/09.
- Kalo, J. J. (2005). North Carolina Oceanfront Property and Public Waters and Beaches: The Rights of Littoral Owners in the Twenty-First Century. *North Carolina Law Review* [2005] Vol 83: 1427-1505.
- Kearns, F. J., Kerr, N. T. and Smith, M. C. (1997). *Boundaries and Boundary Definition*. Department of Surveying and the New Zealand Institute of Surveyors, Dunedin.
- MfE. Ministry for the Environment. (2008). *Coastal Hazards and Climate Change. A Guidance Manual for Local Government in New Zealand. 2nd Edition. Revised by Ramsey, D, and Bell, R. (NIWA)*. ME892. Ministry for the Environment, Wellington.
- Peart, R. (2009). *Castles in the Sand. What's happening to the New Zealand Coast?* Craig Potton Publishing. Nelson.
- Stallworthy, M. (2006). Sustainability, Coastal erosion and Climate Change: An Environmental Justice Analysis. *Journal of Environmental Law* (2006) Vol 18 No 3, 357-373
- Titus, J. G. (1998). Rising Seas, Coastal Erosion, and the Takings Clause: How to save wetlands and beaches without hurting property owners. *Maryland Law Review* [1998] Vol 57 No 4, 1279-1399.
- Turbott, C. (2006). *Managed Retreat from Coastal Hazards: Options for Implementation*. Environment Waikato. Technical Report 2006/48: 89. <http://www.ew.govt.nz/publications/Technical-Reports/Managed-Retreat-from-Coastal-Hazards-Options-for-Implementation/> viewed: 20/11/09.

Adverse Possession of Boundary Land – Lessons from Abroad

Una Woods¹

¹ School of Law, University of Limerick, Limerick, Ireland

Email: una.woods@ul.ie

Abstract:

Boundary disputes are an unfortunate fact of life. This paper illustrates the difficulties currently faced in Ireland by those seeking to demonstrate adverse possession of boundary land and discusses alternative remedies which may be relied upon to resolve boundary disputes. Recently, concerns have been expressed in Ireland and elsewhere in relation to the fairness of the doctrine, in particular the way it enables the deliberate squatter who is fully aware that he is not the owner of the land to acquire title to it. In England and Wales, the Land Registration Act 2002 introduced substantive reforms to the law on adverse possession of registered land which restrict its operation in the context of boundary land to adverse possessors who can satisfy a good faith requirement. In 2005 the Irish Law Reform Commission proposed certain reforms to the law on adverse possession, including the imposition of a good faith requirement in relation to boundary land in language which mirrors that used by the English 2002 Act. This paper discusses the interpretative difficulties which the English good faith requirement may give rise to and reveals a vacuum which this new approach creates in the resolution of boundary disputes that may leave certain ‘deserving’ possessors of boundary land without a remedy. The Irish Law Reform Commission plans to revisit the matter of reform in this area of law in the near future. If a good faith requirement is imposed in relation to adverse possession of boundary land, certain lessons could be learned from the English experience and also the Australian approach to boundary disputes, which relies heavily on mistaken improver and building encroachment legislation.

Keywords:

Boundary dispute; adverse possession; Ireland

1 Introduction

1.1 Boundary Litigation and Alternative Dispute Resolution

Boundary disputes between neighbours are “wretched affairs”² and the financial cost of boundary litigation frequently exceeds the value of the land concerned. In some cases the area of land in dispute will be very small but litigation will ensue because of its strategic importance for the development plans of one or both contiguous owners.³ In other cases the plot may be worth little in terms of commercial value or even in terms of amenity but

¹ School of Law, University of Limerick, Limerick, Ireland

Email: una.woods@ul.ie

² *Childs v Vernon* [2007] EWCA Civ 305 per Toulson LJ at para.1.

³ See *Persian Properties Ltd v Registrar of Titles and Minister for Finance* [2003] 1 IR 450.

the boundary dispute becomes a proxy for another dispute⁴ or an underlying clash of personalities between neighbours. Occasionally, even a court ruling will not resolve matters and the same neighbours end up back in court years later in connection with a similar or related matter.⁵ The Irish Law Reform Commission has recently recommended that parties to such disputes should be advised by their legal representatives to consider mediation or conciliation prior to commencing litigation, as is the case with divorce proceedings (2008, para 9.17). The Commission also condoned the approach taken by a High Court Judge hearing a recent court case, *Charlton v Kenny*,⁶ involving a boundary dispute between a well-known solicitor and a famous chat show host. Clark J urged the parties to consider resolving the matter through mediation. He noted: “You both live in very attractive houses in a very idyllic setting and you have to go back and live there. It won’t be idyllic when the case is over” (Black and Healy, 2008). Following this intervention the parties adjourned the case to reach a mediated settlement. Alternative Dispute Resolution will not work for all such disputes but where appropriate, it will operate more efficiently and effectively if the legal principles and remedies governing the area are unambiguous, familiar to legal advisors and coincide with popular notions of social justice.

1.2 Boundary Identification

The Irish Institution of Surveyors maintains that the incidence of boundary disputes in Ireland may be significantly higher than the international norm (2008, para. 1.6). A contributory factor may be the difficulty in determining the legal boundary of a property. If a boundary dispute reaches court, the court will first examine the description of the property contained in the parcels clause of the deeds.⁷ If the description of the land is unclear, the boundary may be determined by relying on certain rebuttable presumptions which arise in relation to adjoining hedges and ditches, roads, rivers and party walls (Woods and Wylie, 2005, para. 18.58-18.62). In certain circumstances extrinsic evidence may be admitted to clarify any ambiguity⁸ or inferences may be drawn from topographical features which existed or probably existed when the conveyance was executed.⁹ Frequently, deeds incorporate a map outlining the property. The scale of older maps is typically so small and the lines marking the boundaries are so thick as to be useless for any purpose except general identification.¹⁰ In the event of an inconsistency between the boundaries as identified by the verbal description and by the map, the words used to incorporate the map will be examined by the court to assess whether the map was

⁴ See *Boyle v Connaughton* unreported, High Court, 21 March 2000).

⁵ See *Fahy v Dillon* unreported, Circuit Court, 29 July 2005.

⁶ 2006 No 4266 P High Court, 8 to 11 April 2008 (hearing of action) and 15 April 2008 (settlement after mediation).

⁷ See *McCoy v McGill* [2008] IEHC 301, para. 7-8, endorsing the approach taken by Lord Hoffman in the English case, *Wibberley Building Ltd v Insley* [1999] 1 WLR 894 at pp 895-896.

⁸ For example, if the property was described as being ‘formerly in the occupation of’ a named individual or family or ‘known as...’, see *Freeguard v Rogers* [1999] 1 WLR 357. If the adjoining owners entered into an agreement in relation to the location of the boundary, the agreement will be treated as identifying on the ground what the documents describe in words or delineate on plans, see *Neilson v Poole* (1969) 20 P & CR 909; *Burns v Morton* [1999] 3 All ER 646.

⁹ See *McCoy v McGill* [2008] IEHC 301, para. 9, where the judge referred to a cliff face and a pier which would have existed when the deeds were executed and were identified in the plans attached thereto.

¹⁰ See *McCoy v McGill* [2008] IEHC 301, para. 7-8 endorsing the approach taken by Lord Hoffman in the English case, *Wibberley Building Ltd v Insley* [1999] 1 WLR 894 at pp 895-896.

intended to be determinative of the precise boundaries or not (Woods and Wylie, 2005, para. 18.66-18.68).

Surprisingly, boundary identification is no easier if the title to the land is registered in the Land Registry as the conclusive nature of the register¹¹ does not extend to boundaries. Section 85 of the Registration of Title Act 1964 specifies that the outline of the property as it appears on the Land Registry map shall not be conclusive as to the precise line of the boundary or the extent of the land unless the boundaries have been entered as conclusive pursuant to the procedure provided by the 1964 Act.¹² This facility is rarely relied upon (Fitzgerald, 1995). In response to those who wonder why the legislature did not contemplate making boundaries conclusive when the system of title registration was initially introduced, McAllister(1973) comments as follows:

The answer would appear to be that precise boundaries are generally unknown in unregistered conveyancing; and the difficulty and expense involved in making an exact survey for the purpose of registration would make all registration impossible.

Lessons were presumably learned from the first English experiment with the introduction of registered title system in 1862 which was designed to record exact boundaries. This approach was subsequently acknowledged to be a mistake by the Land Transfer Commission on the Operation of the Land Registry Act 1870 - people who were quite content with an undefined boundary were compelled to have it defined, which required notices to be served on their neighbours who immediately began to consider whether some injury was about to be inflicted on them “and a dispute is thus forced upon neighbours who only desire to remain at peace.”¹³

Sperling has recently suggested that a move towards conclusive boundaries is inevitable and necessary (2008). The Irish Institution of Surveyors’ Commission on Land Registration has recommended the introduction of a requirement to register all new boundaries as conclusive, although the conversion of existing boundaries to conclusive would continue to be voluntary (2008, p47). It envisages the gradual conversion of the Registry database from non-conclusive to conclusive boundaries over the next 100 years (2008, p51). Currently, the primary statutory mandate of the Property Registration Authority, which manages the Land Registry, is to extend the registered title system to the entire country.¹⁴ It seems unlikely that an additional statutory obligation to enter new boundaries as conclusive will be introduced in the near future given the additional costs that this would entail. The benefits of such a scheme would have to be carefully ascertained before such a step could be taken. It is also difficult to envisage existing neighbours opting to voluntarily convert their boundaries to conclusive, given the potential for dispute which this could give rise to.

¹¹ S31(1) of the Registration of Title Act 1964 provides that the register shall be conclusive evidence of the title of the owner to the land as appearing on the register and of any right, privilege, appurtenance or burden as appearing thereon.

¹² See s87-88 of the 1964 Act.

¹³ Quoted by Lord Hoffman in *Alan Wibberley Building Ltd v Insley* [1999] UKHL 15.

¹⁴ S10(1)(b) of the Registration of Deeds and Title Act 2006.

1.3 Outline of Paper

In any event, even if a legal boundary can be satisfactorily determined, the potential for dispute is not eliminated as a dissatisfied neighbour may avail of a number of legal and equitable remedies. This paper illustrates the difficulties currently faced in Ireland by those seeking to rely on the doctrine of adverse possession to resolve a boundary dispute and discusses alternative remedies which may be relied on to settle such disputes. Recently, concerns have been expressed in Ireland and elsewhere in relation to the fairness of the doctrine, in particular the way it enables the deliberate squatter who is fully aware that he is not the owner of the land to acquire title to it (Law Reform Commission, 2005, p.8). In England and Wales, the Land Registration Act 2002 introduced substantive reforms to the law on adverse possession of registered land, restricting its operation in the context of boundary disputes to adverse possessors who can satisfy a good faith requirement. In 2005 the Irish Law Reform Commission proposed certain reforms to the law on adverse possession, including the imposition of a good faith requirement in relation to boundary land in language which mirrors that used by the English 2002 Act (2005 p.330). This paper discusses the interpretative difficulties which the English good faith requirement may give rise to and reveals a vacuum which this new approach creates in the resolution of boundary disputes that may leave certain ‘deserving’ possessors of boundary land without a remedy. Following the Grand Chamber decision in *JA Pye (Oxford) Ltd v United Kingdom*¹⁵ the pressure to introduce substantive changes to the Irish law on adverse possession has abated. However, the Law Reform Commission plans to revisit the matter of reform in this area of law in the near future. If a good faith requirement is imposed in relation to adverse possession of boundary land, certain lessons could be learned from the English experience and also the Australian approach to boundary disputes, which relies heavily on mistaken improver and building encroachment legislation.

2 Adverse possession of boundary land in Ireland

A review of reported Irish caselaw reveals that the vast majority of adverse possession claims which reach court are between contiguous landowners. The remaining caselaw typically involves a pre-existing family or contractual relationship between the parties.¹⁶ An adverse possession claim involving neighbours is sometimes preceded by a paper title dispute,¹⁷ which the court resolves by relying on the principles outlined earlier.

¹⁵ (2008) 46 EHRR 45. The Grand Chamber ruled that the English law on adverse possession, as it applied to registered land before the enactment of the Land Registration Act 2002, did not violate article 1 of the First Protocol to the European Convention on Human Rights, which protects property rights.

¹⁶ See *Gleeson v Feehan* [1993] 2 IR 113; *Bellew v Bellew* [1982] IR 447; *Moley v Fee* [2007] IEHC 143; *Griffin v Bleithin* [1999] 2 ILRM 182.

¹⁷ See *Tracey Enterprises MacAdam Ltd v Drury* [2006] IEHC 381; *Dunne v CIE* [2007] IEHC 314.

2.1 Proving Adverse Possession

To demonstrate adverse possession, a neighbouring landowner must prove that his possession was not consensual and that his acts of possession were sufficiently strong, exclusive and coupled with the requisite intention. In *Tracey Enterprises Macadam Ltd v Drury*¹⁸ the defendant's use of the disputed plot for repairing machinery was described by the court as non-exclusive, sporadic and unaccompanied by the necessary *animus possidendi*. Laffoy J attached significant weight to a letter sent by the defendant to the plaintiff which intimated that he only used the plot occasionally and that it was also used by others. If the owner continues to engage in acts of possession in relation to the land, an adverse possession claim will not succeed, regardless of whether the claimant is also engaged in acts of user. In *Browne v Fahy*¹⁹ the defendants' adverse possession claim failed because the court was satisfied that the plaintiffs had occasionally walked over the land which the defendants had grazed cattle on for over 30 years.²⁰ More recently, in *Mahon v O'Reilly*²¹ the owner failed to convince the court that he had continued to engage in acts of possession in relation to his yard during the limitation period and the plaintiff, his neighbour, succeeded in proving adverse possession. In *Doyle v O'Neill*²² the defendant succeeded in demonstrating that he had taken adverse possession of a small triangular section of ground he needed to widen the entrance to his dwellinghouse. He had put a hard surface down on this strip and separated it from the rest of the disputed plot by a boundary wall. However, the court was satisfied that the rest of the disputed plot was only used for occasional and temporary storage and dumping. The following comment by O'Hanlon J has since been referred to in numerous decisions:

In order to defeat the title of the original landowner... the adverse user must be of a definite and positive character and such as could leave no doubt in the mind of a landowner alert to his rights that occupation adverse to his title was taking place. This is particularly the case when the parcel of land involved is for the time being worthless or valueless for the purposes of the original owner.²³

O'Hanlon J also based his decision on a number of incidents which demonstrated an absence of *animus possidendi* on the part of the defendant. The defendant had, for example, sought permission from the plaintiff to store steel on the plot on a temporary basis²⁴ and requesting the plaintiff to stop cutting a hedge along the boundary of the plot rather than commanding him to do so. In *Kelleher v Botany Weaving Mills Ltd*²⁵ Clark J ruled that the defendant, the owner of a neighbouring factory, had failed to establish adverse possession of a plot of land at the back of a garage owned by the plaintiff. Although the defendant had cleared up the area and sown it in grass, Clark J felt that

¹⁸ [2006] IEHC 381.

¹⁹ Unreported, High Court, 24 October 1975.

²⁰ See also *Mulhern v Brady* [2001] IEHC 23; *Feehan v Leamy* [2000] IEHC 118.

²¹ [2010] IEHC 103.

²² Unreported, High Court, 13 January, 1995.

²³ Unreported, High Court, 13 January, 1995 at p 20 of transcript. See also *Fanning v Jenkinson* unreported, High Court, 2 July 1997.

²⁴ See also *Keelgrove Properties Ltd v Shelbourne Development Ltd* [2005] IEHC 238.

²⁵ [2008] IEHC 417.

these were not actions which would leave no doubt in the mind of the plaintiff that occupation adverse to his title was taking place. She also noted that plaintiff had used the plot to access the garage roof to carry out repairs on a number of occasions. Although she was clearly entitled to reject the adverse possession claim on the basis that the acts of possession by the defendant were insufficient and the plaintiff was still engaged in acts of possession, she put forward a third reason for her decision: the director of the defendant company believed he owned the plot. According to Clark J *animus possidendi* could not co-exist with a belief in ownership: 'Logically, if a person in occupation believes he is rightfully there, he cannot have the intent to dispossess.'²⁶

2.2 Critique of *Kelleher*

It is respectfully submitted that the ruling in *Kelleher* on *animus possidendi* is unsound and unlikely to be followed for a number of reasons. First of all, it fails to take into account the Supreme Court decision in *Murphy v Murphy*²⁷ which clearly establishes that an adverse possessor who believes he is the owner can satisfy the requirement for *animus possidendi*. This case involved a farm which the testator's widow and his two sons believed had been split between the two sons under the testator's will. In actual fact a portion of the land had not been dealt with and therefore passed to the widow under the residuary clause. The eldest son bought out his brother and farmed the entire farm as if it was his own throughout the limitation period. Kenny J was convinced that the widow was unaware that she had become the owner of the disputed land and that the eldest son had assumed it belonged to him. He continued:

The ignorance of the testator's widow and of the defendant that the lands north of the road belonged to her does not prevent the Act of 1957 applying and does not prevent the defendant being in adverse possession of these lands. In Wylie's *Irish Land Law* at p. 857 it is stated: "It is also established that the adverse possession may take place without either party being aware of it."²⁸

Instead, Clark J, in her decision in *Kelleher*, referred to the paucity of relevant Irish law discussing whether a person can be in adverse possession while believing that he is the owner. She noted, however, that *Halsbury's Laws of England* cite *Williams v Usherwood*²⁹ as an authority that adverse possession can occur even if the possessor believed he was the legal owner. In that case two neighbouring landowners had mistakenly assumed that the defendant was the owner of a driveway. The court noted that the enclosure of the driveway by a fence and later by a wall was strong evidence of *animus possidendi* and the fact that it was based on a mistaken premise did not aid the paper owner. Clark J proceeded to distinguish the facts of *Williams v Usherwood* on the basis that the adverse possessor in that case was involved in much stronger acts of possession. She felt that the issue of the mistaken belief paled into insignificance when viewed against these acts and the abandonment by the owner of any use of the driveway.

²⁶[2008] IEHC 417 at para. 22.

²⁷[1980] IR 183.

²⁸[1980] IR 183 at 202.

²⁹(1983) 45 P & CR 235.

For this reason, she maintained that the case could not be viewed as any authority for the principle that *animus possidendi* can co-exist with a belief in ownership. It is difficult to follow this reasoning. Either a belief in ownership is fatal to *animus possidendi* or it is not; it is not logical to say that it can be ignored if the squatter is engaged in very strong acts of possession. Both a physical and a mental element must be established to succeed in an adverse possession claim and the presence of one without the other should be fatal to a claim, regardless of whether that one element is very easily established.

It is unfortunate that the court's attention was not drawn to the other English authorities on this point which Jordan refers to in his book on adverse possession (2002). Jordan is unequivocal in his view that an erroneous belief by the adverse possessor that he owns the property constitutes the necessary *animus* (para 9.13). He refers to *Hughes v Cork*³⁰ where the judge at first instance held that the squatters had not acquired title by adverse possession because they believed they owned the land and thus could not have an intention to exclude or oust the true owner. However, the Court of Appeal reversed that decision and noted as follows:

If this were the law then only those who knew they were trespassing... could acquire such a title, while those who did not realise that they were doing anything wrong would acquire no rights at all. I can see no reason why, as a matter of justice or common sense, the former but not the latter should be able to acquire title in this way.... Obviously, if the possessor knows or believes someone else has paper title to the land he must intend to exclude that person along with everyone else. But in the absence of such knowledge or belief it is in my judgment sufficient ... simply to establish a manifest intention to exclude everyone.

In *Newlands (Charmouth) Ltd v Langran*,³¹ as was the situation in *Kelleher*, weak acts of possession in relation to the disputed plot, such as cutting the grass and nettles once a year and occasionally using it for bonfires, were coupled with a belief in ownership. Although the court was satisfied that the neighbour's belief in ownership amounted to proof of an intention to possess, it noted that this was not sufficient to establish adverse possession in the absence of sufficient acts of possession.

It must also be assumed that the court in *Kelleher* was unaware of a recent Northern Irish case, *Donnelly v Doyle*,³² where both parties had mistakenly assumed that the plaintiffs owned a laneway. The plaintiffs had cut the hedges, filled in potholes and concreted the entire length and width of the laneway and the court ruled that they had established adverse possession. The court endorsed the approach taken by the English Court of Appeal in *Hughes v Cork* and the Irish Supreme Court in *Murphy v Murphy* in relation to the impact of a mistaken belief in ownership.³³

³⁰ [1994] EGCS 25.

³¹ Unreported, Court of Appeal, 14 July 1989.

³² [2004] NIQB 66.

³³ [2004] NIQB 66 at para15.

What is most surprising about the judicial ruling in the *Kelleher* decision in relation to the impact of a belief in ownership is that it is completely at odds with reforms to the doctrine imposing a good faith requirement in relation to boundary land recently introduced in England and Wales and proposed by the Irish Law Reform Commission. It is submitted that this development in *Kelleher*, which prevents a good faith adverse possessor from obtaining title by adverse possession, is a retrograde step in terms of aligning the doctrine with popular notions of social justice but, fortunately, one which is unlikely to lead to an established precedent. It is interesting to note that a similar approach at one time held sway in the United States.³⁴ The Maine Rule provided that an adverse possessor who mistakenly believed he was the owner of the boundary plot could not demonstrate the requisite intention unless he could prove that he intended to claim title even if he did not own the plot. However, in more recent times this approach has been rejected by the courts in favour of the Connecticut Rule³⁵ which renders good and bad faith irrelevant and ensures that the honest but mistaken possessor does not receive less favourable treatment than the bad faith possessor.

3 Alternative Remedies for the resolution of boundary disputes in Ireland

3.1 Negligence, Misrepresentation and Breach of Contract

Certain remedies may arise from a transaction in relation to the land. General Condition 14 of the *Law Society of Ireland General Conditions of Sale* (2001) specifies that the vendor shall not be required to define exact boundaries or to specify whether fences, walls, hedges or ditches are of a party nature. However, if a purchaser is buying a site, his solicitor should insist on the inclusion of a special condition to deal with identity and boundaries. In *Doran v Delany*³⁶ the High Court held that the purchaser's solicitor was negligent in failing to ensure that the boundaries were staked out and for agreeing to the deletion of a special condition in the draft contract which would have provided for the furnishing of an Ordinance Survey map showing the site and its boundaries clearly marked. If the boundaries of the property are misrepresented or misdescribed in the contract for sale, the purchaser will be able to sue the vendor for misrepresentation and breach of contract if he suffers a loss as a result.³⁷ In *Doran v Delany*, on appeal to the Supreme Court,³⁸ even the vendor's solicitor was found to be liable to the purchaser of a landlocked site. In response to the standard enquiry raised by the purchaser's solicitor as to whether there was any litigation pending or threatened in relation to the property or any adverse claim made in relation thereto, the vendor's solicitor replied "vendor says no." At the time, the vendor's solicitor had been aware of a boundary dispute with the neighbour. He was found to be negligent, in the particular circumstances of the case, in failing to ascertain the terms on which this dispute had been resolved and conveying that information to the purchaser

³⁴ This approach originated in *Preble v Maine* Cent R 85 Me, 260 27 A 149 (1893).

³⁵ As originally applied in *French v Pearce* 8 Conn. 439 (1831).

³⁶ [1996] 1 ILRM 490.

³⁷ See *Doran v Delany* [1996] 1 ILRM 490.

³⁸ [1998] 2 IR 61.

3.2 Rectification

If the land described in a deed or outlined on the Land Registry map fails to reflect the actual intentions of the parties to the original agreement and one of the parties refuses to execute a deed of rectification, the equitable remedy of rectification may be sought.³⁹ In *Boyle v Connaughton*⁴⁰ the court applied the remedy of rectification to the successors in title of the original parties to the agreement. The original parties had agreed to a sale of a site comprising part of a Land Registry folio. At the time the deed of transfer was executed, the house had already been constructed and Laffoy J was satisfied that the boundaries of the site were properly reflected on the ground, in particular by the walls which were constructed to separate both properties. She held that the map attached to the deed of transfer did not properly depict the boundaries as agreed. According to the Land Registry map both parties were technically encroaching on each other's property at different points along the boundary. Laffoy J noted that the conclusive nature of the register as set out by section 31(1) of the Registration of Title Act 1964 was qualified by section 72 which sets out a number of burdens which affect the land without registration, including the rights of those in actual occupation of the land.⁴¹ She ruled that the plaintiff, when he acquired the property, would have been bound by the defendant's right to have the Land Registry map rectified by virtue of his actual occupation of the relevant portion. This decision has proved slightly controversial in that a right to rectify is generally not considered a proprietary right and tends to be classified as a "mere equity" (Breen, 2000). However, the decision is in line with the approach taken in very similar English case, *Blackrocks v JB Developments (Godalming) Ltd.*⁴²

3.3 Compensation from the State

Where a rectification or refusal to rectify the register arises due to an error originating in the Land Registry it may give rise to a claim for compensation from the State,⁴³ as was the case in *Persian Properties Ltd v Registrar of Title and Minister for Finance*.⁴⁴ This litigation concerned a small strip of land in inner city Dublin representing part of a bottle-neck entrance down the side of a building owned by the plaintiff and at the rear of another building owned by Tomkin Estates Ltd. On the first registration of a neighbouring building which included part of this entrance in the Land Registry, the entire entrance was mistakenly included within the folio. This plot was subsequently acquired by the plaintiff, in reliance on the Land Registry map. The title dispute was concluded by a High Court declaration that Tomkin Estates Ltd was the owner of the strip and the plaintiff succeeded in a claim for compensation pursuant to section 120 of

³⁹ S31(1) of the Registration of Title Act 1964 sets out the conclusive nature of the Land Register but qualifies this statement by setting out that nothing in the Act shall interfere with the general jurisdiction of the court to order the register to be rectified in cases of fraud or mistake. This jurisdiction is supplemented by s32 which enables the Property Registration Authority or the court to order the rectification of the register and its map to correct any error which originated in the Land Registry where this would not result in injustice to either party.

⁴⁰ Unreported, High Court, 21 March 2000.

⁴¹ S72(1)j of the Registration of Title Act 1964.

⁴² [1982] Ch 183.

⁴³ S120 of the Registration of Title Act 1964.

⁴⁴ [2003] 1 IR 450.

the 1964 Act. The Supreme Court ruled that section 85 which states that the register is not conclusive as to boundaries did not relieve the defendants from payment of compensation where the error could be regarded as substantial. Although the error as a percentage of the entire area comprised within the application on first registration was very small, it nonetheless affected two thirds of the entrance to the lane, and therefore could not be regarded as minor.

3.4 Proprietary Estoppel

The doctrine of proprietary estoppel may provide a remedy for a neighbour who can prove that the owner made an inaccurate representation to him in relation to the position of the boundary or acquiesced in his mistaken belief in this regard, provided that the neighbour can also prove that he relied to his detriment on that representation or acquiescence. The mistake limb of the doctrine, as originally set out by Lord Cranworth in *Ramsden v Dyson*,⁴⁵ appears particularly well-placed to remedy the injustice which may arise from the application of the principle *superficies solo cedit* in certain circumstances of boundary error:

If a stranger begins to build on my land supposing it to be his own, and I, perceiving his mistake, abstain from setting him right, and leave him to persevere in his error, a Court of Equity will not allow me afterwards to assert my title to the land on which he had expended money on the supposition that the land was his own.⁴⁶

However, if the neighbour is aware that the land is not his own, no remedy will be available under this doctrine. As Lord Cranworth clarified,

(f)or if a stranger builds on my land knowing it to be mine, there is no principle of equity which would prevent my claiming the land with the benefit of all the expenditure made on it. There would be nothing in my conduct active or passive, making it inequitable in me to assert my legal rights.⁴⁷

Lord Cranworth also pointed out that the real owner must be aware that the land belongs to him at the time of the expenditure.⁴⁸ In an early Irish case, *O'Sullivan v Cork County Council*,⁴⁹ the plaintiffs failed to realize that they owned the land while the council was building a house on it. The house had been built by the time the plaintiffs became aware of their rights. The court refused to accept the council's estoppel claim and the plaintiffs succeeded in reclaiming possession of the land along with the house built on it.

An extremely broad application of estoppel principles took place in *Mac Mahon v Kerry County Council*.⁵⁰ The plaintiffs had purchased a plot from the council with the intention

⁴⁵ (1866) LR 1 HL 129

⁴⁶ (1866) LR 1 HL 129 at 140-141.

⁴⁷ (1866) LR 1 HL 129 at 141. See *O'Callaghan v Ballincollig Holdings Ltd* unreported High Court 31 March 1993.

⁴⁸ (1866) LR 1 HL 129 at 141.

⁴⁹ Unreported, High Court, 18 December 1969.

⁵⁰ [1981] ILRM 419

of building a school but that plan was subsequently abandoned. The plot was located within a council estate and its boundaries were undemarcated. The council built two houses on the plot but the work was completed and tenants were in situ before the plaintiffs became aware of the situation. When the plaintiffs brought proceedings to recover possession, the council sought to rely on the mistake limb of the doctrine. *Finlay P* made an order permitting the council to retain possession of the land on condition that it paid compensation to the plaintiffs to cover the market value of the site and the costs of purchasing an equivalent site. *Finlay P* was satisfied that the council was mistaken, in the sense that it did not wilfully build on the land belonging to the plaintiffs, and although this was not a case where the owners could be described as remaining wilfully passive or guilty of any wrongdoing, the doctrine could still intervene on the particular facts of the case. *Finlay P*'s approach in recognising an estoppel in the absence of any real unconscionability on the part of the legal owner has been described as "extreme" and an example of a hard case making bad law (Mee, pp210-211). It seems unlikely that it will be relied on as a precedent in the future. However, the obvious judicial sympathy for the predicament of this mistaken improver indicates that some other remedy should be available where there is no acquiescence on the part of the owner, thereby precluding reliance on the doctrine of proprietary estoppel, and the mistaken improver cannot demonstrate possession for the limitation period, thereby precluding reliance on the doctrine of adverse possession. The enactment of mistaken improver or building encroachment legislation, akin to that operating in Australia and discussed in the next section, would plug this vacuum.

An advantage of using proprietary estoppel and mistaken improver or building encroachment legislation to resolve boundary disputes is that it facilitates a tailored response to satisfy the equities of the case. The landowner may be ordered to convey the freehold or some other right to the claimant, to pay compensation or simply restrained from enforcing his strict legal rights. On the other hand, mechanical rules such as adverse possession are more straightforward to apply and may result in less litigation (Merrill, p. 1137).

3.5 Prescription

If a landowner has exercised rights over neighbouring land that are less extensive than exclusive possession it may be possible to claim a profit or easement by prescription pursuant to Land and Conveyancing Law Reform Act 2009.⁵¹ The claimant must be able to prove 'user as of right' (ie without force, secrecy or permission) for 12 years without interruption. Jourdan notes that a person who wishes to claim an easement or profit by prescription in the alternative to adverse possession may undermine his ability to prove the necessary elements of the adverse possession claim, in particular that he was engaged in exclusive use amounting to possession and had the necessary *animus possidendi* (2003, para. 23.26-23.28).

⁵¹ S33-38 of the 2009 Act. Until recently there were three methods of acquiring an easement by prescription: common law prescription; prescription under the doctrine of lost modern grant and pursuant to the Prescription Act (Ireland) 1958. Different prescriptive periods and slight variations in the requirements applied to each method. All three methods were abolished by the Land and Conveyancing Law Reform Act 2009 and replaced by a new method.

3.6 Works Order

Finally, certain boundary disputes may be defused by availing of the procedure provided by the Land and Conveyancing Law Reform Act 2009 which enables works, such as repairs or maintenance, to be carried out to party structures.⁵² A party structure is defined as including any fence, wall, ditch, shrub or tree which divides neighbouring properties or is situated so close to a neighbouring property that access is needed to carry out the works. It includes a structure which is situate entirely on one side of the boundary or which straddles the boundary and is co-owned or subject to some division of ownership between neighbours.⁵³ Where an adjoining owner disputes the right to carry out such works or fails to co-operate, a works order may be obtained from the District Court authorising the carrying out of the specified works subject to certain terms and conditions.⁵⁴

The remainder of this paper outlines the role played by mistaken improver and boundary encroachment legislation in Australia in settling boundary disputes. It also assesses the potential impact of the recent English reforms of the law on adverse possession on the resolution of boundary disputes. It concludes by outlining the valuable lessons which Ireland can learn from these experiences.

4 Boundary Disputes in Australia

In all Australian jurisdictions, except Western Australia and Victoria, adverse possession of boundary strips (or ‘part parcel’ adverse possession) is expressly or effectively prohibited. Instead, boundary disputes are typically resolved through mistaken improver or building encroachment legislation. This legislation confers jurisdiction on the court to provide a remedy where a person improves neighbouring land or erects a building that encroaches on his neighbour’s land. While a mistaken belief in ownership is required under mistaken improver legislation, building encroachment legislation may allow for a little leeway in this respect. Although a bad faith encroacher may not be absolutely barred from relief, ‘it would be a rare and exceptional case in which the court would make an order ... in favour of a person who, with full knowledge encroached on his neighbour’s land.’⁵⁵ Such a bad faith encroacher may be required to pay three times the value of the unimproved capital value of the strip if the court makes an order for compensation for the transfer, lease or grant of the land to the encroacher. O’Connor argues that this penalty provision should not apply where the encroachment was built by the encroacher’s predecessor in title (2006, at p. 214; 2007 at p.60). Aside from the conduct of the encroacher, various other factors are considered in deciding whether to grant relief, such as the conduct of the adjoining owner, the nature and extent of the encroachment, the situation and the value of the land and the loss and damage that would be incurred by the encroacher if he was required to remove the encroachment. As with

⁵² See s43-47 of the 2009 Act.

⁵³ S43 of the 2009 Act.

⁵⁴ S45-46 of the 2009 Act.

⁵⁵ *Haddans Pty Ltd v Nesbitt* (1963) 57 QJPR 21, 22 (SCQ, Gibbs J).

proprietary estoppel, the potential remedies are more flexible than adverse possession and include the possibility of orders for compensation.

5 The English Reforms

5.1 Background

The reforms to the doctrine of adverse possession included in the Land Registration Act 2002 stemmed from recommendations made by the Law Commission for England and Wales (1998, chapter 10 and 2001, chapter 14). The proposals were designed to introduce a system of adverse possession which was more consistent with principles of title registration (Law Commission, 1998, para. 10.2) and would confer added protection on landowners (Law Commission, 1998, para. 10.19). The new scheme aims to confine the acquisition of land by adverse possession to cases where it is necessary either in the interests of fairness (and 'land theft' is not involved) or to ensure that land remains saleable (Law Commission, 1998, para. 10.2 and 10.12-16).

5.2 A Qualified Veto System

The 2002 Act allows the registered owner to veto adverse possession applications which confers added protection on the owner against squatters. The veto system is, however, qualified so that the benefit of the doctrine is preserved for an adverse possessor who can satisfy one of three conditions. Two of the conditions seem designed to provide a remedy for applicants who went into possession pursuant to an informal transfer⁵⁶ or can prove that they are entitled to be registered by virtue of an equity by estoppel.⁵⁷ Of course the remedy of specific performance or proprietary estoppel may also be available to such claimants but, if the limitation period has expired, an adverse possession application to the Land Registry may be the most expedient way to update the register. The third condition is more relevant for the purposes of this paper as it concerns boundary land and imposes a good faith requirement. To satisfy the third condition the applicant must own land which is adjacent to the land in question and the applicant (or his predecessor in title) must have reasonably believed that the land in question belonged to him for the limitation period ending on the date of the application.⁵⁸ It was clearly considered desirable to preserve the doctrine's role in resolving certain boundary disputes (Law Commission, 1998, para 10.15-16).

5.3 Utilitarian and Ethical Dimensions

According to law and economics theorists, the doctrine of adverse possession plays a role in limiting the extent to which the land owner can extract 'quasi rents' when a mistake has been made in relation to the boundary (Mirceci and Sirmans, 1995). This is the term used to describe the windfall profits that the owner may make if the possessor has

⁵⁶ Sch 6 para 5(3).

⁵⁷ Sch 6 para 5(2).

⁵⁸ Sch 6 para 5(4).

invested substantial financial resources in the land. In such circumstances, market forces become irrelevant in determining the price which the owner can demand for the plot of land. According to Miceli and Sirmans (1995), when adjudicating boundary disputes the law should seek to encourage landowners to avoid making boundary errors before they undertake a sunk investment by making improvements (the deterrence objective) and limit the costs of boundary errors once they have occurred (the cost minimisation objective). Adverse possession could be viewed as a slightly crude and imperfect method of achieving these objectives. Although rent seeking will no longer be possible on the expiry of the limitation period, it may continue up until that point and the neighbour may ultimately get to keep the property even though he or she may have encroached deliberately or carelessly or even when no sunk investment has been made (O'Connor 2007 p.205). As O'Connor notes, the good faith requirement included in the third condition in schedule 6 of the 2002 Act ensures that Micelli and Sirman's deterrence objective is more effectively achieved (O'Connor, 2007, p.209).

The more obvious reason for the imposition of a good faith requirement was to inject an ethical dimension into the doctrine and ensure that only deserving squatters benefit from it. However, if the registered owner fails to exercise his veto, which will frequently indicate that the land has been abandoned land, a bad faith adverse possessor will succeed in his application to be registered owner. In such circumstances, ethical concerns are trumped by utilitarian needs and the doctrine is preserved to fulfill its traditional role in bringing such abandoned land back onto the market.

6 The potential impact of reform

6.1 Difficulties of Proof

The 2002 Act places the burden of proof on the applicant to establish his good faith and the requirement is made up of a subjective and an objective element. In civil law jurisdictions it is standard to include a legislative provision setting out a rebuttable presumption of good faith (Hargrave, 1998-1999, p.1200), thereby placing the burden on the owner to prove bad faith. In such circumstances the owner's case will tend to be based on the reasonableness of the adverse possessor's belief in ownership, given the difficulties in proving a negative (Hargrave, 1998-1999, p.1200). Although the 2002 Act fails to include a presumption of good faith, the Law Commission's discussion indicated that if the physical features suggest that the applicant owned the land in question, that fact, of itself, will in practice raise a rebuttable presumption that the applicant has satisfied the requirement. In such circumstances, it will be incumbent on the registered proprietor to show that the applicant knew or ought to have known that the parcel of land did not belong to him or her (Law Commission, 2001, para. 14.50-14.52). Another difficulty may arise in proving that good faith existed throughout the limitation period (O'Connor, 2007, p.209). However, if the applicant can demonstrate good faith when he or she initially went into possession, perhaps this will be viewed as giving rise to a rebuttable presumption of good faith throughout the limitation period. Therefore, although the requirements may appear stringent on a first glance, judicious interpretation

may reduce the difficulties faced by the applicant. However, it would clearly have been less risky to make express provision for these legal presumptions in the legislation.

6.2 Easier access to a proprietary estoppel remedy

It is important to bear in mind that a neighbour who is unable to satisfy the requirements of this condition may be able to avail of other remedies similar to those available in Ireland. A neighbour in a position to seek a remedy pursuant to the doctrine of proprietary estoppel may face a number of difficulties in satisfying the requirements for adverse possession as amended by the 2002 Act. If he was on the receiving end of a representation, he may have difficulties proving that his occupation was non-consensual. Also, if he was mistaken about the position of the boundary but discovered his mistake after going into possession, he will not be in a position to prove good faith throughout the limitation period. However, provided he can prove detrimental reliance, eg improvements or a building encroachment, he will be entitled to a remedy pursuant to the doctrine of proprietary estoppel and, since the enactment of the 2002 Act, this remedy has been made more accessible. In considering an adverse possession application based on the estoppel condition, the Land Registry adjudicator has jurisdiction to consider the elements of the proprietary estoppel claim and is expressly authorised to award a remedy less than ownership if it would be unconscionable for the registered owner to seek to dispossess the applicant but the circumstances are not such that the applicant ought to be registered as owner.⁵⁹

6.3 How to plug the remedial vacuum

When it comes to the resolution of boundary disputes, O'Connor appears to favour building encroachment legislation over the adverse possession model adopted in England pursuant to the Land Registration Act 2002 (2007, p.216). However it could be argued that they should not be viewed as exclusive approaches and, instead, should work in tandem, alongside the doctrine of proprietary estoppel, to provide three alternative methods for settling a boundary dispute. The encroacher who was uncertain about the location of the boundary or who discovered his mistake after he built the encroachment will not be in a position to satisfy the good faith requirement and so will be denied an adverse possession remedy pursuant to the 2002 Act. If the legal owner did not make a representation or was not aware of his neighbour's mistake, the encroacher will also be denied a remedy under the doctrine of proprietary estoppel. However, building encroachment legislation would allow the courts to provide a suitable remedy in such circumstances.

Alternatively, it could be argued that the good faith requirement set out in the 2002 Act should be amended so that proof of good faith when the applicant first entered into possession would suffice. However, this would allow for a transfer of ownership in circumstances where there was no building encroachment and the true owner is not in a position to demand quasi-rents. If no sunk investment is involved, it is submitted that the

⁵⁹ S110(4) of the 2002 Act.

doctrine of adverse possession should only be available for those who have developed a very strong psychological attachment to the land (Stake, 2000-2001). Such an attachment is more likely to develop if the occupier believed himself to be the owner over a long period of time. A requirement for good faith throughout the limitation period ensures that the doctrine of adverse possession operates in a more targeted fashion to provide a remedy for those who deserve it and cannot be accommodated by other legal mechanisms.

7 Conclusion

Currently, a broad range of remedies are available in Ireland to resolve the disputes that may arise in connection with boundaries, although proprietary estoppel and adverse possession have been the main focus of this paper. Improvements carried out in reliance on a representation or acquiescence by the owner in a mistake made by a neighbour in relation to the boundary may give rise to a remedy under the doctrine of proprietary estoppel. Alternatively, someone who can prove adverse possession of a boundary plot for the limitation period will acquire title to it. Assuming that the precedent set by the decision in *McMahon v Kerry County Council*⁶⁰ is not followed in the future, a mistaken improver may be denied a remedy if he falls between two stools – if he cannot rely on the doctrine of proprietary estoppel, as the owner has not acquiesced in his mistake, and he cannot demonstrate adverse possession for the limitation period. If Ireland decides to replicate the English good faith requirement in relation to adverse possession of boundary land, this remedial vacuum will become more pronounced as a mistaken improver or an encroaching builder who was uncertain about the position of the boundary or discovered that he was not the owner of the plot during the limitation period will be denied a remedy. It is submitted that any such reform should be supplemented by the introduction of legislation akin to that operating in Australia which would facilitate the consideration of a broad range of factors and provide for flexible remedies to resolve such difficulties. It is also submitted that the legislation imposing such a good faith requirement should specifically provide for the operation of certain legal presumptions to avoid the potential interpretative difficulties outlined earlier in relation to the English approach.

8 References

- Black F. and Healy T., 2008. Judge makes plea for peace in Kenny Land War. *The Irish Independent*, 12 April.
- Breen, O., 2000. Registration of Title and Overriding Interests – Another Crack in the Mirror? *Conveyancing and Property Law Journal* 5(3), 52.
- Commission on Land Registration, 2008. *Green Paper proposing Reform of Boundary Surveys in Ireland* Dublin: Irish Institution of Surveyors
- Fitzgerald, B., 1995. *Land Registry Practice*. 2nd ed. Dublin: Roundhall Press.
- Hargrave, L., 1998-1999. Ruminations on the Revision of the Louisiana Law of Acquisitive Prescription and Possession 73 *Tulane Law Review* 1197-1230.
- Jordan, S., 2003. *Adverse Possession*. London: Butterworths LexisNexis.

⁶⁰ [1981] ILRM 419.

- Law Commission, 1998. *Land Registration for the Twenty-First Century: A Consultative Document* (LC 254). London: Law Commission.
- Law Commission, 2001. *Land Registration for the Twenty-First Century: A Conveyancing Revolution* (LC 271). London: Law Commission.
- Law Reform Commission, 2005. *Report on Reform and Modernisation of Land Law and Conveyancing Law* (LRC 74-2005). Dublin: Law Reform Commission.
- Law Reform Commission, 2008. *Consultation Paper– Alternative Dispute Resolution* (LRC CP 50 2008). Dublin: Law Reform Commission.
- McAllister, D.L., 1973. *Registration of Title in Ireland*. Dublin: Incorporated Council of Law Reporting for Ireland.
- Mee, J. 1998. Lost in the Big House: Where Stands Irish Law on Equitable Estoppel? 23 *Irish Jurist* (ns) 187-219.
- Merrill, T.W., 1984-1985. Property Rules, Liability Rules and Adverse Possession. 79 *Northwestern University Law Review* 1122-1154.
- Miceli, T.J. and Sirmans, C.F., 1995. An Economic Theory of Adverse Possession 15(2) *International Review of Law and Economics* 161-173.
- O'Connor, P., 2006. The Private Taking of Land: Adverse Possession, Encroachment by Buildings and Improvement Under a Mistake. 33 *University of Western Australia Law Review*, 31-62.
- O'Connor, P., 2007 An Adjudication Rule for Encroachment Disputes: Adverse Possession or a Building Encroachment Statute? In Cooke, E., ed. *Modern Studies in Property Law: Volume 4*. Oxford: Hart, 197-217.
- Sperling, D., 2008. *Property Registration in Ireland: The Role of the State Guarantee* Dublin: Dublin Institute of Technology.
- Stake, J.E., 2000-2001. The Uneasy Case for Adverse Possession' 89 *Georgetown Law Journal* 2419-2474.
- Woods, U. and Wylie, J.C.W., 2005. *Irish Conveyancing Law*. West Sussex: Tottel Publishing.

Assessment on problems of issuances and transferences of Strata Title in Malaysia: The Perception of the owner

Kamaruzzaman, S.N., Naziah Muhamad Salleh and AlZawawi, S.A.F.

Department of Building Surveying, Faculty of the Built Environment, University of Malaya, 50603
Kuala Lumpur, Malaysia

Abstract

The ownerships of property as an asset, is a long-term investment for the owner. Strata Titles is a document of title which endorsed by the land administrator for the whole parcels of a subdivided multi storey building. Strata Titles important to purchasers as a proof of the ownership of the parcels they purchased as they will be able to persevere any deal which stated in Part IV of National Land Code that aloud any titles' transferred, mortgage, leasing and any permitted activities to be carried out. However, there are problems in applying strata titles which will affect the ownerships of the parcels in a multi storey building. Normally, everyone will blame the developers as the causes of failure to appraise the issuance of the Strata Titles. This study will be focused on the problems which caused the failure of the issuances and the transferred of the strata to the purchasers. The study analyse from the perceptions of the owner of the parcels.

The findings of the survey found that the service and the maintenance fee are slightly high compared to the services offered by the Management Corporations. Besides, the owner refused to pay any outstanding bills on maintenance of their building and the mortgage and stamp duty disbursements are another factor for them to refuse the transference of ST. Based on these findings, several recommendations identified to reduce the problems of the issuance and transfer of the Strata Titles to the purchasers by providing transparent information, campaign and education regarding the importance and the process of issuances and transferences of ST

Keywords: Strata Titles, issuance, transference, perception, owner

1 Introduction

Development of strata property has rapidly grown in recent years. It is a trend of 20th century to increase the urbanisation in most countries including Malaysia. From the perspective of supply and demand, rapid urbanisation, high population growth, increase in investment opportunities as well as increasing job opportunities in the cities are making land scarce resources. The logical response is to maximise land use by building upward (Bahari.A 2001). As evidence, for the first quarter of 2008 Malaysia has 2,423,105 units of strata residential, 8,735,173 square meters shopping complex and 14,881,383 square meters purpose built office (NAPIC Q1 2008).

The accelerated in numbers of strata property has lead to the development and change on the pattern of property ownerships. High rise buildings have escalating built and its offer different types of property ownerships than landed property. This type of property gives a right to a property with the absence of strata title. As in order to fulfil demands of the right on high-rise properties owners, Malaysia has applied Strata Title Act (STA 1985) since late 1985. It is part of the National Land Code 1965 (NLC) which dealt with subsidiary titles. Basically, subdivision under the STA 1985 means the registration and issuing of a separate strata title for every unit in a building of two storeys or more. When the building is subdivided, title to the land which the building is subdivided. Title to the land on which the building erected continuously exists but it is automatically established on the

registration of strata titles to the parcels of the building, where the high rise properties were managed by the management corporation appointed by developer.

Unconsciously, undelivered Strata Title (ST) will confer the troublesome to the owners of the parcels. The documents might cover the owner on various uncertainties. It is seriously effects on the ownerships, difficulties in obtaining financing and uncertainties of the future (Khadijah, 2002). In Bercham, Ipoh Perak, the owner of the low cost apartment was unable to receive the strata title for their parcels even though their houses have been occupied since 1994 (The Star, 24 November 2003).

2 Strata Title: The Ownership documents

Strata title means the ownerships of a unit in a subdivided building. The ownership of the unit is evidenced by a separate title called the strata title under STA, which issued in respect of the unit. The management corporation, which consists of all the owners of the units, is the medium through which the owner control and manage the strata scheme. The management corporation is responsible for the maintenance and the management of common properties such as open spaces, lifts, corridors, gardens and community facilities other than the units' respect of which of the individual titles have been issued and registered.

Construction of condominiums, apartments, flats, and high-rise buildings in the urban areas such as Kuala Lumpur, Johor Bahru and Penang is a common phenomenon due to scarcity of lands in those areas and high land costs. The property market boom supported the massive development carried out in urban areas. This has encouraged some developers to construct their projects before the ownership of title has been issued. The effect is that the transfers of such properties can be problematic where there may be delays in transferring property rights to the respective purchasers. Since the transfer of title properties involves some interested parties such as developers, financiers, buyers and solicitors, the problem can have certain implications to one or more (Hussin, 2002).

As per third quarter of 2008, The National Property Information Centre's (NAPIC) reported that the strata residential stock in Malaysia consists of 1,232,848 units. The statistic on the strata residential includes Low Cost Flat, Flat, Service Apartment, Condominium / Apartment and Service Apartment. The detail of the statistic is tabulated as table 1

State	Low Cost Flat	Flat	Service Apartment	Condo / Apartment	Total (Unit)
Selangor	182,424	145,722	8,843	177,694	514,683
WP KL	82,519	48,136	5,507	126,269	262,431
Pulau Pinang	50,351	98,986	839	33,468	183,644
Johor	44,353	19,162	2,948	33,468	99,931
WP Putrajaya	-	2,538	- - 2,		538
WP Labuan	1,980	-	-	308	2,288
Perak 7	,735	2,168	1,259	4,492	15,654
Neg.Sembilan 10	,381	6,210	5,407	13,395	35,393

Melaka 5,	919	6,073	345	10,359	22,696
Kedah 4,	135	1,272	748	630	6,785
Pahang 3	,768	2,589	647	7,479	14,483
Terengganu 3	,120	1,176	-	625	4,921
Kelantan 514		436	-	778	1,728
Perlis 898		96	-	96	1,090
Sabah 14	,258	7,992	904	15,970	39,124
Sarawak 13	,728	2,497	9,234	-	25,459
Total 426	,083	345,053	36,681	425,031	1,232,848

Table 1: Strata Residential Stock in Malaysia
Source: NAPIC, Ministry of Finance

The statistic above shows that the strata residential are mostly built in the city that is Wilayah Persekutuan Kuala Lumpur, Selangor, Penang and Johor. However the statistic details the quantity of unit and not the strata development by project. Therefore Dewan Bandaraya Kuala Lumpur on FABIABCI Malaysia & PPK Seminar (2008) informs that the total number of strata development are 9,449 which consists of 9,319 of residential projects and 130 of commercial project. The figure 1 illustrates the strata development by project in Malaysia.

Strata Development in Malaysia

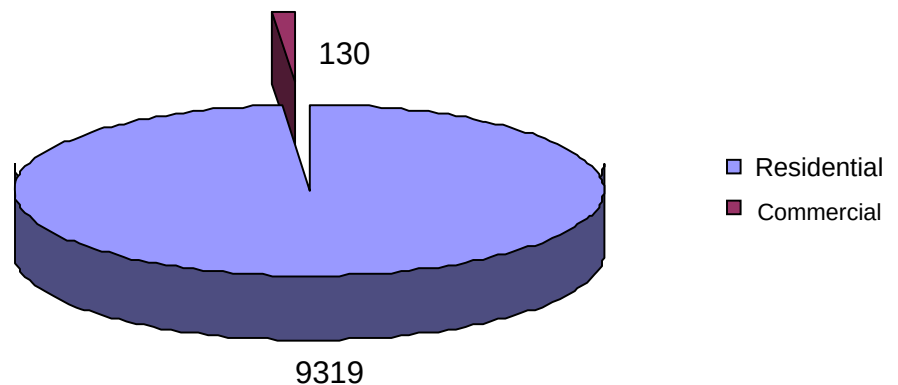


Figure 1: Strata Development by Project in Malaysia In 2008
Source: Dewan Bandaraya Kuala Lumpur (2008), "Property Managing Agent in Malaysia", FABIABCI Malaysia & PPK Seminar

2.1 Applications of Strata Tittles

Application for subdivision of building has to be made by the proprietor of land. The current process takes 268 days to issue the ST (Director's of Lands and Mines (2008). Figure 2 explained the flows of application's process for ST in Malaysia. The process will be extended once the application is rejected and it will start all over again. This consumes another time to process another application.

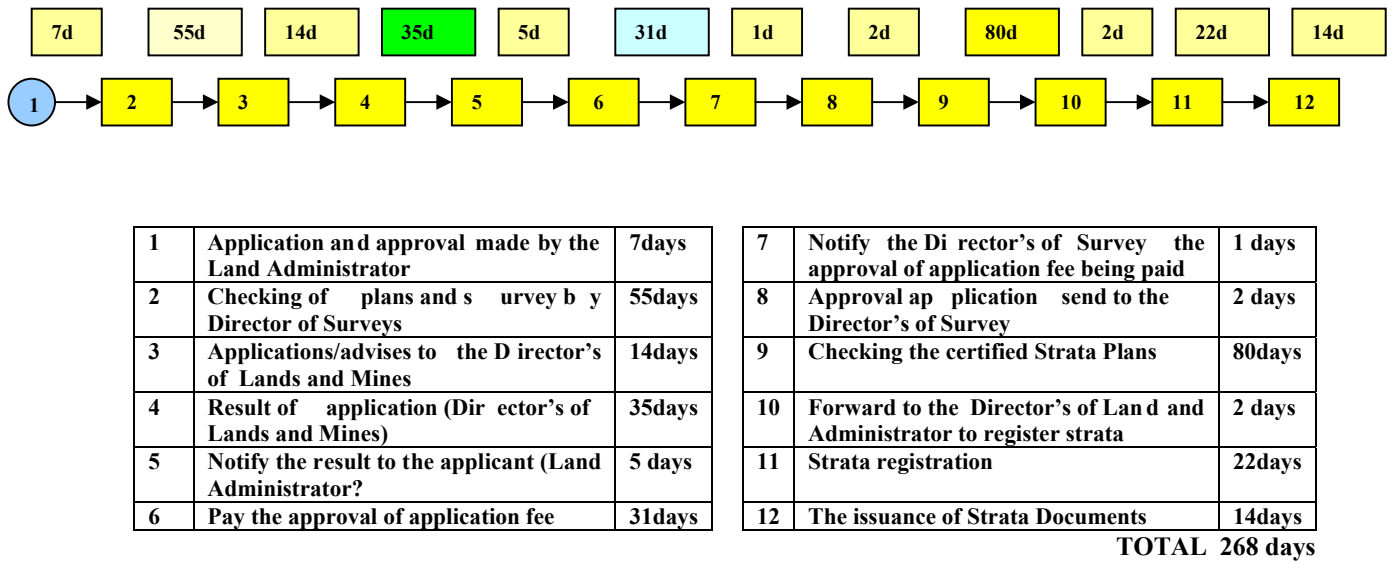


Figure 2: The Current Scenario of Processing Strata Titles

Source: Director's of Lands and Mines Office (Felderal) Departments (2008), "Panduan Pelaksanaan Penambahbaikan Pengurusan Tanah", Circulation 1/2008

At 2008, the Director's of Lands and Mines has proposed the new flow of processing ST where it takes only 123 days to issue. Figure 3 shows the flow of the new process of ST's application.

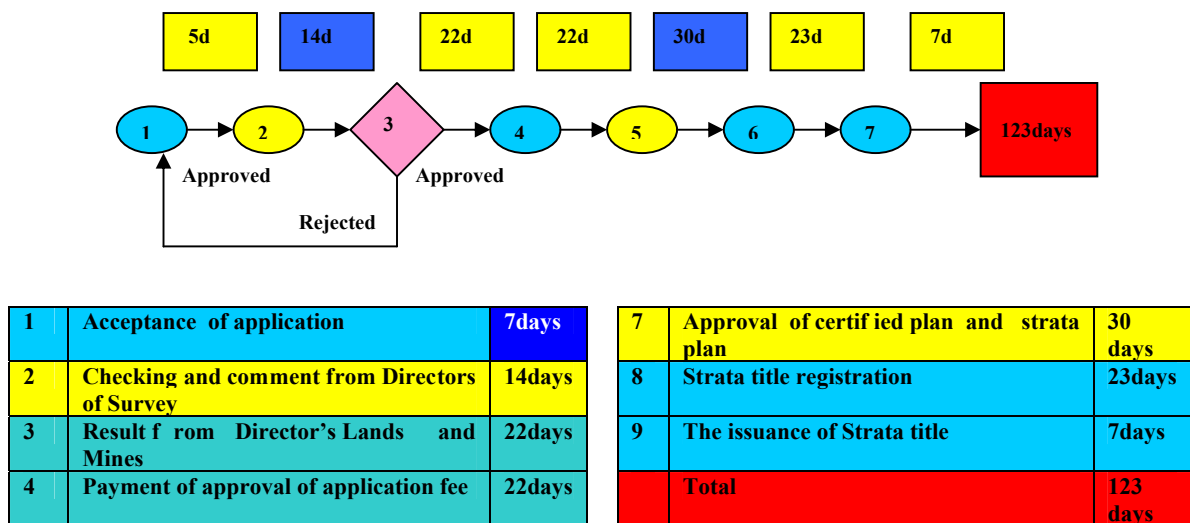


Figure 3: The New Proposal of Processing Strata Titles

Source: Director's of Lands and Mines Office (Felderal) Departments (2008), "Panduan Pelaksanaan Penambahbaikan Pengurusan Tanah", Circulation 1/2008

2.2 Failure of application and transference ST

Azimuddin (2007) said that a total of 481,277 strata titles were successfully issued from 1993 till 2006. Of this, the highest number was for Selangor (159,453 titles), Penang (120,170), Kuala Lumpur (115,036) and followed by Johor, Negeri Sembilan, Malacca, Perak, Pahang, Terengganu, Kedah, Kelantan and Perlis (Homefinder, 2007).

The strata title applications received between 1993-2006, Selangor again topped the list with 6,092 applications for 319,102 unit of parcels were sent in, of which 3,385 applications for 195,274 units were approved, 59 applications for 3,341 units were rejected and another 2,648 applications for 120,487 units are still being processed. Kuala Lumpur sent in the second highest number of applications 2,570 applications for 171,870 units, of which 2,134 applications for 142,954 units were approved and 241 applications for 12,038 were rejected. Another 195 applications for 16,878 units are still being processed. Penang sent in 1,368 applications for 148,286 units, of which 1,223 applications for 128,572 units were approved, 30 applications for 287 units were rejected and a further 115 applications comprising 19,427 units are still being processed (Homefinder, 2007).

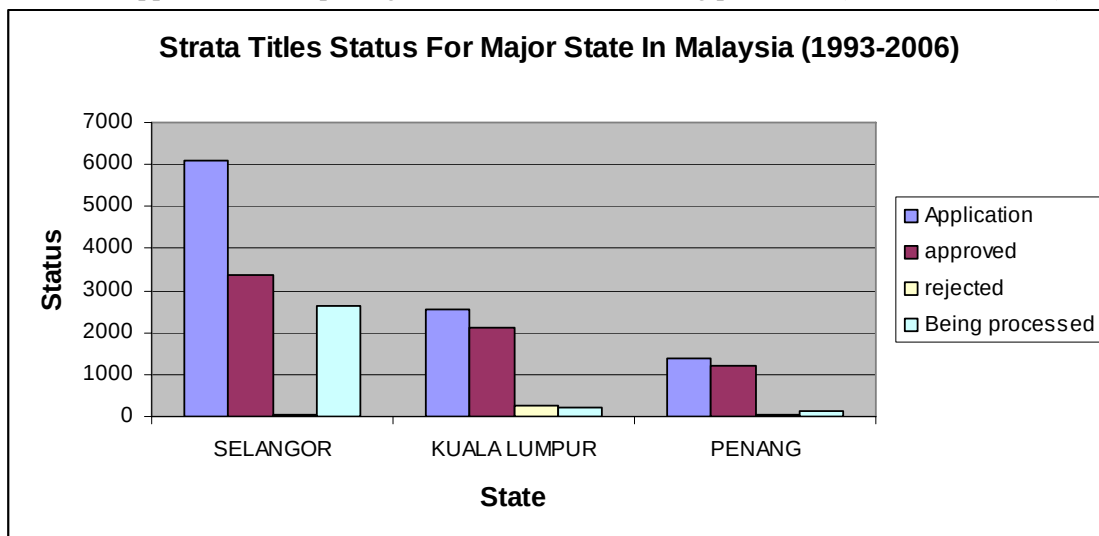


Figure 4: Strata Titles

Source: http://homefinder.com.my/index.php?option=com_content&task=view&id=523&Itemid=71

When one compares to the number of high rise development approved and completed each year as reported by the Valuation Department versus the number of strata titles issued by the Ministry since 1993, these numbers seem really small (NAPIC, 2006).

Exploring the reason of failure for the titles to be transferred, Za in (1997), Helmi (1992), Robiah (1991) alleged that the applications were not accompanied by the requirements of section 9 and 10 of STA 1985. They were not accompanied with satisfactory documents such as;

- i. schedule of parcels
- ii. architect's/engineer's certificates
- iii. Certificates of fitness (C.F/CCC) for occupation or issued by public or local authority
- iv. Approved buildings plans (including approved amended building plans)
- v. differences in details of the plans with actual building erected on the land
- vi. conflicting details in the building plan pertaining to common property with the actual building being used
- vii. buildings erected on lands being amalgamated

- viii. buildings do not have adequate access
- ix. Share units of each parcel allotted were not equitable
- x. permits to use space above the reserved land if there is any eave, awning or balcony projecting over road reserve

However, the above mentioned were normally affected by developer. Naziah (2004), Magendran (2000) and Sukeri (1999) found that the problems of issuance and transference ST are caused by the owners too. The problems appeared when the owner refused to disburse the mortgage and stamp duty, lack of understanding towards the importance of the ST, dissatisfaction on the service and the maintenance fee. Thus, it will present big effects to transfer the ST to them.

Figure 5 shows the flows of the transference of ST to tenants. The dispute occurs when the tenant refuse to pay any outstanding bills. This force the developer to hang the process with the financier

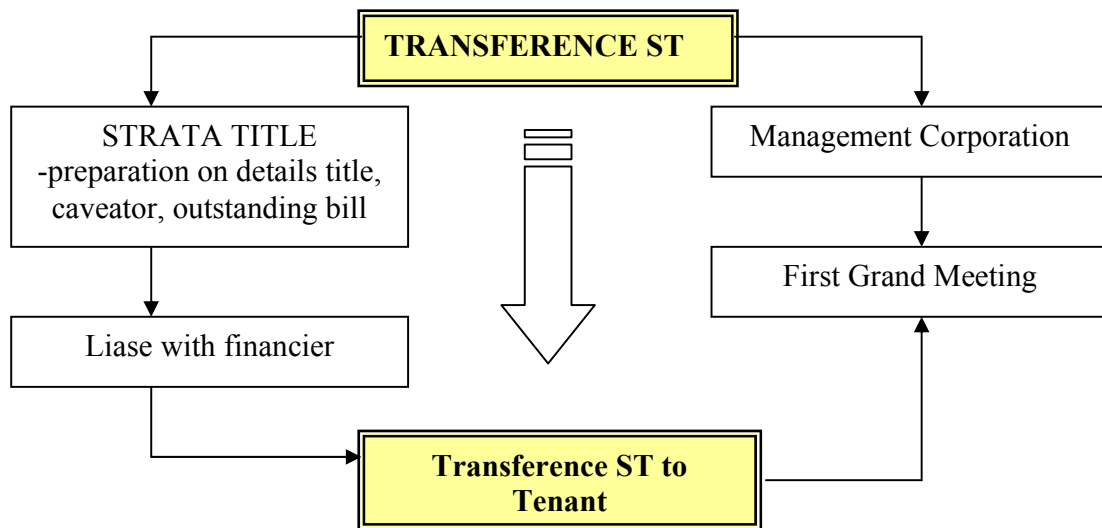


Figure 5: The process of transference ST
Source: inner circulation: developer (2008)

The service charge is the means by which an owner is able to recover from the occupiers the cost of providing the services for the benefit of those premises (Singh, 1993). However, it would be irrelevant for the tenant themselves if they wouldn't care less on what they are paying ever since. Tawil (2007) in his study revealed the difficulties face by management in collecting service charge. Those residents were not satisfied with the service charge fare and they need to know the disbursement of service charge even they paid on time. There are lots of complaints and also they think that they are paying more than the quality of service given to them. Schedule H falls under the Standard Sales and Purchase Agreement (SPA) has mentioned the service charges are paid to the vendor or developer to run the maintenance of the said building.

The SPA also describes the reimbursed fees should be settled during the transference of Certificate of Fitness (CF) and ST. include in the mortgage and stamp duty disbursement.

3 Methodology

The framework of this research is as shown in figure 6.

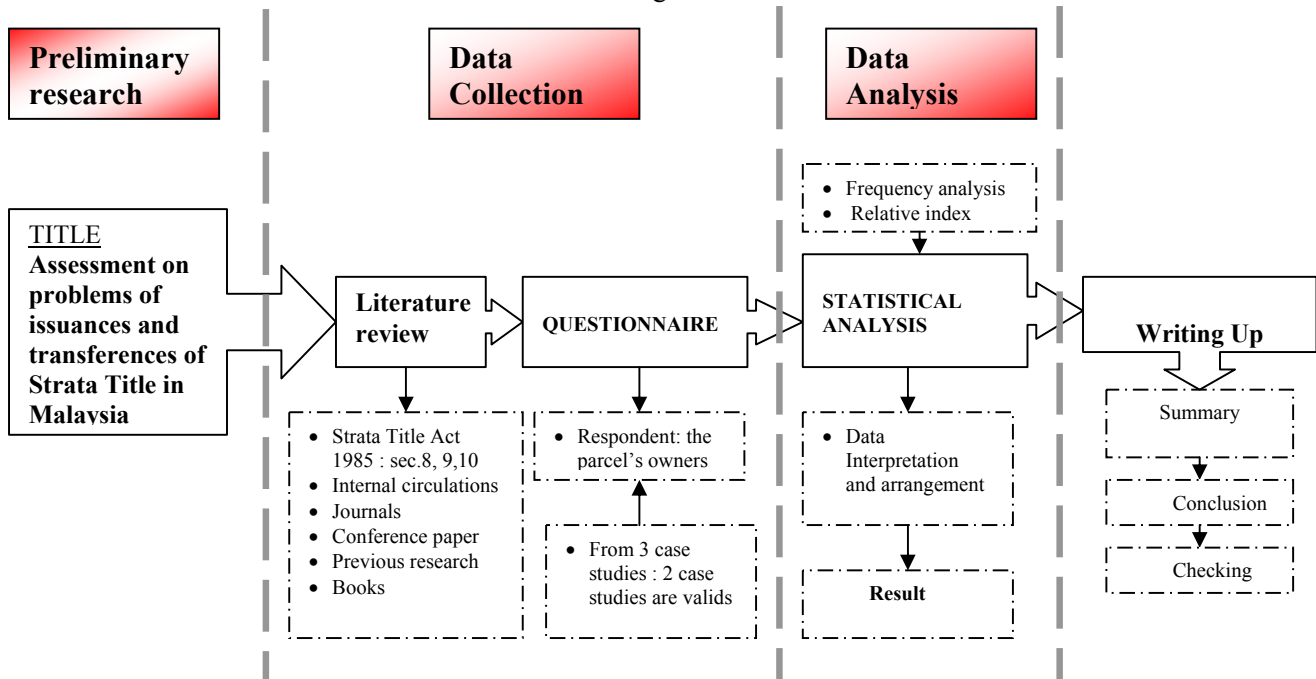


Figure 6: Framework

3.1 Case study of two residential strata developments

The research strategy adopted is quantitative in nature and data collecting technique is through personally assisted questionnaire. Respondents were randomly selected from two of three case studies as one is a service apartment and is owned by a developer. The rest case studies involved the owners of parcels of medium cost apartment in Selangor and townhouses apartment resorts in Negeri Sembilan. Both developments have received their Certificate of Fitness (CF) within 2001 to 2003 and confront with the same problems while processing the ST. The problems perceived are;

- buildings erected on lands being amalgamated
- caveat on the master title.
- Letter of concerns from the financier of each caveat

The age of the strata developments provides a good comparison with regard to any bylaw issues during the applications of ST

3.2 Information of case studies

3.2.1 Case Study 1: Selangor

Phase development involved 3 blocks of apartment. The CF issued within different timeframes, the details as mentioned in table 2.

Phase	Nos of Unit	Date of CF	Nos of Tenant refuse to pay outstanding bills
55A 22	0	31.05.2003	26

55B 24	0	06.06.2002	30
55C 24	0	25.06.2001	22
TOTAL	700		78

Table 2: Informations of case study 1
Source: project department of developer office, 2008

3.2.2 Case Study 2: Negeri Sembilan

Phase development involved 4 blocks of apartment. The CF issued within different timeframes, the details as mentioned in table 3.

Blocks	Nos of Unit	Date of CF	Nos of Tenant refuse to pay outstanding bills
I 24		08.09.2001	6
J 24		08.09.2001	9
K 36		02.04.2001	-
L 36		02.04.2001	8
TOTAL	120		23

Table 2: Informations of case study 2
Source: project department of developer office, 2008

Based on interviews with the Project Manager and Managing Agent of the two developments, the respondents then selected among tenants who have refused to accept the transference of ST on certain purposes. The respondents are among that are denied to pay outstanding bills (mortgage disbursement, stamp duty and maintenance fee). The selection made assuming the rest tenant who has resolve the outstanding bills are prepared and fully committed to accept the transference of ST.

Between May 2009 to June 2009, 101 questionnaires distributed to the selected tenant. Only 47 questionnaires/ from the tenants agree to answer. 5 of the tenants prefer to answer the questionnaire unassisted, while the rest are assisted. The collection of primary data was done by using a structured questionnaire. The questionnaire for this research consisted of a few different types of questions, namely open and closed questions. The large part of this questionnaire consisted of closed questions, where the respondent was asked a question and required to answer by choosing between a limited numbers of answers. The answers for some parts of the closed questions were based on Likert Scaling. The Likert scale chosen was range from 1 to 5, where 1 means 'not satisfactory at all', and 5 means 'most satisfactory'. Generally, the questionnaire designed for this study consisted of 7 parts. Part 1 covered basic owner's profile. Part 2 provided the information on understanding towards the Concept of ST, part 3 on maintenance fee, Part 4 : mortgage disbursement, Part 5 : Stamp duty disbursement, part 6: dissatisfaction on the service and maintenance and part 7 is the conclusion.

Data obtained from the feedback will be analyzed using Frequency Analysis and Relative Index. It is then written back in terms of Bar Chart, Column-Chart and Pie Chart to generate findings. The data will be converted into percentages to make more understandable. The results is then evaluated and used in the findings of objectives in the study. A brief summary was produced to conclude the outcome of the survey.

3.3 Frequency Analysis

Frequency Analysis depends on the percentage of respondents giving the same answers. It is also used to measure the degree of agreement for certain statements. The formula of Frequency Analysis is as below:

$$\text{Percentage (\%)} = (n/N) \times 100\%$$

Where:

n = Number of respondents

N = Total number of respondents received

3.4 Relative Index

To evaluate the ranking of different factors that cause the problems of issuances and transferences of Strata Title from the point of view from respondents, rating is made against the five-point scale described previously were combined and converted into relative important indices for each factor, adopting the Relative Index (RI) ranking technique. This determined the relative ranking of the different factors by comparing the individual value of the relative importance indices for each factor.

The highest ranking referred to the highest RI value. The individual numerical rating for each of the identified factors (from the Likert Scale) was transformed to relative factors, by using the following formula:

$$RI = \frac{n1(1) + n2(2) + n3(3) + n4(4) + n5(5)}{5N}$$

Where:

n1 = Number of respondents for “Strongly disagree”

n2 = Number of respondents for “Disagree”

n3 = Number of respondents for “Neutral”

n4 = Number of respondents for “Agree”

n5 = Number of respondents for “Strongly agree”

N = Total number of respondents

4 Results and Findings

4.1 Respondents Background

The questionnaires have accepted 44 respondents. The respondents then being categorized into races and the year they have tenured the parcels.

Race	Period of living in the parcels (Year)				Total
	1-2	3-4	5-6	>6	
Malay	0	6	20	6	32
Chinese	0	2	2	2	6
Indian	0	0	6	0	6
Others	0	0	0	0	0
Total	0	8	28	8	44

Table 4: Respondents Background

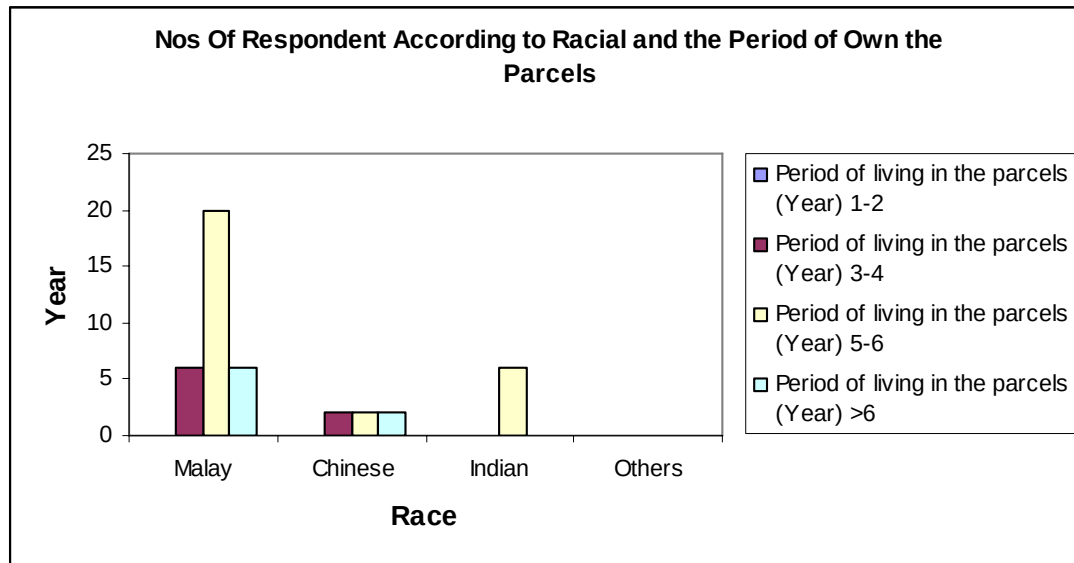


Figure 7: Respondents Background

The Malay owner is the highest numbers of respondent who have occupant the parcel for 5-6 years.

4.2 Information of HS

According to Table 5. Most of the respondent obtained the information about ST from the Sales and Hire Purchase documents (44%) once they purchased the parcels. The notification for the developer to forward the ST application as mentioned in Section 8 of the STA 1985 that developers must apply for strata titles for each individual project development within 6 months of receiving the Certificate of Compliance and Completion (CCC) (previously Certificate of Fitness for Occupation) or face a fine of not less than RM 10, 000 or not more than RM100, 000 for any failure of forwarding the ST documents. What's more, the developer is also subject to a minimum fine of RM100 and a maximum fine of RM1, 000 for each day the offence continues.

Source of Information	Frequency	Percentage
Sales & Hire Purchase Documents	22	44%
Sales Agent	16	32%
Media	0	0%
Reading	2	4%
Authority Circulations	0	0%
Friends	4	8%
Experience	6	12%
Others	0	0%
TOTAL	50	100%

Table 5
Sources of Information on Strata Titles

This is followed by the information given by the Sales Agent (32 %) and the respondent's own experience (12%)

For high rise residential buildings, the first encounter with the term 'service charges' is when the purchaser reads through the standard sale and purchase agreement. The provision of importance here is Section 16 Schedule H of the *Housing Developers (Control and Licensing) Act 1966, Regulation 1989* which reads as follows:

"The Purchaser shall be liable for the maintenance and management of the common property and for the services provided by the Vendor prior to the establishment of a management corporation under the Strata Title Act 1985."

Schedule H falls under the Standard Sales and Purchase Agreement. The service charges are paid to the vendor or developer. There is no explicit definition of service charges but when it is read in conjunction with other provisions of Schedule H it can be taken to mean the expenditure incurred in the maintenance and management of the common property. It also mentioned the term of disbursement of mortgage and stamp duty fee during the transference of ST. the tenant/purchaser are well informed that without paying the outstanding bills will lead them trouble to get the ST.

4.3 Problems features among Owner On Issuance and Transference of ST

Although the respondents are well-informed about the importance of ST, most of them found it hard to comply with the regulation of getting the ST as being mentioned earlier. Figure 4 unveiled majority of respondents (39%) feels that service and maintenance fee of their apartments were slightly high and it does not comply with the satisfaction of their service (23%). Anyway, the respondent found the disbursement of stamp duty (28%) and the outstanding bills of service and maintenance fee (22%) might also affect the issuance of strata title. Those matters should be resolved by respondents before the transference of ST.

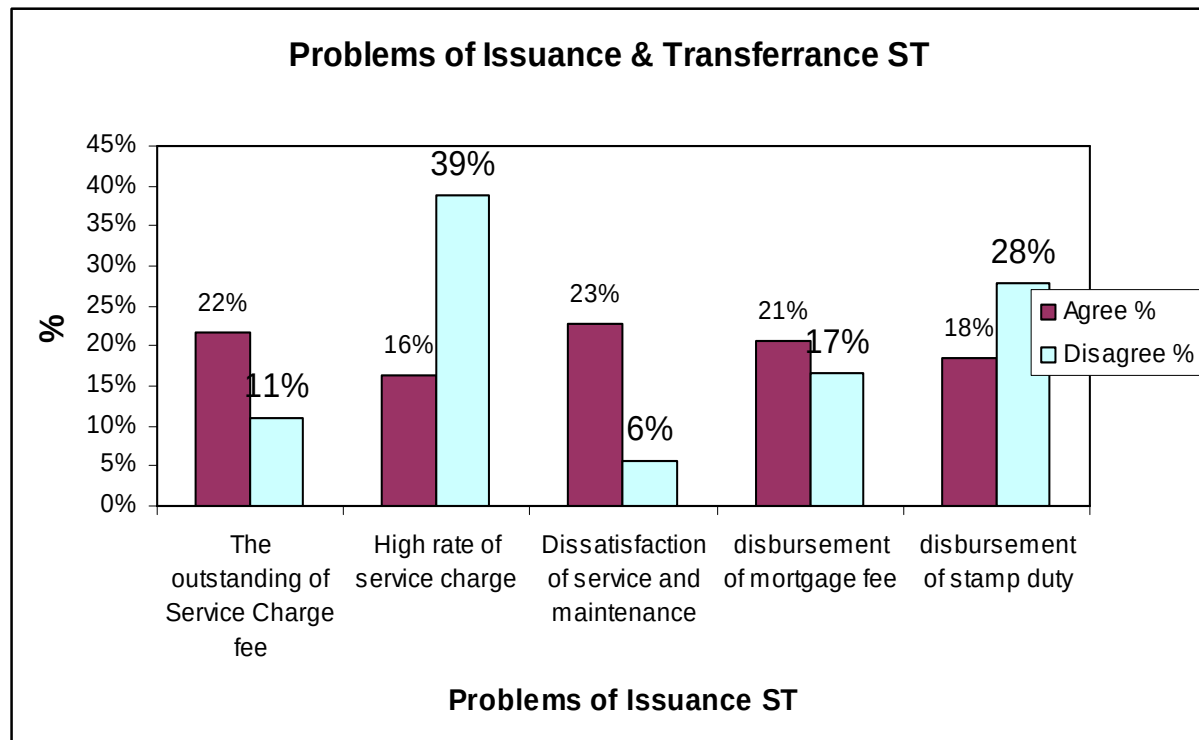


Figure 8
Problems of Issuance and Transference Strata Title

Service charge provisions are incorporated in tenancy agreements to ensure that the landlord's capital asset is being maintained at the tenant's expense. However, tenants often complain that certain services are not being provided or that the services are of a poor standard (Noor, 2007). The tenant always suspects that he is paying too much. The constituents of service charge will vary according to the type of building, type of usage, size, and term of tenancy or lease agreement as well as any explicit definitions in the agreements.

4.4 The Significance of Transference the ST

The most common problem associated with the transference of ST is the maintenance and service fee that run by the management corporation assigned by developer is the highest factors (0.73 Relative Index). Followed by the dissatisfaction on building services and maintenances that are run by the management corporations (0.68 relative Index). Besides, there are some residents who still have the outstanding bills of maintenance and service fee that should be settle down before the transference of ST (0.55 Relative Index). The summary of the problems and refusal of respondents to receive the ST can be viewed at table 4 below.

DESCRIPTION	Likert Skill					Frequency analysis		RI	Rank
	1	2	3	4	5	<3	>3		
	No of Respondent					%	%		
Outstanding maintenance and Service Fee	2	6	4	30	0	18%	68%	0.55	3
Maintenance and service fee too high	0	2	6	20	16	5%	82%	0.73	1
Dissatisfaction on building services and maintenances	0	6	6	10	22	14%	73%	0.68	2
Mortgage fee disbursement	2	8	14	16	4	23%	45%	0.38	4
Stamp duty disbursement	6	8	16	12	0	32%	27%	0.22	6
Unfamiliar with Strata Tittle	4	8	30	2	0	27%	5%	0.04	7
Developer to settle the Strata Tittle	0	8	22	14	0	18%	32%	0.25	5
note: Scale used : 1=Strongly Disagree, 2=Disagree, 3=Neutral, 4=Agree, 5=Strongly Agree									

Table 6
Problems of Issuance and Transference Strata Title

5 Conclusions

The main objectives of this research to determined the factors of the issuance and the transference of ST on the owner's perspectives. Taking into account of the above discussion, we found the owners understand the concept of ST but seems refused to accept the transference of ST based on a few factor namely the service and the maintenance fee are slightly high compared to the services offered by the Management Corporations. Besides, the owner refused to solve any outstanding bills on maintenance of their building while feel the mortgage and stamp duty disbursements are another factor of them to refuse the transference of ST. Based on these findings, several recommendations identified to reduce the problems of the issuance and transfer of the Strata Titles to the purchasers by providing transparent information, campaign and education regarding the importance and the process of issuances and transferences ST.

REFERENCES

- A.K. Robiah (1991), "Processing and Registration of Strata Titles Within the Seven Months Time Frame: an Appraisal", Strata Titles and Management Corporations Seminar, Kuala Lumpur
- Dewan Bandaraya Kuala Lumpur (2008), "Property Managing Agent in Malaysia", FABIABCI Malaysia & PPK Seminar
- Directors of Lands and Mines Office (Federal) Departments (2008), "Panduan Perlaksanaan Penambahbaikan Pengurusan Tanah", Circulation 1/2008

- Gurjit Singh (1993). "Condominium Management in The Surveyor Vol 28 No 1". Petaling Jaya.
- Gurjit Singh (1996). "Property Management in Malaysia." Federal Publications. Selangor. pp.53 – 81.
- Hussain (1999), *Strata Title in Malaysia*", Pelanduk Publications, Kuala Lumpur
- Hussin, K., & Pardi (2002), F. *Apartment Law In Malaysia Buying Properties Without Title: Legal Point Of View*.
- Magendran (2000), "Faktor-faktor yang mempengaruhi Permohonan Hakmilik Strata dan Pembentukan Perbadanan Pengurusan Oleh Pemaju Bangunan Bertingkat". Unpublished. Universiti Sains Malaysia, Pulau Pinang
- Mohd-Tawil, N., Ramly, A., Che-Ani, A. I., Usman, I. M. S., Tahir, M. M., Zain, M. F. M., et al. (2009). Service Charge Collection of High-Rise Residential in Kuala Lumpur, Malaysia; Owner's Perspective. *European Journal of Social Sciences*, 10(1).
- M.S Khalid (1999), "Keengganan Pembeli Menerima Pindahmilik Strata". Unpublished. Universiti Sains Malaysia, Pulau Pinang
- M.S Naziah (2004), "Kajian Permohonan dan pindahmilik strata: Persepsi Pemaju dan Pemilik Petak". Unpublished. Universiti Sains Malaysia, Pulau Pinang
- Noor, M. N. M (2007), Adaptation of best value service on rationalization of service charge for purpose built office (PBO) in Malaysia, 13th Pacific-Rim Real Estate Society Conference Fremantle, Western Australia, 21 to 24 January 2007.
- Star Newspaper, "Resident Want State to Help Get Strata Titles for Flat", 24 November 2003, Malaysia
- The National Property Information Centre's (NAPIC) (2006), Property Market Report, Ministry of Finance, Malaysian Government
- The National Property Information Centre's (NAPIC) (2008), Property Market Report, Ministry of Finance, Malaysian Government
- http://homefinder.com.my/index.php?option=com_content&task=view&id=523&Itemid=71 viewed 1 June 2010
- <http://www.rehda.com/bulletin/10/bulletin-1002>. Pdf viewed 1 June 2010