Volume 18, No 2 December 2006

This journal was launched in 1989 with the generous assistance of our Foundation Sponsor:

McLaughlin Gordon & Lennon
Gold Coast, Australia

The **BOND LAW REVIEW** is a fully refereed journal published twice each year by the Bond Law Review Editorial Committee, Faculty of Law, Bond University, Australia.

The Review is available on direct subscription. Please send your order to:

Administrator
Bond Law Review
Faculty of Law
Bond University
Gold Coast QLD 4229
AUSTRALIA

Telephone:	07 5595 2039
Facsimile:	07 5595 2036
International Facsimile:	+ 61 7 5595 2036
Email:	blr@bond.edu.au

We welcome the submission of articles, notes and comments for publication in the journal. Please submit manuscripts to the Administrator, Bond Law Review at **blr@bond.edu.au**. Citation style should conform to the most recent edition of the *Australian Guide to Legal Citation* as published by the Melbourne University Law Review Association Inc, also available electronically at http://mulr.law.unimelb.edu.au/aglc.asp

This issue is to be cited as (2006) 18.2 Bond LR and was published in December 2006.

ISSN 1033-4505

CONTENTS

Articles

DIRECTORS' LIABILITY TO CREDITORS – WHAT ARE THE ALTERNATIVES?

*Helen Anderson**

Precis

It is timely to examine alternatives to imposing liability on directors for creditor losses when a company is unable to pay its debts. Directors' personal liability for corporate fault and corporate social responsibility are currently being investigated by parliamentary committees, but there are dangers of risk aversion and unfair legal complainace burdens on directors if liability is extended too far. This article examines alternatives to director liability, such as mandatory capitalisation or insurance, shareholder liability and government funded schemes, to assess whether they provide appropriate and sufficient compensation for losses in times of corporate failure. It is concluded that imposing liability on directors appears to be a superior way in which to ensure creditor protection because it not only potentially gives creditors access to funds for compensation but also deters the adverse behaviour which may be a cause of loss to creditors.

Introduction

The Corporations and Markets Advisory Committee has recently[1] released a discussion paper examining the personal liability of directors for corporate fault.[2] It is

* Dr Helen Anderson LLB (Hons) Grad Dip Bus (Acc) LLM; Barrister and Solicitor of the Supreme Court of Victoria; Senior Lecturer, Department of Business Law and Taxation, Monash University. This article formed part of my Ph D thesis on directors' liability to creditors.

1 May, 2005.

2 Corporations and Markets Advisory Committee, Parliament of Australia, *Personal Liability for Corporate Fault* (2005). Senator Campbell's reference spoke of directors being 'subject to a range of both general common law and statutory duties'. He also suggested that CAMAC could consider 'whether this potential liability would result in a disincentive for persons accepting or continuing to hold directorships; and directors engaging in entrepreneurial but responsible risk taking'. This Discussion Paper has, however, concentrated on direct and derivative personal liability for corporate criminal fault. A later reference to CAMAC,

based on a reference to the Committee from Senator Ian Campbell, who was 'concerned that duties being imposed on directors by various pieces of legislation may result in inconsistent compliance burdens and increased costs for business'.[3]

Many of these duties are imposed for the protection of creditors, and often it appears that the debate on creditor protection in the event of a company's inability to pay concentrates on choosing between letting the loss lie where it falls or imposing liability on directors for that loss.

There are, however, alternative ways of ensuring recovery for creditors, such as imposing liability on shareholders or government subsidised schemes of compensation. Indeed, employees currently enjoy the benefits of such a scheme, outlined below. Nonetheless, imposing liability on directors appears to be a superior way in which to ensure creditor protection because it not only potentially gives creditors access to funds for compensation but also deters the adverse behaviour which may be a cause of loss to creditors.

However, imposing liability on directors is not without cost, and Senator Campbell's concerns are real ones. Inconsistent or excessive compliance burdens may deter qualified people from accepting directorships and may discourage appropriately risky behaviour which is beneficial to shareholder wealth maximisation, the creation of employment and the economy as a whole. It is therefore timely to consider alternative ways of protecting creditors and assess whether they are effective means of achieving equal or better compensation for creditors without the need to impose liability on directors.

The article is divided into parts. Part II will briefly address economic theory to determine whether in fact creditors require compensation or whether the loss should lie where it falls. It will be seen that some creditors are arguably already protected by contract against the risk of loss ex ante, and it is maintained by economic theorists that

---

dated 23 March, 2005 and entitled 'Reference in Relation to Directors' Duties and Corporate Social Responsibility' has now come from the present Parliamentary Secretary to the Treasurer, the Hon Chris Pearce. One of its questions is '[s]hould the Corporations Act be revised to require directors to take into account the interests of specific classes of stakeholders [other than shareholders] or the broader community when making corporate decisions?' The Parliamentary Joint Committee on Corporations and Financial Services is currently calling for submissions on Corporate Responsibility and Triple Bottom Line Reporting, and will report on 29th November, 2005

3 Reference from Senator the Hon Ian Campbell, Parliamentary Secretary to the Treasurer to the Corporations and Markets Advisory Committee, 9 July 2002. http://www.camac.gov.au/camac/camac.nsf/print/Directors%27+Duties+and+Personal+Liability+%28July+2002%29?opendocument.

these creditors do not require further compensation ex post. Others, however, are more vulnerable. These include small trade creditors and employees who are not fully compensated by contract, and tort creditors who may have no contract with the company.

Part III looks at the means by which the debts of all types of creditors could be protected against the risk of loss. These include companies holding mandatory debt insurance, mandatory minimum capital requirements, and imposing liability on shareholders, especially on holding companies of insolvent subsidiaries.

Part IV examines the methods by which the creditors noted above as vulnerable could be protected. The means of compensation for them include priority in a winding up, mandatory insurance held by the company and the provision of government assistance schemes, funded either by levies on companies or from taxpayer revenue.

Part V weighs the various methods of creditor recovery and protection against the imposition of liability on directors, and considers the relevance of risk aversion and professional indemnity insurance to director liability. It will be concluded that while lifting the corporate veil on holding companies and some means of creditor protection targeted towards vulnerable creditor groups may be desirable, the imposition of liability on directors where they are at fault is a superior means of protecting creditors from loss.

Creditors' Need for Recovery

Not all creditors are vulnerable to the risk of loss in the event of a company's insolvency. Economic theorists argue that creditors can self protect against the risk of non-payment by the company. This can be done in a number of ways. For example, protection can be afforded by the terms of the contracts they negotiate.[4] Easterbrook noted that '[a]s long as these risks are known, the firm pays for the freedom to engage

[4] Wishart asserted that '[c]reditors charge interest for the service they render. Built into that fee is compensation for the risk of loss they bear. The greater the risk of loss, the more is charged to compensate for that risk. Creditors cannot complain that insolvency as such has caused them loss because they have contracted to bear that risk, and have built compensation for bearing it into the cost of credit. If creditors do not charge for the probability of certain events happening, they should not be supported in their foolishness. They should not survive to charge less than wiser people.' David Wishart, 'Models and Theories of Directors' Duties to Creditors' (1991) 14 *New Zealand Universities Law Review* 323, 336.

in risky activities. … The firm must offer a better risk-return combination to attract investment'.[5]

In addition to the capacity to price protect, some creditors can also be protected by devices such as loan covenants, restricting the company's ability to sell or further pledge its assets, security over the corporation's major assets or personal guarantees from the directors.[6] Other means of protection against loss include the ability of some creditors to diversify their client base so that non payment by one does not lead to their own financial ruin. Additionally, creditors may have short term credit periods, which allow them to carefully assess credit worthiness with current information about the company's financial stability.

This ability to provide self protection would indicate that creditor protection at common law or under statute is unnecessary. However, the self-protection argument is based on the 'efficient markets' hypothesis. This holds that 'all relevant information will be available to the market and that the market will rapidly, if not instantaneously, digest all information as it becomes available'.[7] But even the proponents of this theoretical outlook are prepared to admit that markets do not always work efficiently,[8] because it does not take into account situations where there is not full information about the investment or the borrowing company's financial position. Moreover, the size of the company[9] with which creditors are dealing can have a

---

5 Frank Easterbrook and Daniel Fischel, *The Economic Structure of Corporate Law* (1991) 50. See also Ross Grantham, 'Directors' Duties and Insolvent Companies' (1991) 54 *Modern Law Review* 576, 579. Posner also commented that 'the interest rate on a loan is payment not only for renting capital but also for the risk that the borrower will fail to return it': Richard Posner, 'The Rights of Creditors of Affiliated Corporations' (1976) 43 *University of Chicago Law Review* 499, 501. However, Keay noted research by Cheffins (published in Brian Cheffins, *Company Law: Theory, Structure and Operation* (1997) 501) which establishes that 'there is little evidence that creditors charge a higher interest rate when dealing with a limited liability company, compared to other creditors'. Andrew Keay, 'Directors' Duties to Creditors: Contractarian Concerns Relating to Efficiency and Over-Protection of Creditors' (2003) 66 *Modern Law Review* 665, 689.

6 Posner, 'Affiliated Corporations', above n 5, 504.

7 Jeffery Gordon and Lewis Kornhauser, 'Efficient Markets, Costly Information and Securities Research' (1985) 60 *New York University Law Review* 761, 770.

8 Wishart, 'Models and Theories', above n 4, 336. Also Ross Grantham, 'The Judicial Extension of Directors' Duties to Creditors' [1991] *Journal of Business Law* 1, 2-3.

9 The vast majority of companies in Australia are not listed on the Stock Exchange. As at the 30/6/04 there were approximately 1,309,870 companies in Australia, of which approximately 1,291,110 were proprietary companies and 18,670 were public companies. Approximately 1,400 public companies are listed on the Stock Exchange. Email from Debbie Cowley, Product Team, ASIC, to Helen Anderson 6 December, 2004.

significant impact on their ability to obtain information about solvency, as small closely held companies are more likely to deprive creditors of vital information than larger companies with mandated public disclosure or a board well separated from its shareholders.[10] In addition, directors of a company may choose to take additional risks,[11] which creditors have not foreseen and for which no premium has been charged.[12]

Certainly, not all creditors are able to protect themselves[13] and indeed, the ability of some creditors to protect themselves, for example, with charges over company assets or loan covenants, increases the risk to weaker parties who cannot negotiate such protection.[14] Small trade creditors may lack the knowledge and expertise to make accurate assessments of risk and in any event, would be unable to calculate an appropriate premium to compensate for it. Because the size of their individual debts are comparatively small, the cost of obtaining information about the risk may be prohibitive. They lack information about their fellow trade creditors to enable them as a group to negotiate collectively[15] for fuller particulars of risk.[16] Van der Weide argued

---

10 Wishart, 'Models and Theories', above n 4, 336.

11 Eisenberg noted: 'It is almost impossible to deal adequately with this potential for ex post opportunism by ex ante contracting'. Melvin Eisenberg, 'The Structure of Corporation Law' (1989) 89 *Columbia Law Review* 1461, 1465. See also Mark Byrne, 'An Economic Analysis of Directors' Duties in Favour of Creditors' (1994) 4 *Australian Journal of Corporate Law* 275, 277.

12 Arrow commented that '[i]t is a plausible hypothesis that individuals are unable to recognise that there will be many surprises in the future; in short, as much … evidence tends to confirm, there is a tendency to underestimate uncertainties'. Kenneth Arrow, 'Risk Perception in Psychology and Economics' (1982) 20 *Economic Inquiry* 1, 5. See also Posner, 'Affiliated Corporations', above n 5, 504-5.

13 Lipson labelled these creditors 'low VCE creditors', who 'lack volition, cognition and exit'. This describes creditors who lack voluntariness in their dealings with the company (tort creditors, taxing authorities, terminated employees); lack information (cognition) about the true state of company affairs; and lack the ability to exit from these relationships because of the absence of a market to sell their rights against the company. Jonathan Lipson, 'Directors Duties to Creditors: Power Imbalance and the Financially Distressed Corporation' (2003) 50 *UCLA Law Review* 1189, 1193.

14 Keay, above n 5, 688. Judith Freedman, 'Limited Liability: Large Company Theory and Small Firms' (2000) 63 *Modern Law Review* 317, 351.

15 Reinier Kraakman et al, *The Anatomy of Corporate Law* (1st ed, 2004) 72.

16 The ability to negotiate collectively may explain why employees have been given protection under the law. Part 5.8A of the *Corporations Act 2001* (Cth), hereinafter referred to as the *Corporations Act*, was introduced by the *Corporations Law Amendment (Employee Entitlements) Act* 2000 (Cth). It aims 'to protect the entitlements of a company's employees

that short term creditors 'can quickly respond to bad firm behaviour by taking their business elsewhere'[17] but Keay described this as 'typical of the gross overstatements that pervade some works that have contributed to the law and economics literature'.[18]

Employees are in a particularly vulnerable position, despite being labelled 'voluntary' creditors in the economic literature. In times of high unemployment, employees may be faced with a difficult decision between unemployment and a financially unstable employer. Unlike trade creditors, employees lack the ability to diversify their risk.[19] In some circumstances, employees may be compensated for their lack of ability to diversify by having superior access to information about their employer's financial position[20] and senior employees may be able to seek added remuneration in exchange for running the risks associated with possible financial instability. However, not all employees are in this favourable position, and the ones who are most likely to need the protection of the law are also the ones least likely to be privy to the company's private financial troubles.[21]

The tort claimant,[22] such as a person injured by the negligence of the defendant company, is the most vulnerable of all, when the company is insolvent, uninsured[23] and unable to provide compensation an injured plaintiff. These people have no ability to diversify their risk and no ex ante information about the financial position of the company.

---

from agreements and transactions that are entered into with the intention of defeating the recovery of those entitlements.'

17 Mark Van der Weide, 'Against Fiduciary Duties to Corporate Stakeholders' (1996) 21 *Delaware Journal of Corporate Law* 27, 49.

18 Keay, above n 5, 697.

19 Paul Halpern, Michael Trebilcock and Stuart Turnbull, 'An Economic Analysis of Limited Liability in Corporation Law' (1980) 30 *University of Toronto Law Journal* 117, 149.

20 Ibid 143.

21 See above n 16.

22 Victims of corporate torts are often referred to as 'tort creditors'. However, more correctly, they are tort claimants or even simply victims of tort, as they may not have initiated action and proved their case against the defendant company due to its insolvency.

23 Easterbrook and Fischel argued that economic theories take into account the protection of tort creditors. They contend that if the compensation of tort creditors could affect the financial viability of the company, the directors will ensure that adequate insurance is maintained. Easterbrook and Fischel, 'Economic Structure', above n 5, 52-4.

Protection for the Benefit of All Creditors

The foregoing indicates that vulnerable cohorts of creditors are undercompensated or uncompensated in their dealings with companies ex ante and therefore are arguably entitled to some form of recovery ex post. Even so called 'strong' creditors who have adopted some means of protection against the risk of loss may not be entirely compensated for the risks to which they are exposed. The following discussion will therefore examine methods of protecting creditors in the event of a company's insolvency which are for the benefit of all creditors. These include mandatory minimum capitalisation, imposing liability on shareholders and mandatory debt insurance. It will be seen that while they are, or could be, of some benefit to creditors, none are a satisfactory means of ensuring that creditor protection, particularly of vulnerable creditor groups.

Mandatory Minimum Capital Requirements

Undercapitalisation of a company is a significant problem in insolvency. Sometimes all assets of a company are already charged to secured creditors and there is nothing left for unsecured creditors. Sometimes a company is set up deliberately undercapitalised, perhaps because it is a vehicle by which a holding company conducts risky ventures. Sometimes it is set up simply because its promoter wants the benefit of limited liability.

These different reasons for undercapitalisation will be examined separately. This section will deal with the general question of undercapitalisation, and the next two will deal with the holding company situation and the issue of whether, or in what circumstances, the privilege of limited liability should be removed from shareholders.

The *Corporations Act* already has substantial capital maintenance provisions. The rule in *Trevor v Whitworth* [24] has been incorporated into Part 2J of that Act and regulates share capital reductions and share buy backs, self-acquisition and control of shares, and the provision of financial assistance for the acquisition of shares. The aim of the provisions is to maintain companies' capital because its reduction could prejudice the rights of shareholders and creditors.[25]

---

[24] (1887) 12 App Cas 409.

[25] *Corporations Act* s 256A says ' … The rules [in pt 2J.1] are designed to protect the interests of shareholders and creditors by

(a) addressing the risk of these transactions leading to the company's insolvency …'

However, the law does not prescribe initial minimum capital requirements.[26] Therefore it is possible for a company to be seriously undercapitalised without ever breaching any of the statutory or common law rules relating to the maintenance of share capital.

Easterbrook and Fischel point out that 'the lower a firm's capitalization, the higher the probability that it will engage in excessively risky activities'.[27] Often the risk associated with a debt increases after the debt is incurred.[28] Even if the capitalisation of the company had been checked and was deemed adequate, later borrowing which prejudices creditors' ability to recover may take place where creditors are unable to obtain negative pledges to prevent it.

The report entitled 'Corporate Insolvency Laws: A Stocktake',[29] released in 2004 by the Parliamentary Joint Committee on Corporations and Financial Services, included recommendations about 'assetless' companies. It said:

> The Committee is of the firm belief that the problem of assetless companies must be addressed. It recommends that the Government establish an assetless company registration fund to finance preliminary investigations of breaches of directors' duties and fraudulent conduct using the skills of registered insolvency practitioners.[30]

Further recommendations were that an empirical study of assetless companies be commissioned[31] and that statistics be collated and published.[32] While these measures will help to determine the extent of the problem, clearly more is needed to overcome its detrimental effects.

Nonetheless, it must be recognised that any attempt to impose minimum capital requirements involves costs to the company. These may be in the form of administrative costs, as well as the cost of having capital sit idle or being subject to

---

26 Even listed companies do not have minimum capitalisation requirements to qualify for listing on the Australian Stock Exchange. While one ASX Listing Rule looks at assets (Listing Rule 1.3), an alternative to qualify for listing is the company's annual profit history (Listing Rule 1.2). However, it should be noted that periodic disclosure requirements apply to these companies under Listing Rule 4.1.

27 Easterbrook and Fischel , above n 5, 60.

28 Grantham, 'Judicial Extension' above n 8, 3.

29 Parliamentary Joint Committee on Corporations and Financial Services, Parliament of Australia, *Corporate Insolvency Laws: A Stocktake* 2004.

30 Ibid [7.50].

31 Ibid [7.56].

32 Ibid [7.57].

statutory limits on its use.[33] Such requirements may also interfere with the company's ability to be flexible in its financing, if, for example, maximum debt/equity ratios are mandated. The question of how much capital is adequate is also difficult to ascertain and presumably would alter with the company's growth. Even if sufficient capital were required to be set aside to cover present trade debts and employee entitlements, these sums may not be sufficient to compensate a plaintiff badly injured by the company's negligence.

Holding Company Liability for Debts of an Undercapitalised Subsidiary

The recent James Hardie asbestos claims illustrate the possible extent of tort liability that can arise in the context of the unpaid liabilities of a subsidiary of a solvent holding company.[34] The Report of the Special Commission of Inquiry into the Medical Research and Compensation Foundation[35] stated:

> One of the main public policy objectives of tort law is to discourage activity that is needlessly harmful to people, by imposing the cost of compensation on the wrongdoer. That policy is undermined where wrongdoers can externalise their risk. In reality, it is substantially undermined if a company about to undertake an activity that poses serious health risks for mere bystanders or ultimate consumers can ensure it will never have to satisfy any claims for compensation by the simple technique of carrying out the operations through a company with no capital, funded by loans from the parent secured by a debenture over its assets. Indeed, the limited liability principle as it presently operates actually *encourages* managers so to act, because it is in the 'shareholders' interests' to do so.[36]

---

33 For example, banks are subject to minimum capital requirements, and these are monitored by the Australian Prudential Regulation Authority, which has the responsibility, under Division 2 of the *Banking Act 1959* (Cth), of ensuring the protection of depositors.

34 Spender notes that 'Australian asbestos manufacturers arranged their corporate groups after the dangers of asbestos had already been recognised overseas. Therefore, from its inception, the mining and manufacture of asbestos in Australia tended to be conducted by the undercapitalised subsidiaries of firms'. Peta Spender, 'Blue Asbestos and Golden Eggs: Evaluating Bankruptcy and Class Actions as Just Responses to Mass Tort Liability' (2003) 25 *Sydney Law Review* 223, 234.

35 David Jackson, The Report of the Special Commission of Inquiry into the Medical Research and Compensation Foundation, 2004, *Annexure T The Concept of Limited Liability – Existing Law and Rationale*,< http://www.cabinet.nsw.gov.au/hardie/Volume1.pdf >.

36 Ibid 419-20.

Easterbrook and Fischel described the problem of 'moral hazard' which can exist with groups of companies as follows:

> If limited liability is absolute, a parent can form a subsidiary with minimum capitalisation for the purposes of engaging in risky activities. If things go well, the parent captures the benefits. If things go poorly, the subsidiary declares bankruptcy [to the detriment of its outstanding unsecured creditors] and the parent creates another with the same managers to engage in the same activities. The asymmetry between benefits and cost, if limited liability is absolute, would create incentives to engage in a socially excessive amount of risk activities.[37]

Nolan also commented:

> There may be powerful commercial or fiscal incentives for a parent company to allow one or more group members to operate at a loss, or to deny them economic opportunities or to undercapitalise them, or to adopt integrated financing techniques characterised by the shuffling of assets (particularly funds) and liabilities between them as the need arises. Thus, interest-free or otherwise uncompetitive inter-company loans may be made, or one group member may guarantee another's borrowings without specific regard to the guarantor company's liquidity. Resources may be shifted between companies under the label of 'interest', 'dividends' or 'management fees'. Indeed, as one of the American academics has explained:
>
> > the managers of the enterprise would be acting irrationally if they failed to use the resources of one company to salvage another if such assistance would ultimately enhance the profitability of the corporate enterprise. [38]

---

[37] Frank Easterbrook and Daniel Fischel, 'Limited Liability and the Corporation' (1985) 52 *University of Chicago Law Review* 89, 111. See further Halpern, Trebilcock and Turnbull, above n 19, 124. They suggest that a more appropriate way to avoid moral hazard in groups of companies is to impose liability of the directors, rather than the shareholders of the holding company, at 149.Directors themselves are aware of this conflict. Tomasic and Bottomley report that '[t]he vast majority of Australian directors recognise that the group context of corporate life can and does create significant legal problems for directors'. Roman Tomasic and Steve Bottomley, 'Corporate Governance and the Impact of Legal Obligations on Decision Making in Corporate Australia' (1991) 1 *Australian Journal of Corporate Law* 55, 63.

[38] Anthea Nolan, 'The Position of Unsecured Creditors of Corporate Groups: Towards a Group Responsibility Solution Which Gives Fairness and Equity a Role' (1993) 11 *Company and Securities Law Journal* 461, 485, quoting Jonathan Landers, 'A Unified Approach to Parent, Subsidiary and Affiliate Questions in Bankruptcy' (1975) 42 *University of Chicago Law Review* 589, 648-9.

To overcome this situation, the Harmer Report recommended that the New Zealand model of liability be adopted. [39] This allows for the making of contribution orders[40] and pooling orders[41] against the holding company. Instead, Australia followed the British wrongful trading model,[42] with the addition of a provision for holding company liability, s 588V of the *Corporations Act*.[43] As with the liability for insolvent trading of directors under s 588G of the *Corporations Act*, s 588V of that Act only provides for liability by a holding company with respect to the incurring of contract debts. Tort liabilities are ignored.

The lifting of the corporate veil[44] to impose liability on the holding company which is, after all, a shareholder in the subsidiary, is doubtless justified by the frequent reality of control of the subsidiary by the holding company. In doing so, the legislature has recognised that companies can avoid liability for debts by incorporating undercapitalised subsidiaries which can trade into insolvency without damaging the

---

39 Australian Law Reform Commission, *General Insolvency Inquiry*, Report No 45 (1988) (The Harmer Report) [336] and [857] (Vol 1), D13 and PR9 (Vol 2). See further John Farrar, *Corporate Governance in Australia and New Zealand* (2001) 240.

40 A contribution order requires a solvent company in a group to contribute to the debts of an insolvent company within the group.

41 A pooling order requires the assets of the group to be pooled for the benefit of the creditors of companies within the group.

42 *Insolvency Act 1986* (UK) s 214.

43 Nolan commented that 'The pooling order recommendation appears to have been forgotten entirely even though the Harmer Committee "received no submissions opposing this proposal" [Harmer Report, above n 39, [857]]. The contribution order recommendation has been recast as an insolvent trading provision … In the Explanatory Memorandum to the Corporate Law Reform Bill ... it is claimed that this provision implements "the Harmer Report's recommendations in relation to available assets" even though … the Harmer Report's recommendations are quite different to those enacted under the Reform Act. At most, the Government has simply accepted the *philosophy* underlying the Harmer Reports' contribution order proposal'. Nolan, above n 38, 464-5 (emphasis in original).

44 The Companies and Securities Advisory Committee, Parliament of Australia, *Corporate Groups Final Report* (CASAC Report) (2000) in Chapter Six: Liquidation of Group Companies examines a number of ways in which holding companies can be held liable. These include as shadow directors, at [6.15], liability for misfeasance, including liability for knowingly assisting another company to breach its fiduciary duty under the doctrine of *Barnes v Addy* (1874) LR 9 Ch App 244, at [6.16], agency, at [6.17], letters of comfort, at [6.18] and a variety of situations where transactions may be set aside, at [6.19] to [6.25].

parent company and that sometimes the subsidiary is incorporated for the very purpose of hiding the true and parlous financial state of the holding company.[45]

Rogers CJ in *Qintex Australia Finance Ltd v Schroders Australia Ltd* [46]said:

> It may be desirable for Parliament to consider whether this distinction between the law and commercial practice should be maintained. … there is a great deal to be said for the suggestion … that assets and liabilities of the parent and the subsidiaries should be aggregated. It may be argued that there is justification for preserving the same attitude in relation to the demised companies as was displayed during their active commercial life.[47]

Ramsay[48] also considered that holding company liability for the subsidiary's debts may be appropriate. He adopted a 'law and economics' perspective on the section, asking first, whether it 'creates incentives for individuals and firms to behave efficiently'[49] and secondly, whether the holding company is the best party to bear the risk of the subsidiary's insolvency.[50] He found these economic criteria substantially satisfied.[51]

Nonetheless, the imposition of liability on a shareholder poses a major theoretical dilemma for corporate law. Section 588G of the *Corporations Act*, which provides for liability of a director for the insolvent trading of a company, does not consider the shareholding of the director. A human shareholder with a large shareholding will frequently but not necessarily be represented on the board of directors, but it is the directorship rather than the shareholding that is the basis of insolvent trading liability. Under s 588V of that Act, on the other hand, it is the holding company rather than its directors which is liable for the insolvent trading.

---

45 Ian Ramsay, 'Holding Company Liability for the Debts of an Insolvent Subsidiary: A Law and Economics Perspective' (1994) 17 *University of New South Wales Law Journal* 520, 525-6. Ramsay noted that there are also legitimate business reasons to incorporate subsidiaries, at 533-4.

46 [1991] 3 ACSR 267, 269.

47 See also Robert Baxt, 'The Duties of Directors of Public Companies – the Realities of Commercial Life, the Contradictions of the Law and the Need for Reform' (1976) 4 *Australian Business Law Review* 289; Robert Baxt, 'Duties to a Corporate Group – One Step Forward or Two Steps Backwards?' (1994) 22 *Australian Business Law Review* 138.

48 Ramsay, above n 45.

49 Ibid 521.

50 Ibid.

51 Ibid.

In no other area of the law does a shareholder bear liability for the acts of a company unless that person is also a director of the company. Removing limited liability for any group of shareholders would appear to undermine the rationale of incorporation, as outlined by theorists such as Easterbrook and Fischel,[52] and Halpern, Trebilcock and Turnbull.[53] Limited liability enables shareholders to avoid excessive monitoring costs, both of the company and their fellow shareholders, and thus encourages a wide diversification of investment to both new and existing businesses. Limited liability allows the free transfer of shares which also promotes wide diversification of shareholding. All of these are vital to the continuation and expansion of a capitalist economy.[54] Parliament's lifting of the corporate veil is inconsistent with the reification of the company as a legal entity separate from the directors of that company, and is also highly inconsistent with the approach shown by courts in the common law actions.

Farrar expressed reservations about the consequences of placing liability on a holding company:

> These provisions represent a limited piercing of the corporate veil but will give rise to concern where the holding company is not the principal operating entity. The risk of liability of the holding company will be a cause for concern amongst the boards of large corporation, who will be required to monitor the affairs of subsidiaries more closely. Where the ultimate holding company is in an overseas jurisdiction there may be practical difficulties with doing this.
>
> The practical implications for directors are twofold. The first is to pursue a strategy of 'ignorance is bliss' regarding the affairs of the subsidiary. Do not share directors or employees or information, if this is at all practicable. The second is to avail itself of one of the defences (eg reasonable grounds to expect solvency and reliance on information from a competent and reliable person.)[55]
>
> A more elaborate reform of enterprise liability based on some concept of group legal personality or *automatic* group responsibility would probably create as many problems as it would solve. Since we are currently seeking to escape from the straightjacket of separate legal personality, it seems a mistake to seek refuge

---

52 Easterbrook and Fischel, 'Limited Liability' above n 37; also Easterbrook and Fischel, '*Economic Structure*' above n 5.

53 Halpern, Trebilcock, and Turnbull, above n 19.

54 Easterbrook and Fischel, 'Economic Structure', above n 5, 41-4.

55 John Farrar, 'Legal Issues Involving Corporate Groups' (1998) 16 *Company and Securities Law Journal* 184, 192.

> in a larger concept of group legal personality or responsibility, enticing though this may be because of its apparent reflection of commercial reality.[56]

In 2000, the Corporate Groups Final Report (CASAC Report) looked at the issues arising from corporate groups closely. Unlike the Harmer Report, the CASAC Report did not endorse the introduction of contribution orders[57] but did recommend that liquidators be permitted 'to pool the unsecured assets, and the liabilities, of two or more group companies in liquidation with the prior approval of all unsecured creditors of those companies.'[58] It also recommended that the law should give courts the power to make pooling orders in the liquidation of two or more companies.[59]

However, the Report had noted that imposing liability on the parent company has disadvantages:

> … the value of recovery rights of unsecured creditors of a parent company could be diminished through that company's exposure to the financial risks of all the group's activities. At the same time, creditors of controlled group companies could gain a windfall from the parent company providing, in effect, a form of liability insurance against any defaults by its group companies.[60]

---

56 Ibid 201. Farrar also notes a number of British cases which adopt the two different approaches, ibid 185-6. The more flexible approach was taken by *DHN Food Distributors Ltd v Tower Hamlets LBC* [1976] 1 WLR 852; *Aiglon Ltd and L'Aiglon SA v Gau Shan Co Ltd* [1993] BCLC 1,321; *TSB Private Bank International SA v Chabra* [1992] 1 WLR 231. The traditional separate legal entity approach was taken by *Adams v Cape Industries plc* [1990] Ch 433; *Woolfson v Strathclyde Regional Council* (1978) 38 P & CR 521. See also Farrar, 'Corporate Governance' above n 39, 229. Baxt and Lane also note the English High Court decision in *Re Polly Peck International plc* [1996] 2 All ER 433 which upheld the strict legal position rather than recognising the economic reality of the companies in a group. Robert Baxt and Timothy Lane, 'Developments in Relation to Corporate Groups and the Responsibilities of Directors – Some Insights and New Directions' (1998) 16 *Company and Securities Law Journal* 628, 633. See also John Farrar, 'Frankenstein Incorporated or Fools' Parliament? Revisiting the Concept of the Corporation in Corporate Governance' (1998) 10 *Bond Law Review* 142, 158-160.

57 CASAC Report, above n 44, Recommendation 21, [6.59].

58 Ibid [6.85]. Note the confirmation of the ability of liquidators to pool assets in the recent decision in *Tayeh & De Vries; re The Black Stump Enterprises Pty Ltd* [2005] NSWSC 475. Barrett J held that liquidators under a creditors' voluntary winding up who wished to pool the assets of companies in a group for the purpose of distribution to the creditors of all of the companies must obtain the unanimous consent of all creditors prior to the court permitting the pooling of those assets under s 511(1)(a).

59 Ibid [6.97].

60 Ibid [1.64].

Therefore, the Report suggested that the judicial power to make pooling orders must 'make clear that pooling orders do not affect the rights of external secured creditors' and to 'permit individual external creditors to apply to have a pooling order adjusted to take their particular circumstances into account.'[61]

Fridman also believed that group liability would cause more problems than it would solve.[62] He suggested that the oppressive conduct remedy [63] is a more appropriate way to deal with abuses of the separate legal entity rule by companies in a group. In discussing the judgment in *Qintex Australia Finance Ltd v Schroders Australia Ltd* [64] he said that '[c]learly the problem of identifying which member of a corporate group should be responsible for a given obligation should not be resolved by destroying the notion of corporate personality that has been enshrined in the law since *Salomon v Salomon* …'.[65]

Nonetheless, creditors taking action against companies in a group face considerable difficulties. One is information asymmetry. Creditors frequently are not sure of the particular entity with which they are contracting, and rely on the credit worthiness and reputation of the group as a whole. Transactions within the group can increase the risk without the knowledge of creditors. The Harmer Report considered '[t]here is little sense in promoting a law that is at odds with commercial reality'.[66] It found that often it is 'difficult or impossible to reconstruct the financial position of a company at or about a particular time'[67] because a liquidator has to 'cope with either inadequate or meaningless company accounting records or no accounting records at all'.[68]

The foregoing discussion largely focuses on the arguments for and against holding company liability for the unpaid *contract* liabilities of its subsidiaries. Ramsay highlighted the frequently overlooked inadequacy of the current law by saying '[t]he section is seriously deficient in that it provides no protection for tort claimants of

---

61 Ibid [6.97]. The CASAC Report also considered whether courts should be permitted to subordinate intra-group claims in the insolvency of a group company, but rejected the idea, at [6.112]. This was also considered by John Farrar, 'Corporate Group Insolvencies, Reform and the United States Experience' (2000) 8 *Insolvency Law Journal* 148, 153.

62 Saul Fridman, 'Removal of the Corporate Veil: Suggestions for Law Reform in Qintex Australia Finance Ltd v Schroders Australia Ltd' [1991] *Australian Business Law Review* 211.

63 Currently found in *Corporations Act* s 232.

64 [1991] 3 ACSR 267.

65 Fridman, above n 62, 213.

66 Harmer, above n 39, [33].

67 Ibid [297].

68 Ibid [290].

insolvent subsidiaries'.[69] These creditors are particularly vulnerable given that the holding company may deliberately incorporate an undercapitalised subsidiary to conduct dangerous activities which may lead to tort claims.[70]

In some cases, the holding company may be held liable in tort because of its own breach of duty to the plaintiff. This occurs where the holding company has such a degree of control over the activities of the subsidiary that it is in effect conducting those activities itself.[71] Where there is no evidence that the subsidiary is a mere façade, however, courts have held that the fact that a parent company exercises control and influence over its subsidiary does not of itself justify lifting the corporate veil.[72]

Yet despite the fact that most holding companies will not attract liability for their subsidiary's tortious actions, and that creditors of their undercapitalised subsidiaries will remain uncompensated, CASAC's *Corporate Groups Final Report*[73] concluded that:

> The introduction of a general tort liability for parent companies in corporate groups is undesirable, as:
>
> > this liability would undermine the separate legal entity principle and could have negative consequences for the economy
> >
> > this area should be dealt with by specific legislation where the extension of liability beyond the tortfeasor company is desirable in the public interest.[74]

Shareholder Liability

Any discussion of the removal of the privilege of limited liability from shareholders necessarily involves a consideration of economic theory. As noted above, the concept of limited liability for shareholders is central to the purpose of incorporation. However, as the foregoing section has shown, the holding company as a shareholder

---

69 Ibid. A noteworthy example of this is *Briggs v James Hardie & Co Pty Ltd* (1989) 7 ACLC 841.

70 Ramsay, above n 45, 542.

71 *CSR Ltd v Young* (1998) Aust Torts Reports [81-468], 64,952; *CSR Ltd v Wren* (1998) Aust Torts Reports [81-461].

72 *James Hardie & Co Pty Ltd v Hall* (1998) 43 NSWLR 554, 579-84. Rogers AJA in *Briggs v James Hardie & Co Pty Ltd* remarked: 'Rare indeed is the subsidiary that is allowed to run its own race.' (1989) 7 ACLC 841, 858.

73 Above n 44.

74 Ibid [4.20].

can be stripped of the benefit of limited liability where parliament considers it to be warranted.

The question here is whether the privilege of limited liability should be withdrawn from the shareholders of small, undercapitalised companies,[75] where the temptation to use the corporate form to defeat the claims of creditors is most prevalent. The shareholders of these companies are the beneficiaries of the risky behaviour of the director if it yields the desired results, and the directors of such companies are frequently the principal shareholders. Therefore their risky behaviour is not curbed by the fear of losing their positions or the threat of actions for breach of fiduciary duty. They may undercapitalise the company so that their losses as shareholders are minimised, while financing the business by means of debt, sometimes in the form of secured loans from themselves and others.

This situation leaves creditors of these companies particularly exposed if the company becomes insolvent. Of course, if a sole trader becomes bankrupt, the same situation may occur, especially if they transfers their main assets to another person or entity for safekeeping. Nonetheless, but it is easier for a company to exist with shareholder equity of $2 than it is for a sole trader.

Freedman[76] suggested three tests for evaluating the efficacy of limited liability and therefore the desirability of retaining it in a given situation:

> First, does limited liability allocate risk to those most capable of bearing it? Secondly, does it result in optimal levels of risk taking, ensuring that ventures with a net positive value to society, but not others, are undertaken? Thirdly, does it reduce transaction and monitoring costs? In [economic analysis] literature, these measures of 'efficiency' operate within an overall framework of profit maximisation.[77]

Under Freedman's first test, it is not possible to generalise that creditors are more capable than shareholders of bearing the risk of loss. For example, limited liability is not justified for shareholders where the claims being denied are those of the uncompensated and undercompensated creditors identified by this article as deserving recovery. Employees, tort claimants and small trade creditors suffering from contractual powerlessness or information asymmetry are no more capable of bearing the risk of loss than shareholders are.

---

75 Reinier Kraakman, 'Corporate Liability Strategies and the Costs of Legal Controls' (1984) 93 *Yale Law Journal* 857, 873. Also Robert Clark, 'The Duties of the Corporate Debtor to Its Creditors' (1977) 90 *Harvard Law Review* 505, 547-50.

76 Freedman, 'Limited Liability', above n 14.

77 Ibid 319.

Under the second test, limited liability may encourage shareholders acting in their capacity as directors to take excessive risks which bring a net positive benefit to themselves if no adverse consequences result but not to society if the risks do eventuate.[78]

Under the third test, while limited liability may reduce transaction and monitoring costs for shareholders, it increases them for creditors, who have far less power to demand information than members of the company do. Limited liability is even less justified under this test for closely held companies. Since the shareholders are likely also to hold the company's directorships, monitoring costs for shareholders are minimal. Limited liability increases those costs for creditors, particularly of small closely held companies, where information is unlikely to be publicly available.

However, if limited liability were to be removed[79] because it is considered to be unjustified in certain circumstances, such as in the case of undercompensated trade creditors and employees, it would require legislation to do so. It cannot be done by individual contract as it is only the strong creditor who will have the economic power to negotiate a reversal of the limited liability default, such as by the obtaining of personal guarantees from directors. It is beneficial for the company to make this type of arrangement with large creditors as it reduces the risk, and therefore cost, of their transaction. Large creditors are kept happy at the expense of small ones.

Most small trade creditors and employees have little or no bargaining power to insist on such terms, if they wish to continue to do business with the company. Obviously tort creditors[80] are also unable to negotiate ex ante a state of unlimited liability. In

---

78 Leebron commented that 'an investment may be undertaken even though from society's point of view it is not worthwhile. In addition, the full risk of the enterprise will not be reflected in the required rate of return. The tort victim, or society at large, may be quite averse to the prospect of the catastrophic loss. The purely rational investor, however, will continue to regard the enterprise as being only moderately risky since the worst possible outcome is the loss of the investment'. David Leebron, 'Limited Liability, Tort Victims and Creditors' (1991) 91 *Columbia Law Review* 1565, 1585.

79 The justification for the removal by the state of limited liability lies in the 'fictive entity' theory, also known as the 'concession' theory. This sees the corporation as a concession of the government, and thus subject to government regulation. 'What the state gives … the state can take away'. David Cohen, 'Theories of the Corporation and the Limited Liability Company: How Should Courts and Legislatures Articulate rules for Piercing the Veil. Fiduciary Responsibility and Securities Regulation for the Limited Liability Company' (1998) 51 *Oklahoma Law Review* 427, 435.

80 Macey suggested that the position of tort victims can be protected by 'common law judicial craftsmanship', that is, by lifting the corporate veil. Jonathan R Macey, 'The Limited Liability Company: Lessons for Corporate Law' (1995) 73 *Washington University Law*

addition, such individual contracts involve costs, which the small trade creditors are least able to afford. Freedman noted:

> Limited liability with partial reversal through contract does not seem to ensure allocation of risk to those most capable of bearing it. … Nor does the risk shift onto superior risk bearers. Small creditors least able to monitor and assess risk and to contract out of limited liability may in fact pick up any remaining losses. It is mainly the sophisticated creditor with bargaining power who seems to gain.[81]

She concluded:

> Other values must also be weighed in the balance. … These underlying considerations, reflecting society's values, need to be exposed and discussed in order to ensure that legal policy does properly reflect moral and political criteria. … There may come a point at which we are prepared to choose certain principles, such as fairness, over and above profit maximisation.[82]

Therefore, contracting around limited liability does not appear to be a feasible way of ensuring creditor protection. Another way to achieve this would be to remove limited liability for all small companies. However, this presents its own difficulties. Defining companies as small is problematic.[83] Such entities are unlikely to grow if unlimited personal liability makes shareholders reluctant to delegate management to a board of directors because it would involve them incurring substantial monitoring costs.[84] New enterprises may struggle to attract capital for innovative, possibly worthwhile ventures.

In 1989, the *Close Corporations Act 1989* (Cth) was passed.[85] Its purpose was to reduce financial and reporting costs by simplifying the corporate law rules for small businesses. The maximum membership was ten natural persons.[86]

Quarterly 433, 449. In reply to this notion, Freedman commented that lifting the corporate veil 'is both uncertain and cumbersome, and operates only in extreme cases.' Freedman, 'Limited Liability', above n 14, 350.

81 Freedman, 'Limited Liability', above n 14, 332.

82 Ibid 319-20.

83 Judith Freedman, 'Small Businesses and the Corporate Form: Burden or Privilege' (1994) 57 *Modern Law Review* 555, 566-71.

84 Easterbrook and Fischel, 'Economic Structure', above n 5, 60.

85 It was based on the recommendations of the Companies and Securities Law Review Committee, Parliament of Australia, *Report to the Ministerial Council on Forms of Legal Organisation for Small Business Enterprises*, 1985.

86 Close Corporations Act 1989 (Cth) s 60.

Under this Act, the members were exempt from liability for company debts except for certain specified situations.[87] One of these was where the company unreasonably delays initiating winding up proceedings.[88] Another was where the company acquires its own shares when there were not reasonable grounds for the required declaration of solvency.[89] However, the Act was never proclaimed. Since then the Corporations Law Simplification Program Task Force has released a report entitled 'Small Business Proposal to Simplify Proprietary Companies'[90] but it did not take up the ideas of the *Close Corporations Act 1989* (Cth).

Mandatory Debt Insurance

Another possible way to protect creditors in general is by forcing companies to obtain insurance for their creditors against losses.[91] While some companies might chose to insure without compulsion, there is little incentive[92] to do so on the part of the company's owners. They would see the company's undercapitalisation and their own limited liability as a cheaper and more effective means of 'insurance' against claims.

Mandatory professional indemnity insurance has recently been adopted as a condition of allowing auditors to incorporate. The *Corporate Law Economic Reform Program (Audit Reform and Corporate Disclosure) Act 2004* (Cth) introduced provisions into the *Corporations Act*[93] permitting auditors to incorporate and be registered as authorised audit companies, subject to a number of conditions. One of these is the obtaining of 'adequate and appropriate professional indemnity insurance for claims that may be made against the company in relation to the audit of companies'.[94]

---

87 Close Corporations Act 1989 (Cth) s 106.

88 Close Corporations Act 1989 (Cth) s 110.

89 Close Corporations Act 1989 (Cth) s 111.

90 Corporations Law Simplification Program Task Force, Parliament of Australia, *Small Business Proposal to Simplify Proprietary Companies*, 1994.

91 Keay, above n 5, 696.

92 The greater the perceived risk, the greater the premium charged by 'powerful' creditors with the ability to do so, such as banks. Insurance would reduce both the risk and therefore the cost of credit provided by these creditors. However, as noted above, not all creditors have the ability to charge such a premium, so insurance against losses by *all* creditors is an inefficient way to reduce the cost of credit provided by 'powerful' creditors. A cheaper and more efficient way is to provide security to those creditors by way of charges over company property or personal guarantees by directors.

93 *Corporations Act* pt 9.2.

94 *Corporations Act* s 1299B.

What amounts to 'adequate and appropriate insurance' was outlined in the ASIC Policy Proposal which preceded the *Corporate Law Economic Reform Program (Audit Reform and Corporate Disclosure) Act 2004* (Cth).[95] It was suggested that where the maximum engagement fee is less than $50,000, the insured amount be $500,000; where the maximum is over $50,000, the insurance must be at least ten times the estimated maximum engagement fee, up to a maximum of $20 million.[96]

However there are a numbers of reasons why mandatory insurance against creditor losses by companies in general may be hard to implement. Setting an appropriate amount of insurance is not as easy for other types of business as it is for professional services businesses. In addition, auditors are willing to undertake adequate insurance as the price of their right to incorporate; other businesses take the right to incorporate for granted.

The cost of the insurance is also clearly significant. It will be seen in the discussion of mandatory insurance for employee entitlements below that industry objected vigorously to the cost of the proposal. The cost for general creditor insurance would be considerably higher than insurance purely for employee entitlements. This cost would have to be socialised, that is, passed on to the public by way of increased costs for goods and services.[97] As a result, some small traders could be forced out of the market.[98] If the cost of the insurance impacted upon dividends to shareholders, it could lead to a withdrawal of capital from the market, which would be highly detrimental for economic activity.

In addition, debt insurance might encourage unnecessarily risky behaviour.[99] This could result in higher claims, leading to either higher insurance premiums where the insurer was aware of the additional risk, or else exposing the insurer to losses, where

---

95 ASIC Policy Proposal CLERP 9 Bill *Authorised Audit Companies: Insurance Arrangements*, June 2004. This was replaced by ASIC on 21 December, 2004 with PS 180 *Auditor Registration*. PS180.32 deals with the adequacy and quantum of professional indemnity insurance, which is a condition of registration for authorised audit companies.

96 Ibid PS180.33.

97 Many auditors would currently hold professional indemnity insurance, so the cost of their insurance is already passed on to their clients and socialised in the cost of their client's goods and services.

98 Easterbrook and Fischel, *Economic Structure*, above n 5, 60.

99 See further Keay, above n 5, 685. It will be noted below that insurance held by directors does not encourage risky behaviour because directors have reputational disincentives against improper behaviour leading to claims against them. Debt insurance, held by the company, does not carry such disincentives because improper behaviour by the director is not the trigger for the insurance claim to be paid.

it was unaware.[100] Small businesses, especially in risky areas such as new technologies, may have difficulty obtaining insurance.[101] This could stifle innovation and entrepreneurship. Fixing the appropriate level of insurance and ensuring that it is obtained are also problematic. The issue of insurance of directors against their liability will be discussed below.

Other Methods of Protection of Creditors

A brief note should be made of other schemes to protect creditors. There have been many initiatives to improve and simplify corporate law in Australia and to make it safer for shareholders and creditors. The current CLERP program was preceded by the Corporations Law Simplification Program.[102] ASIC's National Insolvency Co-ordination Unit (NICU) initiated a National Insolvent Trading Program in July 2003 to review companies suspected of trading while insolvent.[103] ASIC staff conduct surveillance visits, which can result in the appointment of voluntary administrator or liquidator.

Federal government financial assistance in the form of social security is also available for those who are unemployed or ill.[104] While these socialise to some extent the cost of loss of employment and injury, they do not ensure adequate compensation for those who lose their employment entitlements or suffer injury due to the insolvency or negligence of companies. In addition, they do not provide any payments whatsoever for trade creditors whose debts are not paid by insolvent companies or for tort creditors who suffer no personal injury.

---

100 Halpern et al pointed out that information for monitoring is costly and that 'the insurance payment is typically independent of the actions of the insured'. Halpern, Trebilcock and Turnbull, above n 19, 140.

101 Freedman also noted that small businesses with no track record may have difficulties obtaining insurance at an affordable price. These are the businesses which are most likely to fail, yet the economy relies on the incorporation of new, small businesses. Freedman 'Limited Liability' above n 14, 340.

102 Parliament of Australia, Attorney-General's Department, *Corporations Law Simplification Program*, Simplification Task Force, 1994 – 1996.

103 ASIC, 'Court Decision a First for ASIC's Insolvent Trading Program' (Press Release, 3 February 2004).

104 *Unemployment and Sickness Benefits Act 1944* (Cth).

Protection for the Benefit of Specific Groups of Creditors

The forms of protection described above are ones which were or could be available for the benefit of creditors generally. What follows is a discussion of those types of measures which could be targeted for the benefit of creditors identified as vulnerable. They include Government initiatives and actions available to certain groups of creditors under the Act. In addition, the issue of priority distributions by the liquidator will be examined. The aim of this discussion is to evaluate the adequacy of the methods by which vulnerable creditors can be protected in the wake of a company's insolvency.

The justification for studying targeted solutions is twofold. First, as this article has maintained, certain creditor groups, in particular tort creditors, are not given sufficient compensation ex ante for the risk of non-payment by an insolvent company. Therefore, it is possible that the best way to protect them is to devise remedies expressly for them.[105] Secondly, creditor rights of recovery are not costless – they are at the expense of other parties.[106] Where this other party is the director of the company, the burden of potential liability may arguably have adverse economic effects on the entrepreneurial function of the director. This will be examined further in Part V. Therefore, a targeted solution may be beneficial if it ensures acceptable levels of recovery for vulnerable creditors, while minimising levels of liability on other parties.

Small Trade Creditors

It was noted above that small trade creditors may be undercompensated by the terms of their contracts with companies, and therefore are arguably deserving of recovery in the event of the company's inability to pay. However, it is difficult to categorise ex ante those creditors who are undercompensated, and therefore to devise rules which ensure their compensation ex post. This is so, even if the trade creditors are particularly disadvantaged in terms of price protection, information asymmetry and inability to diversify. To determine the extent of their disadvantage would be both a complex question and a costly one. Therefore they are treated *pari passu* as unsecured creditors and recover from the liquidator after the distribution of preferential

---

105 Clearly, some specific remedies already exist. Employees are given special protection under *Corporations Act* pt 5.8A, outlined below.

106 Lipson, above n 13, 1251. See also Rizwaan Mokal, 'Priority as Pathology: The *Pari Passu* Myth' (2001) 60 *Cambridge Law Journal* 581, 609.

payments pursuant to s 556(1) of the *Corporations Act*.[107] As a result, they do not enjoy any priority in winding up or other special privileges.

Employees

Employees are the most well treated of the undercompensated groups of creditors. Possibly this is because they have powerful political representation by trade unions, and because business closures and job losses are frequently newsworthy.

Few employees, with the exception of senior management, have the ability to price protect against the risk that their entitlements will not be paid, or to diversify away the risk of non-payment by holding multiple jobs. However, currently, employees enjoy a degree of priority in the distribution of the assets of a company when it is wound up. Section 556(1)(e) of the *Corporations Act* provides priority for wages and superannuation contributions of employees, with limits applicable to directors and their spouses.[108] Leave entitlements and retrenchment payments are covered by s 556(1)(g) and (h) of the Act respectively.

The Harmer Report noted that employee priority 'was first introduced into insolvency legislation for social welfare reasons to ease the financial hardship caused to a relatively poor and defenceless section of the community by the insolvency of their employer'[109] but that 'the principal rationale for the employee priority has been significantly diminished by the development of a sophisticated social welfare system.'[110]

However, it is likely that the lowest priorities relate to 'non-distribution' as there is rarely enough funds to satisfy the highest categories of priority payment, each of

---

107 Mokal commented that *pari passu* 'is best seen as a fall-back provision. It is the rule which takes over when it would be pointless to provide any other. … Recall that most insolvency proceedings (75% of them or more) yield nothing for general unsecured creditors. And when they do bring some returns, the yields are fairly small (about 7 pence in the pound on average). So there simply is no point in deciding how *these* claims should rank vis a vis each other. For such claims to be governed by the pari passu rule makes very good sense, since the cost in terms of time, effort and resources required to determine *their* appropriate (fair and efficient) rankings would far exceed any benefits'. Mokal, above n 106, 613 (emphasis in original). These statistics are from the United Kingdom.

108 *Corporations Act* s 556(1A) refers to excludes employees, which is defined in s 556(2) of that Act.

109 Harmer, above n 39, [721].

110 Ibid [722].

which must be paid in full before later categories receive anything.[111] Overcompensation of employees as a result of priority in a winding up is therefore unlikely to be a problem.

Another issue, raised by the Harmer Report, is whether giving preference to employees deprives other worthy creditors.

> Further, the effect of the priority is to deprive other unsecured creditors of their claim to a share of the available assets. Included in that class of unsecured creditors may be small traders who were substantially dependent upon the insolvent for their business and persons who were in an employee-like relationship with the insolvent but who are classified (in a strict legal sense) as independent contractors. These creditors may be as vulnerable as employees in the event of bankruptcy or liquidation but enjoy no protection. [112]

In particular, those with claims against the company for compensation for injury rank behind employees.[113] This issue will be discussed further in the next section.

Another point to note is that the priority of employees in a winding up does not extend to other forms of insolvency administration. Taylor[114] noted that

> [i]n liquidation, the legislative priority is fairly exhaustive, but in controllership it is less so, and in the two remaining administrations under Part 5.3A of the *Corporations Law* [voluntary administration and deeds of company arrangement] it is or may be almost non-existent.'[115]

Under Part 5.2 of the *Corporations Act*, when the company is in receivership or subject to other controllership, employees have a degree of priority under s 433(3)(c) of that Act.[116] One of the recommendations of the recent report entitled *Corporate Insolvency Laws: A Stocktake* 2004 was 'that the law be amended to make it mandatory for a deed of company arrangement to preserve the priority available to creditors in a winding up under s 556(1) unless affected creditors agree to waive their priority.'[117]

---

111 Rizwaan Mokal, 'On Fairness and Efficiency' (2003) 66 *Modern Law Review* 452, 459.

112 Harmer, above n 39, [722].

113 *Corporations Act* s 556(1)(f).

114 Terry Taylor, 'Employee Entitlements in Corporate Insolvency Administrations' (2000) 8 *Insolvency Law Journal* 32.

115 Ibid 34.

116 This section only applies where the receiver is appointed to act on behalf of holders of debentures that are secured by a floating charge, not a fixed charge: *Corporations Act* s 433(2).

117 Stocktake, above n 29, [11.20].

In 2001, the Federal Government proposed[118] that employee entitlements be a maximum priority and that they rank ahead of secured creditors. The fact that this proposal was made appears to provide evidence that employee entitlements are not fully paid out under the current level of winding up priority, and that the joint initiatives of 2000 were not sufficient.[119] Despite strong support from the trade union movement and others, criticisms of the proposal were expressed to the Parliamentary Joint Committee.[120] Reasons included the uncertainty the proposal would have on the cost and administration of secured lending, the complexity it would cause during administrations and the incentives for avoidance by secured creditors.

As a result of these criticisms, the Stocktake Report concluded:

> The Committee recommends that the maximum priority proposal not be adopted. The emphasis in any reform proposals in relation to employee entitlements should be on preventative measures to minimise the risk of loss of employee entitlements and modifying current behaviour to ensure directors and managers of companies take greater responsibility in meeting the cost of employee entitlements in the event of business failure.[121]

The scope of this proposed action is unclear; nonetheless, it is plain that the Report considered that the present situation is unsatisfactory and that some action is warranted. The Report said that

> [t]he committee considers that the protection of employee entitlements in the circumstances of employer insolvency is an important public policy and it is appropriate for governments to explore options for better protecting employee entitlements.[122]

Therefore it is clear that priority in winding up is insufficient to ensure the adequate protection of employee entitlements.

A national insurance scheme to protect employee entitlements on the event of liquidation was mooted in 1999.[123] In a survey by insolvency firm Benfield Greig,

---

118 It was announced by the Prime Minister at a press conference on 14 September, 2001, and reiterated in the Government's November 2001 election policy statement entitled 'Choice and Reward in a Changing Workplace': Stocktake, above n 29, [10.29].

119 These are *Corporations Act* pt 5.8A and the Employee Entitlements Support Scheme, discussed below.

120 Stocktake, above n 29, [10.33] – [10.51].

121 Ibid [10.55].

122 Ibid [10.53].

123 Commonwealth of Australia, *The Protection of Employee Entitlements in the Event of Employer Insolvency*, Ministerial Discussion Paper (1999). The Harmer Report, above n 39, also

commissioned by the New South Wales Government,[124] the authors observed that annual losses of entitlements could be up to $464 million.[125] It was suggested that the scheme be funded from a levy on businesses calculated in accordance with their wages bill, similar to workers' compensation, except where businesses could prove that they had provided protection for employee entitlements. However, the proposal was opposed by industry groups and was abandoned in favour of the Employee Entitlements Support Scheme (EESS), discussed below.

An insurance scheme funded by employers has certain advantages. The cost to the taxpayer is minimised. It also provides an incentive to business to reduce the risk of loss of employee entitlements. Many countries have adopted such a scheme.[126] However, the scheme as proposed in Australia was not without its problems. To ease the cost burden to small businesses (defined as those with less than 20 employees), they would be exempt from the proposal, with a separate government-funded safety net provided for their employees. Hammond[127] noted research by the ACTU which placed the average number of employees per company in 1998 as 8.75, meaning that a great many businesses would be exempt under the scheme.[128] In addition, the differential treatment of small and other businesses would add a further layer of administration and complexity for both businesses and employees affected by the scheme, especially for those businesses close to the employee limit. The scheme assumed that the insurance industry would have the capacity to absorb whatever losses occurred from business failures. In times of recession, this may not be realistic.

Trust funds established by specific industries funded by levies on employers are another alternative frequently contemplated when compulsory insurance is analysed. The Stocktake report concluded that:

---

suggested a wage earner protection fund, at [727]. See further Celia Hammond, 'Insolvent Companies and Employees: The Government's Year 2000 Solutions' (2000) 8 *Insolvency Law Journal* 86, 88.

124 'National Insurance Scheme to Protect Employee Entitlements: Preliminary Feasibility Study', noted in Steve O'Neill and Bronwen Shepherd, *Corporate Insolvencies and Workers' Entitlements* (2002) Parliament of Australia Parliamentary Library <http://www.aph.gov.au/library/intguide/econ/insolvencies.htm>,(date of retrieval) 6.

125 More recent estimates of potential losses put the figure much higher. Institute of Actuaries of Australia, *Protection of Employee Entitlements in the Event of Employer Insolvency*, 2003, 18.

126 Argentina, Austria, Belgium, Canada, Denmark, Finland, Greece, Germany, Ireland, Italy, Israel, Japan, Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, United Kingdom and Oregon in the USA.

127 Hammond, above n 123, 90.

128 As noted above n 9, nearly 99 per cent of companies in Australia are proprietary limited companies.

> the proposals for the establishment of insurance schemes or trust funds are a major departure from the current system and would require a thorough examination and extensive consultation with industry before even preliminary models could be produced. The Committee believes that the proposals are worthy of further attention but suggests that much ground work would need to be done before any serious consideration could be given to the proposals.[129]

The Employee Entitlements Support Scheme (EESS)[130] was introduced by the Federal Government on 8 February, 2000. It adopted the recommendations of the Ministerial Discussion Paper entitled 'The Protection of Employee Entitlements in the Event of Employer Insolvency'.[131] Its purpose was to provide a safety net for employees who lose their jobs due to the insolvency of their employers. As noted above, the alternative proposal for an insurance scheme contained in that Discussion Paper was rejected.

The EESS scheme involved a 50% contribution from the states collectively, but initially support was not forthcoming.[132] Unions also expressed considerable concern about the scheme.[133] It was replaced by the General Employee Entitlements Redundancy Scheme (GEERS) on 11 September, 2001, which does not rely on contributions from the states. In addition, a scheme purely for Ansett employees called the Special Employee Entitlements Support Scheme (SEES) was introduced, funded by a levy on airline tickets.[134]

GEERS provides for the payment of all unpaid wages, all accrued annual leave, all accrued long service leave, all accrued pay in lieu of notice and up to 8 weeks of redundancy entitlements. The restrictive cap on total payments under the EESS[135] was removed, and a much more generous one derived from the *Workplace Relations Act 1996* (Cth)[136] put in its place.

The protection of employee entitlements by a government scheme has a number of disadvantages. Those who are not classified as employees, such as independent

---

129 Stocktake, above n 29, [10.86].

130 See further Hammond, above n 123; also, Christopher Hughes, 'Towards Pinstriped Unionism: Protecting Employee Entitlements Through Securitisation' (2000) 12 *Bond Law Review* 7.

131 Ministerial Discussion Paper, above n 123.

132 For a full discussion of the history of the scheme, see O'Neill and Shepherd, above n 124.

133 Ibid.

134 *Air Passenger Ticket Levy (Collection) Act 2001* (Cth) and the *Air Passenger Ticket Levy (Imposition) Act 2001* (Cth).

135 $20,000.

136 $75,200 for 2001-2002, $81,500 for 2002-2003 and $85,400 for 2003-2004.

contractors, miss out on the scheme's benefits. It may also encourage a degree of carelessness on the part of directors, since they know that their workforce's entitlements will be protected. However, the recent Stocktake report was of the view that 'GEERS is an important aspect of the overall arrangements for the protection of employee entitlements in Australia and should continue to be a feature of those arrangements.'[137]

Tort Claimants

Tort claimants are in a unique position amongst creditors in that they have no capacity to price protect or to diversify in order to spread the risk that their claims will be unpaid. Even where there is also a contract between a creditor and the company, and where the tort committed is, for example, negligent misstatement with respect to the contract, it is unlikely that the contract price will include a premium to protect against tortious, as opposed to contractual, breach. Diversification is also clearly not possible.

Tort claimants, due to their involuntary nature, are also disadvantaged in terms of information. They are unlikely to have any prior knowledge of the company's financial position or the fact that the company's insolvency may intervene to prevent full payment of their claims.

There is a limited degree of priority in a winding up for some tort creditors. Under s 556(1)(f) of the *Corporations Act*, amounts due in respect of injury compensation rank as a category of preference in a winding up. However, as noted above, the entitlements of employees rank ahead of injury compensation payments, and may not be fully compensated by liquidator distributions in a winding up. If they were, government initiatives such as GEERS and the maximum priority proposal would not have been considered. Therefore, given that the higher ranking categories of priority are to be paid in full before lower categories receive any distribution, it is unlikely that tort creditors will obtain adequate recovery as a result of this priority.

In addition, the priority only extends to injury compensation. There are many types of torts that do not involve injury, and therefore that are left with no priority for payment. This lack of priority might be considered appropriate, given that tort claims which do not concern personal injury are treated more strictly by courts,[138] than those which do. However, it should be remembered that the claims of trade creditors and

---

137 Stocktake, above n 29, [10.79].

138 For example by the 'salient features' requirement of a duty of care, as shown by *Graham Barclay Oysters Pty Ltd v Ryan* (2002) 194 ALR 337.

employees also entail economic loss rather than physical injury, and in both of those situations, there is at least some consent to deal with the company and to a greater or lesser extent, the ability to price protect or diversify. Tort claimants, whether suffering personal injuries or some form of economic loss, lack all of these. Therefore, it can be seen that the small amount of priority which tort claimants do enjoy under s 556(1) of the *Corporations Act* is clearly inadequate as well as inconsistent with the treatment of employees.

Currently there is no specific government scheme targeted at the compensation of those claiming in tort against insolvent companies. Schwarcz considered it a viable alternative to imposing liability on directors. He said:

> Where contracting parties fail to internalize externalities, government can mitigate distributional inequities through taxation and transfers. For example, government could give tort claimants, the most common type of non-adjusting creditor, priority over other unsecured creditors; or it could create a fund for paying unpaid tort claims; or it could give tax breaks to unpaid tort claimants.[139]

The advantages of such a scheme would be that compensation should be available more quickly and easily, without protracted and costly litigation. Complex questions of fault would also be irrelevant. However, as was noted above with respect to the GEERS scheme, a government 'safety net' may undermine the deterrence aspect of the imposition of liability on directors and also socialises the cost to those not at fault.

In addition, since there is no general no-fault tort compensation in Australia, it is arguably unfair to allow directors acting tortiously to escape the consequences of their actions simply because of their status as directors where other individuals do not. Moreover, those schemes which do exist, such as those for transport or workplace injury compensation,[140] generally focus on personal injury losses and ignore economic loss. Claimants under these schemes do not usually receive compensation for all their losses.

One possible way of ensuring tort claimant recovery could be by lifting the corporate veil to make shareholders liable for unpaid corporate torts. Hansmann and

---

[139] Steven Schwarcz, 'Collapsing Corporate Structures: Resolving the Tension Between Form and Substance' (Paper presented at the Corporate Law Teachers Association Conference, Canberra, 10 February, 2004) 17.

[140] See http://www.workcover.vic.gov.au/dir090/vwa/home.nsf/pages/about and http://www.tac.vic.gov.au/.

Kraakman[141] saw no economic justification for shareholder limited liability for corporate torts, regardless of the size of the company, and they identified a number of reasons why limited liability may cause sub-optimal decision making. Limited liability externalises the true costs of engaging in hazardous activities for companies without sufficient funds to meet judgments. Therefore it reduces the incentives for companies to take appropriate precautions and for shareholders to vote for a board prepared to pay for safety measures. Following this reasoning, it also encourages shareholders to invest in risky undercapitalised enterprises where their dividends may be large, because neither the company nor the shareholders will bear the liability if the risk eventuates. These enterprises may have a positive value for shareholders but not for society as a whole.[142]

The authors, however, acknowledged that in the case of publicly traded companies, lifting the corporate veil to impose shareholder liability may undermine the free trading of the shares which is so fundamental to the market in securities. In addition, the timing of the attribution of liability is problematic for a number of reasons. If liability attaches to the shareholders at the time of the tort, this time may be difficult to ascertain if the damage accrues, as in the asbestos cases, over a long period of time. If it is attributed to those at the time of the first claim being made, there is a great incentive in mass tort cases for shareholders to sell their shares to avoid liability for further claims. The same reasoning applies to liability attaching to those who are shareholders at the time of judgment.[143]

Moreover, the imposition of liability creates an incentive for shareholders to minimise their own assets, in the same way that companies may deliberately undercapitalise.[144] Leebron commented that '[i]n fact, there is no such thing as truly unlimited liability'.[145]

Comparison of Methods of Creditor Protection and Compensation

The aim of this article was to explore the array of options for ensuring that creditors are compensated for their losses in the event of the insolvency of their debtor companies, without imposing liability on directors. These range from imposing

---

141 Henry Hansmann and Reinier Kraakman, 'Towards Unlimited Shareholder Liability for Corporate Torts' (1990) 100 *Yale Law Journal* 1879.

142 Ibid 1882-3.

143 Ibid 1896-7.

144 Ibid 1910.

145 Leebron, above n 142, 1575.

personal liability on various parties such as shareholders, to socialising the cost of compensation through debt insurance and government assistance.

However, each of the options for general creditor protection examined in Part III were either insufficient for vulnerable creditors or were problematic. Mandatory insurance of debts may be excessively costly and may encourage risky behaviour. In addition, setting appropriate levels of debt insurance is difficult, and insurance may not be available to all companies. Setting mandatory levels of capitalisation is also complex. Apart from the external costs of regulating and enforcing any requirements, companies face costs in both administration and in the restrictions on their ability to raise and use finance. Again, setting appropriate levels is difficult.

Making holding companies liable for the debts of their undercapitalised subsidiaries is sometimes suggested as an alterative to overcome the problems associated with mandatory capitalisation levels. It would discourage the practice of holding companies incorporating undercapitalised subsidiaries for the purpose of engaging in excessive risk taking, and would surmount the present obstacle facing creditors of knowing which company in a complex group should be sued.

Section 588V of the *Corporations Act* is arguably a partial solution to the problem but it is too narrow for general creditor protection, as it merely targets those debts incurred while the subsidiary company was insolvent. Some would argue that the section is too wide because it lifts the corporate veil to impose liability on shareholders. In addition, the imposition of liability under the current provision may encourage the holding company to turn a blind eye to subsidiary practices and may also have a detrimental effect on recovery by holding company creditors. Nonetheless, some measure of holding company liability or liability of holding company directors may be a part of a solution to the need for creditor protection.

The imposition of liability on shareholders generally is highly contentious. The limited liability of shareholders is seen as the bedrock of incorporation and thus the economy. Lifting the veil by statute on non-corporate shareholders, even in a partial way, is unlikely to be acceptable as a way of protecting creditors. For example, removing the privilege of limited liability from shareholders of small, undercapitalised companies would possibly stifle the incorporation and growth of new enterprises, essential for economic and social progress and development. Removing limited liability with respect to certain types of action may appear to be a possible answer but is arguably unworkable. For example, the removal of limited liability for tort claims would be hard to implement due to problems associated with the identification of liable shareholders.

Solutions for the benefit of all creditors, therefore, are inadequate in ensuring recovery for deserving creditors. Likewise, the solutions targeted specifically to

undercompensated and uncompensated creditors, outlined in Part IV of this article, were also seen to be unable to offer the protection which those vulnerable creditors require.

For example, providing priority in a winding up for undercompensated trade creditors would be troublesome. It is difficult to generalise about trade creditors in their ability to price protect against the risk of loss and to diversify their risk. Some will be particularly vulnerable to loss and others not at all, with many degrees of undercompensation in between. Since the fact and degree of their undercompensation is not uniform, to distinguish the infinite variety of situations may well consume resources which are much better distributed to creditors in general.

It is easier to make generalisations about the inability of employees and tort creditors to self-protect and therefore about their worthiness for a targeted remedy. As noted above, employees frequently lack information about the financial viability of their employer. They are unable to diversify away their risk of loss, and they may lack the contractual power to negotiate a premium to compensate for this risk. Tort claimants are the most vulnerable of all, lacking all means of ex ante self-protection.[146]

Therefore this article maintains that tort creditors and employees are deserving of special rights to recover their entitlements, even if the claims of these plaintiffs takes funds away from other classes of creditor. The means of providing these rights, however, is not straight forward, as all the methods of recovery outlined above have some shortcomings. Giving employees and tort creditors priority in a liquidation, for example, does not guarantee sufficient funds for recovery. This is clear from the further measures to cover employee entitlements which have been implemented or suggested since the late 1990s.

For employees, a taxpayer-funded or even employer-funded compensation scheme administered by the government, such as GEERS, appears to be a simple way of ensuring that they receive the full amount of their entitlements, but the existence of the scheme has the potential to undermine the deterrence of improper behaviour of the imposition of liability on directors which Part 5.8A of the *Corporations Act*[147] might otherwise achieve.

---

146 Clearly, tort victims could insure themselves against losses, particularly for those resulting from personal injuries. It would be more difficult, and possibly very costly, to insure against losses from negligent misstatements, breaches of copyright or other tortious conduct.

147 Pursuant to s 596AC(1) of the *Corporations Act*, a person who has breached s596AB(1) by entering into a transaction with the intention of preventing or significantly reducing the amount of the recovery of employee entitlements is liable to pay compensation. There are

In contrast, tort claimants against insolvent companies are presently ignored by government protection schemes, possibly because tort creditors of bankrupt individuals are not protected. The New Zealand experience shows that accident compensation schemes are not always economically effective,[148] despite the cost savings from avoiding litigation that such schemes bring. Even if government protection of tort claimants were to be considered, it is highly unlikely that tax payer funded assistance would be provided for the payment of tort claims for economic loss unrelated to personal injury or property damage.

The reason for the absence of tort compensation recovery mechanisms under the *Corporations Act* has never been articulated. Economic theorists have consistently acknowledged that tort claimants are neglected by contractarian analysis of corporations,[149] so their lack of ability to be compensated by the market for the risk of non-payment comes as no surprise. Perhaps the reason for the neglect is that tort creditors are involuntary and their claims will occur regardless of the fairness of the regime that deals with them. Contract creditors, on the other hand, are optional – the more their interests are protected, the more they will be willing to engage in commercial activities to the benefit of all society.

Another possible reason for the present disregard of tort claimant's rights is that their numbers are much smaller than those of trade creditors or even employees. Unlike employees who are represented by trade unions, tort creditors have isolated and individual claims, and thus are rarely able to mobilise in sufficient numbers to force action by the government or the company for the protection of their interests.[150]

---

difficulties with actions against directors under this section, one of which is the need to prove a subjective intention on the part of the director.

148 New Zealand has had a no-fault accident compensation scheme since 1972. Its *Accident Compensation Act 1972* (NZ) established the Accident Compensation Corporation (ACC). The *Injury Prevention, Rehabilitation, and Compensation Act 2001* (NZ) is now the principal Act under which ACC operates. However, the scheme has been criticised as being uneconomical and subject to widespread fraud. See Stephen Todd, 'Negligence Liability for Personal Injury: A Perspective from New Zealand' (2002) 25 *University of New South Wales Law Journal* 895, 900, who noted that in 1997, the unfunded liabilities from the scheme was estimated at NZ$8.2 billion.

149 Leebron remarked: 'Tort claimants differ from contract creditors in important ways. Indeed almost every commentator has paused to note that limited liability cannot be satisfactorily justified for tort victims ("involuntary creditors") and then moved on as though there is nothing to do about this unfortunate wrinkle in the economic perfection of the law'. Leebron, above n 142, 1601 (footnote omitted).

150 Exceptions to this are the class action torts, such as the James Hardie action, outlined in Part IV above. However, it is noteworthy that those tort claimants were former employees

The shortcomings of alternative ways of protecting creditor focus attention on the director as a source of compensation for creditors. Presently, there are a number of possible actions to recover funds for the benefit of creditors available against directors. Some are enforceable only by the company's liquidator.[151] These have the advantage of avoiding a multiplicity of actions and cost, and are therefore of benefit to plaintiffs with small claims for whom the expense and risk of litigation would be prohibitive. Some actions can be taken by creditors directly, although often liquidator consent is needed.[152]

The principal advantage of imposing liability on directors is that it provides two forms of protection for creditors – it deters a director from unacceptably risky conduct which can lead to creditor losses, whilst also providing a measure of compensation for creditors if the director is not deterred. Before concluding, however, that directors' liability is a superior means of creditor protection, two important matters need to be considered. The first is the adverse effect that the possibility of actions against directors might have on appropriate risk taking and the attractiveness of directorships. The second is the availability of insurance.

What amounts to appropriate risk taking alters during the life of a company. Directors may increase the level of their risk taking when the company approaches insolvency. Yeo and Lin remarked that

> [t]he argument that the debtor's self-interest will restrain unnecessary risk taking does not stand when the company is in financial distress. As the company may have no future to think about, accordingly it is less likely to be concerned about its credit rating. Self-interest may cause the company to take only a short term perspective of the gain from high-risk activity.[153]

The rationale for director liability is that it ought to curb this behaviour, to the benefit of creditors' ability to recover their debts from the company. Directors will avoid decreasing the few assets left for the creditors or incurring further debts which will compete for payment.

---

of the defendant company and were assisted by trade unions in their claims against the company.

151 These include the fiduciary duty to consider creditor interests, and statutory actions under *Corporations Act* pt 9.4B, s 598(2) and s 592(3).

152 For example, s 588R and s 596AC(3) of the *Corporations Act*. An exception is a claim for damages or an injunction pursuant to s 1324 of that Act.

153 Victor Yeo and Joyce Lin, 'Insolvent Trading – A Comparative and Economic Approach' (1999) 10 *Australian Journal of Corporate Law* 216, 230.

However, the response of economic theorists and legal commentators, outlined below, is that the effect of liability is not only to deter unduly risky behaviour, but also to discourage appropriately risky behaviour, which behaviour could benefit the company, shareholders and creditors alike. Directors may concentrate on strategies to minimise the risk of possible liability, rather than on the growth and prosperity of the company for the benefit of its shareholders. Byrne reasoned that

> the more serious cost is the effect the liability regime will have on the performance of the director. Their inability to efficiently cope with the liability would logically mean further incentive to avoid the riskier ventures which raise the potential losses. It is this cost which may be seen to be of significant social consequence. It is extremely difficult to measure the size of such cost and, therefore, whether or not it will outweigh the benefits to creditors … [154]

Yeo and Lin[155] agreed:

> The economic model on optimal sanctions is this: a person engages in the wrongful conduct because the expected benefits of the wrong to him or her exceed the expected costs. The law cannot vary the level of utility of the wrongful conduct for a director but it can increase the cost to him. ... The logic of imposing personal liability on directors assumes the law has a significant impact on directors' decision making. But a logical rule does not necessarily mean an efficient rule. The important question is what the total social costs involved in imposing liability on directors are.
>
> A significant social cost is the adverse effect potential liability may have on directors' performance. It is essential to bear in mind that … directors are expected to make risky decisions for the benefit of shareholders of the company. The impact of potential liability … is likely to be a function of the scope of the provisions, severity of sanctions imposed and avenues through which the directors may shift the risk from themselves to third parties.
>
> Where the scope of the legislative provisions is very wide, the spectre of being subjected to personal liability is all the more real and unpredictable for directors. … The standard of care required of directors tends to be objective in all the legislative provisions, which means the directors will labour under the possibility that ex post judicial review will find that what they have done or not done is insufficient. The fear of personal liability may create hedging behaviour which will yield a social loss because capital is diverted to more inefficient but less risky uses. Seen in this light, directors' liability can seriously jeopardise allocative efficiency goals.[156]

---

154 Byrne, above n 11, 283.

155 Yeo and Lin, above n 153.

156 Ibid 234.

Oesterle commented that '[l]egislatures, to catch a few more wrongdoers, have erected bars to legitimate business judgments in recurring and significant situations'.[157] Much has been written about risk aversion in the context of insolvent trading and indeed, 'catching a few more wrongdoers' appears to be the deliberate intention of that law. The insolvent trading provision, s 588G of the *Corporations Act*, was tightened from 'reasonable grounds to expect'[158] to 'reasonable grounds to suspect'.[159] This was a recommendation of the Harmer Report[160] and was designed to encourage directors to be more rigorous in considering the company's financial affairs, and where appropriate, to initiate insolvency administration.[161] This test is an objective one, assessed by courts with the benefit of hindsight. It should also be noted that the defences were tightened in 1992. Deterring inappropriate risk taking was therefore the intention of the law.

---

157 Dale Oesterle, 'Corporate Directors' Personal Liability for "Insolvent Trading" in Australia, "Reckless Trading" in New Zealand and "Wrongful Trading" in England: A Recipe for Timid Directors, Hamstrung Controlling Shareholders and Skittish Lenders' in Ian Ramsay (ed), *Company Directors' Liability for Insolvent Trading* (2000) 19, 28. This statement overlooks the existence of the court's power to excuse a person from liability. The court has power under s 1318 of the *Corporations Act* to grant relief in connection with any civil proceeding for negligence, default or breach if the person has acted honestly and ought fairly to be excused. Therefore it is arguable that only directors acting improperly will suffer the consequences of their actions. Nonetheless the fact remains that directors may still fear liability – with the attendant effects of risk aversion, the need for additional compensation and the reluctance to act as directors – because they cannot be sure that their behaviour will be excused under this provision.

158 *Corporations Law* s 592(1)(b).

159 *Corporations Act* s 588G. In *ASIC v Plymin, Elliott and Harrison* [2003] VSC 123, [427], Mandie J adopted the opinion of Kitto J from *Queensland Bacon Pty Ltd v Rees* (1965) 115 CLR 266, 303: 'A suspicion that something exists is more than a mere idle wondering whether it exists or not; it is a positive feeling of actual apprehension or mistrust.' Mandie J said: 'Section 588G(2)(a) does not require the plaintiff to establish that the particular director had an actual suspicion that Water Wheel was insolvent, nor does it, in my opinion, require the plaintiff to prove that the director had an awareness that the specified facts and matters which were within his knowledge in fact constituted grounds, or reasonable grounds, for suspecting insolvency. Such a requirement would be no different in practice to a requirement that it be proved that the director had an actual suspicion of insolvency. What s588G(2)(a) requires is proof of a subjective awareness by the director of grounds, whether or not the director had a "subjective suspicion" of insolvency, which grounds may be objectively characterised as reasonable grounds for suspecting such insolvency.'

160 The Harmer Report, above n 39, [287].

161 Niall Coburn, 'Insolvent Trading In Australia: The Legal Principles' in Ian Ramsay (ed), *Company Directors' Liability for Insolvent Trading* (2000) 73, 100.

Not everyone agrees with this approach. Discussing the tightening of insolvent trading liability, Dabner commented, with respect to the word 'suspect':

> Its precise scope is unclear and arguably too onerous, especially when viewed in conjunction with the other extensions to the application of the defaulting officer provisions. Arguably, it shifts the balance too far towards protecting creditors by promoting a risk-averse culture in directors performing a risk taking function. Certainly directors ought to be encouraged to contemplate the expected or probable outcomes of their decisions. However, to extend liability to them for failing to appreciate or act on a concern as to a *possible* outcome is, in the context of a risk taking venture, an obtuse responsibility.[162]

Oesterle pointed out:

> The legal conundrums have real effects. Whenever a jurisdiction adopts an insolvent trading provision, business people, concerned about the potential breadth of the remedy and about the difficulty courts have in accurately assessing after the fact whether trading falls outside the provisions' prohibition, will take extra precautions to stay out of court.[163]

In addition, the fear of liability may deter people from becoming company directors.[164] Oesterle remarked that

> executives on boards will be more likely to resign at the first sign of trouble. Firms may find themselves looking for directors to fill vacancies and to make critical decisions just when good business people will slam the door on inquiries.[165]

> Expose (non-executive) directors to personal liability and one will see many resign from all but the healthiest of companies. Firms cannot pay them enough to compensate them for the personal risk. Sadly, outside directors are the least needed in the best running companies and are the most needed in companies that are suffering through difficult times.[166]

---

162 Justin Dabner, 'Trading Whilst Insolvent – A Case for Individual Creditor Rights Against Directors' (1994) 17 *University of New South Wales Law Journal* 546, 562.

163 Oesterle, above n 157, 34.

164 Dabner, above n 162, 561; also, Oesterle, above n 157, 29.

165 Oesterle, above n 157, 30.

166 Ibid 31. The Age newspaper reported that non-executive directors 'face legal risks (they can be sued) and reputational risks (they are vilified if the company goes bust). … And while their pay packet might appear to be nominally decent to the average worker, it seems it is not enough to attract and keep non-executive directors.' Gabrielle Costa, 'More Non-Execs Wonder if the Pay is Worth the Pain', *The Age* (Melbourne), 25 September 2004, Business 5.

Directors who do remain with the company may demand additional compensation for the risks to which they are exposed. Byrne contended that directors

> are extremely poor risk bearers. Directors, particularly when bound in service to one company, are unable to diversify their investment and spread their risk. Their personal liability may be unlimited. It would necessarily follow, in the same way that creditors seek compensation for the increased risks [due to] limited liability, that the directors would need proper compensation for their risk. Given the inability of the director to avoid the potential liability or reduce its impact as an inefficient risk bearer, the compensation would have to be quite high.[167]

Alternatively, directors, fearing personal liability, may prematurely put a company into liquidation, even where it may be possible for the company to trade out of its difficulties.[168] Yeo and Lin declared:

> Making insolvency a prerequisite to the imposition of obligation on directors may unduly constrain the directors' decision making. In order to avoid having personal liability, the directors may feel the pressure to cease business before the company becomes insolvent, which under some circumstances may bring about premature liquidation. … Highlighting insolvency or the last act that pushed the company into insolvency does not help in establishing the critical link with the extent of the risk that the creditors are prepared to accept.[169]

While the above discussion deals with the effect of the insolvent trading provisions on director behaviour, it is also necessary to look at the effect of the availability of the remedy on creditor behaviour and whether it leads to economically efficient behaviour. Oesterle maintained that 'it encourages creditor complacency'.[170] Byrne noted:

> If we accept that predominantly creditors accept the higher risks, given that in total it is cheaper for them to bear the associated costs, then to ignore the steps they should take to protect their interest is to leave out half of the answer and to solely place the expectations on only one of the parties involved.[171]

Despite the dire warnings of these commentators, it is interesting to note that none cited any empirical research to support the contention of risk aversion. In addition, given that the law has gradually been increasing the extent of director liability, there

---

167 Byrne, above n 11, 282. (footnotes omitted)

168 Ian Ramsay, 'An Overview of the Insolvent Trading Debate' in Ian Ramsay (ed), *Company Directors' Liability for Insolvent Trading* (2000) 1, 9.

169 Yeo and Lin, above n 153, 231-2.

170 Oesterle, above n 157, 41.

171 Byrne, above n 11, 281-2.

has been no statistical evidence of a detrimental effect of that liability on economic growth or employment growth. Companies, including ones with a high probability of business failure,[172] continued to be incorporated.[173] It could be argued therefore that the risk aversion effect of the tightening of, for example, the insolvent trading legislation has had a scarcely noticeable effect on society. In respect of the British provision for insolvent trading, Keay noted:

> While it might be argued that from a normative perspective section 214 itself is not defensible for the same reasons that contractarians challenge the existence of a duty to creditors, it is interesting to note that there is no evidence that the advent of section 214 has caused a reduction in the amount of risk-taking that occurs in British markets.[174]

Another issue relevant to the imposition of liability on directors is insurance and whether its availability undermines the objective of deterrence of excessively risky behaviour that liability may otherwise produce. Directors and officers' insurance against liability whilst acting as a director can be obtained by the director himself or else paid for by the company, subject to limits.[175] It is often suggested as a way to ameliorate the harshness of imposing liability on directors whilst still ensuring that creditors are compensated. Certainly, it may reduce the pecuniary aspects of attributing blame to a director.[176] Finch noted that

> insurance proposes a quite different model of justice in which notions of cause and blame are replaced by the idea of a distributive sharing of a collective burden. Insurance makes directorial wrongdoing a collective or social burden to be imposed at the end of the day upon shareholders and /or consumers. If not to the retributivist's taste, it may more easily be justified in the name of efficiency.[177]

---

172 For example, the 'dotcom' boom of the 1990's.

173 As at 31 December, 2002, there were 1,232,150 companies in Australia. That figure had grown to 1,309,870 by 30 June, 2004. This shows a increase of 77,720 companies in 18 months. Emails from Debbie Cowley, Product Team, ASIC, to Helen Anderson, 9 April, 2003 and 6 December, 2004.

174 Keay, above n 5, 685.

175 *Corporations Act* s 199B, discussed further below.

176 Finch noted that 'the risk spreading effect of insurance may dilute the punishment so that it does not fit the crime or the breach.' Vanessa Finch, 'Personal Accountability and Corporate Control: The Role of Directors and Officers' Liability Insurance' (1994) 57 *Modern Law Review* 880, 887. However, any 'crime' for which retribution seems appropriate is unlikely to be covered by insurance. This is discussed further below.

177 Ibid 891.

However, the payment of the actual damages claim is only one aspect of a finding of liability against a director. When directors act in breach of the law, they may be subject to prosecution.[178] They risk damaging their reputations and their ability to obtain other directorships. They risk increased insurance costs in the future as well as the possibility of insurance being unattainable. Many types of improper behaviour are excluded from coverage of insurance policies.[179] Therefore, insurance of directors is not considered to be contrary to the aim of deterrence that the imposition of liability seeks to achieve.

Insurance is a pragmatic solution to the problem of directors choosing to have no assets. It would be difficult and costly for the law to ensure that directors have sufficient personal assets to satisfy a judgment debt against them. The imposition of liability may encourage directors to shift their assets into the ownership of spouses or other entities.

Insurance also allows the cost of risk to be borne more efficiently. The cost of the insurance is likely to be borne by the company and its shareholders whether the policy is taken out by the director or the company. If directors pay for it themselves, they will pass on the cost to the company in the cost of their services.[180] As mentioned previously, directors are unable to diversify their risk, and so many demand substantial compensation for bearing the risk of liability while acting for the company. However, an insurer has the capacity to diversify away that risk, because it is insuring many different directors from diverse companies. Therefore, the same risk will be borne but by the insurer rather than the director. Thus the premium paid should be less than the amount of compensation demanded by the directors for bearing the same risk. In addition, the insurance payout pursuant to the policy is a certainty, if the claim

---

178 However, in this regard the research of Tomasic and Bottomley should be noted. They found that '[t]he paradox of corporations law is that although corporations are creatures of law, the vast body of corporation law seems to have limited impact upon public company directors in general. Reliance upon the vague notion of good corporate citizenship helps to keep the law at bay, although this has little real impact upon the basic motivations of corporate executives and directors. This is not to suggest that most directors are not law abiding, but rather, that they pay relatively little attention to the formal legal rules which govern their positions'. Tomasic and Bottomley, above n 37, 83.

179 This is discussed further below.

180 Michael Whincop, 'Reintroducing Releases of Officer Liability into Australian Corporate Law' (2000) 26 *Monash University Law Review* 14, 28.

is within the terms of the policy, whereas the asset level of directors when sued is not.[181]

There are a number of benefits from insurance being held by the directors rather than debt insurance being held by the company. Kraakman[182] made the following argument as to why directors ought to bear liability in certain circumstances[183] and be permitted, in some of these circumstances, to insure against it:

> [T]he risk shifting opportunities that top managers enjoy in the form of indemnification and insurance … can force the firm and its shareholders to internalise the expected liability costs that undercapitalisation would otherwise impose …
>
> From this perspective, the real defect of personal liability as a check on undercapitalisation lies in the danger that agents will *not* pressure the firm into providing adequate coverage of their personal liability risks.[184]
>
> … [S]enior managers and directors are ideal targets for incentives aimed at ultimately prodding the firm to cover its potential liability. Their position normally provides them with information about the need for insurance; their power assures that they can act on their knowledge of risk levels; and their personal assets and risk preferences are likely to encourage them to seek adequate insurance coverage. Even when they are not the firm's cheapest harm avoiders, they are likely to be its most reliable insurers. … When [directors] are not principal shareholders, their demand for insurance may be even greater, since they do not receive the full return of gambling their personal assets unless they are specifically compensated for their risk bearing.[185]

Thus the imposition of liability provides directors with a powerful inducement to either be insured against liability themselves,[186] or else make certain that the company is insured or properly capitalised. Kraakman maintained that director liability

---

181 It should be noted, however, that insuring directors' personal liability does not reduce the cost of litigation. The insurer takes over the handling of the claim from the director and is subrogated to his rights.

182 Kraakman, above n 75.

183 These include asset insufficiency (undercapitalisation), sanction insufficiency and enforcement insufficiency. Ibid 867-8. Sanction insufficiency refers to behaviour, such as deliberate crimes, where Kraakman recommended that directors be liable and that insurance not be available, because it would undermine the law's power to deter. Enforcement insufficiency is where neither individual or corporate penalties are sufficient to achieve compliance with the law.

184 Ibid 870.

185 Ibid 871.

186 Leebron, above n 142, 1636.

> protects against legislative over- and under-provision for tort risks, and it permits [directors] to select the optimal strategy for covering risk from among insurance [of the company], self-insurance and risk reduction through control of the firm's activities.[187]

In addition, the protection provided by insurance may encourage non-executive directors to serve on boards. These directors can act as effective monitors and therefore as control mechanisms on the company's behaviour.[188] The temptation on directors to delegate tasks to lesser staff as a device to avoid the possibility of personal liability is removed or reduced when the directors have insurance.

Currently, however, the ability for the company to insure directors against liability is constrained. Section 199B of the *Corporations Act* provides:

> (1) A company … must not pay, or agree to pay, a premium for a contract insuring a person who is or has been an officer… of the company against a liability … arising out of:
>
> (a) conduct involving a wilful breach of duty in relation to the company; or
>
> (b) a contravention of section 182 or 183.

This provision would therefore prevent a company insuring its directors against liability for a wilful breach of the duty to consider the interests of creditors when the company approaches insolvency, improper use of position or improper use of information. In addition, it was noted in the recent Directors and Officers Insurance Report by the Corporations and Markets Advisory Committee (CAMAC Report)[189] that in practice, insurers frequently have a host of other exclusions, whether the insurance is taken out by the company or by directors themselves. These may include prospectus liability, insider trading liability, liability for shareholder claims, dishonesty or fraud, as well as liability for insolvent trading.[190] It appears, therefore, that the exclusions relate to many of the types of actions where it is likely that directors would be sued by creditors.

After taking into account risk factors, insurers may also be reluctant to extend cover to certain applicants. New businesses, technology related enterprises, and speculative mining ventures may find insurance hard to obtain. Directors expressed concerns over

---

187 Kraakman, above n 75, 874.

188 Finch, above n 176, 889.

189 Corporations and Markets Advisory Committee, Parliament of Australia, *Directors and Officers Insurance*, 2004.

190 Ibid [2.3.4].

the limits on the availability and extent of insurance, in a survey cited in the CAMAC Report. The respondents to the survey pointed out that the lack of insurance for directors of companies in economically risky sectors

> affected the ability of these companies to attract experienced persons willing to serve on their boards. Indeed, 41% of respondents considered that their concerns regarding insurance acted as a disincentive for them to take up directorships and engage in risky activity.[191]

Similarly, those without an established financial track record or whose company is suffering a deteriorating financial position may be unable to find insurance. Public company directors, facing a wider range of possible breaches than their private company counterparts, may be further disadvantaged.[192] The cost of insurance has also increased. The CAMAC Report notes that 'premiums for D&O insurance have increased since 2001 on average between 35-50% on an annual basis. Less attractive risks … have faced much higher increases'.[193]

Therefore, while insurance is seen by some commentators as meeting the compensation needs of creditors in an economically efficient way, the unavailability of insurance for some directors means that it cannot be relied on for every type of director liability. For example, it is likely that insurance would not cover errant directors found liable to trade creditors with respect to uncommercial transactions, civil penalty breaches, insolvent trading or breaches of fiduciary duty. Therefore such insurance that is available is unlikely to have any adverse impact on the deterrence effect of imposing liability on directors where they behave improperly.

Conclusion

Frequently, it is assumed that if a company is unable to pay its debts, the loss should either lie where it falls or else directors should be liable for it. After all, directors control the company and if the loss to creditors is a result of some wrongdoing on their part, liability serves as a means of both compensation and retribution. The fear of personal liability also acts as a powerful deterrent to improper behaviour.

However, there is a perception that imposing liability on directors will lead to excessive risk aversion, to the detriment of the company's profitability, as well as legal compliance burdens. Experienced business people may refuse directorships, except

---

191 Ibid [4.2.1].

192 Ibid [2.4].

193 Ibid [2.5].

for the safest of companies who are arguably least in need of their expertise. For these reasons, it is worth considering alternatives to director liability.

However, it was noted above that both current and proposed alternative means of creditor protection suffer from deficiencies. Mandating capital requirements may result in burdensome rules about the composition and use of capital which unduly restricts companies in their business activities and would involve costly administrative requirements. Legislation already exists to impose liability on holding companies for the debts of their undercapitalised subsidiaries, but this only occurs where there has been acts of insolvent trading by the subsidiary's board and does not necessarily protect the creditors most in need of the law's protection, namely the subsidiary's tort claimants. Other forms of shareholder liability have the potential to undermine the concept of limited liability which is fundamental to incorporation and the market for capital.

Certain groups of creditors were identified as particularly vulnerable in the event of a company's inability to pay its debts. These were the small trade creditors, employees and tort creditors of the company. Various means of protecting them exist, such as some priority in a winding up and GEERS. However, while employees are comparatively well protected, small trade creditors and in particular, tort creditors are not.

This brings the discussion back to the imposition of liability on directors as a means of providing compensation for such creditors. The main disadvantage of actions against directors is the cost of the litigation and the possible detrimental effects that the imposition of liability will have on directors' primary role as risk takers. Risk taking[194] is seen in economic analysis as the key factor in the maximisation of shareholder wealth and its flow on effects for the economy.[195]

---

194 Len Sealy, 'Directors' "Wider" Responsibilities – Problems Conceptual, Practical and Procedural' (1987) 13 *Monash University Law Review* 164, 181 said that '[a]ny reformulation of directors' duties to take account of the interests of creditors and others has to accommodate the concept of risk, and allow for the fact that directors must be free to take risks and to judge what risks their business should take. We must not lose sight of the fact that it is the principal function of the limited liability company, and of company law, to facilitate this risk-taking'.

195 The Cooney Committee noted that '[t]he more productive the corporate sector, the more secure the economic well-being of Australia. Directors are crucial to its success. To restrict unnecessarily the operation of their skills, their industry, their enterprise, is to threaten unnecessarily a factor vital to economic growth. Any regulation of directors' activities must be warranted and a sensible balance must be found between measures necessary to promote corporate activity in a way which will be of benefit to all, and measures necessary to protect the bona fide shareholder, worker, consumer, financier, and the public at large.

Yet the risk aversion argument lacks empirical support. In addition, imposing liability on directors only where they are at fault as defined by legislation has the benefit of making apparent what amounts to excessive risk taking. This provides certainty to directors, reduces 'testing' litigation where liability boundaries are not clearly defined and ensures that they are only subject to liability where their own improper conduct has caused the creditors' losses. The notion of retribution appears lost where directors have the ability to be insured against claims, but in reality insurance may not be available for claims involving blameworthy behaviour. It is for the benefit of all present and future creditors, as well as shareholders, of companies if actions against directors for improper conduct deters the behaviour which may cause their losses. This makes director liability superior to other forms of compensation.

Profitability is but one basis for good corporate citizenship'. Senate Standing Committee on Legal and Constitutional Affairs, Parliament of Australia, *Company Directors' Duties: Report on the Social and Fiduciary Duties and Obligations of Company Directors* (1989) [2.39].

TRACING, RESTITUTION AND INNOCENT DONEES: WHO WANTS TO BE A VOLUNTEER ANYWAY?

Susan Barkehall Thomas *

The interaction between property principles and restitution for unjust enrichment is controversial. Property theorists resent the attempts by the proponents of restitution theory to assert that property is subordinate to restitution. Restitution theorists are critical of property theorists for refusing to accept their new ordering of the law. In addition, the issues involved are technical and difficult. Burrows has said that:

> the relationship between the law of property and the law of unjust enrichment/restitution has long been regarded as fiendishly complex and problematic. Indeed one can regard it as the last great unsolved mystery for those working in the law of restitution.[1]

The dispute between the property and restitution theorists is at the core of the problem of the innocent volunteer. What happens when beneficiaries of a trust attempt to trace their equitable property entitlement into the hands of a volunteer who has since changed position on the faith of the receipt?

The traditional property approach permits the claimants to assert their continuing proprietary interest in the traceable proceeds of their stolen property, unless and until it is dissipated or received by a bona fide purchaser for value without notice. Accordingly, the volunteer recipients will be liable provided that they retain the traceable proceeds.

The restitutionary approach permits the volunteer recipient to claim the defence of change of position. If the defendants retain the traceable proceeds, but have otherwise satisfied the requirements of the restitutionary defence, they will not be liable to repay.

The opposing views have been well-stated before, but little attempt has been made to go beyond theory.[2] This article moves beyond the theory and seeks to solve the dispute through economic analysis, and by the presentation of alternative solutions. Part 1 of this article sets the scene by summarizing the approaches and examining the

---

* BA.LLB (Hons) LLM (Mon), Senior Lecturer, Monash University, Faculty of Law.

1 A Burrows, *The Law of Restitution*, 2nd ed, (Butterworths Lexis Nexis, London, 2002 at 412.

2 Those who have attempted some resolution are Burrows in *The Law of Restitution,* n 1 above, and C Rotherham, *Proprietary Remedies in Context* (Hart, Oxford, 2002).

particular implications which rise from preferring one analysis over the other. In part, this will be achieved by establishing a hypothetical variation on the facts of the House of Lords decision in *Foskett v McKeown*.[3]

In Part 2 economic analysis will be applied to the problem to determine whether one approach can be preferred over the other. Fundamentally it is necessary to determine which party should bear the loss. The analysis in Part 2 will demonstrate that there are no compelling reasons from economic analysis to depart from the property-based liability rule.

Part 3 involves discussion of alternative models of beneficiary protection that restructure the problem. If the beneficiary is sufficiently protected by other means, there should be no objections to making the defence available to the volunteer recipient. This Part explores what models might be available, and how they might work.

The Competing Analyses in Depth

The House of Lords decision in *Foskett* brought to a head the debate between property law theorists and unjust enrichment theorists. The case involved a claim by trust beneficiaries to trace their misappropriated property into the hands of innocent volunteers.[4] The core question involved in the decision was whether the beneficiaries could recover a proportionate interest in the payout, rather than merely the actual sum of their funds diverted towards payment of the premiums.

Central to the reasoning of the members of the House of Lords were propositions regarding the basis of the plaintiffs' claim. Two fundamentally opposed views were canvassed. The first, conventional, view is that the plaintiffs were merely tracing their existing equitable property rights into a new form. In this article, this will be called the 'property approach'.

The second view, argued but not accepted, was that the case should be viewed as a question of proprietary rights raised to reverse a claim in unjust enrichment. This analysis of the case proceeds on the basis that (1) the defendants were enriched by the receipt of the insurance premiums; and (2) the enrichment was unjust, because the

---

3 [2001] 1 AC 102.

4 In this case, a wrongdoing trustee had diverted trust funds to his own use, paying premiums on his life insurance policy. On his death, the proceeds of his policy were paid out to his children.

plaintiffs were ignorant of the trustee's misappropriation. This is the 'unjust enrichment approach'.[5]

The Property View

There was clear adoption of the property approach (with a corresponding rejection of the unjust enrichment approach) in *Foskett*.

Lord Millett stated:

> The transmission of a claimant's property rights from one asset to its traceable proceeds is part of our law of property, not of the law of unjust enrichment. There is no 'unjust factor' to justify restitution (unless 'want of title' be one, which makes the point). The claimant succeeds if at all by virtue of his own title, not to reverse unjust enrichment. Property rights are determined by fixed rules and settled principles.[6]

Further, in relation to the relationship between property and tracing, his Lordship stated:

> A beneficiary of a trust is entitled to a continuing beneficial interest not merely in the trust property but in its traceable proceeds also, and his interest binds every one who takes the property or its traceable proceeds except a bona fide purchaser for value without notice.[7]

The conventional property analysis adopted by the House of Lords can be broken down into a number of steps, as follows:

1. Beneficiaries under a trust have equitable title to trust property.
2. Equitable title is good against the world except for a bona fide purchaser for value without notice.
3. The beneficiary's equitable rights to the property subsist notwithstanding any change in form, or transmission to a third-party.
4. The equitable tracing rules exist to identify the continuing equitable property rights of the beneficiary's notwithstanding changes in form.
5. The property rights are defeated if the third-party is a bona fide purchaser for value or if the property is dissipated.

---

5 The terms 'property approach' and 'unjust enrichment approach' are intended as nothing more than convenient labels. In this area, there is a danger in the overuse of labels to the detriment of analysis, but some convenient reference point is required.

6 *Foskett v McKeown* [2001] 1 AC 102 at 127.

7 Ibid. See also, for example, the judgment of Lord Browne-Wilkinson at 108-9.

6. If an application of tracing rules permits a conclusion that the property can be traced, the plaintiffs are considered to be asserting their original and ongoing equitable title.
7. The equitable property rights of the plaintiff beneficiary defeat any rights of a volunteer.

The property analysis is supported by, among others, Virgo[8] and Rickett and Grantham.[9]

Fictions inherent in the property approach

One of the criticisms of the property approach is that it contains an important and fundamental fiction. The fiction is that tracing involves no new property rights, but only the continuation of existing property rights, although in a new object. The fiction underlies points 3 and 6 of the above summary. Burrows offers a colourful example to illustrate the point.

He argues:

> Property is non-fictional and does have explanatory force when the claim being asserted is "I want that back because it is mine". That is the true *vindicatio* claim. It is a pure proprietary claim… No new proprietary rights are being created. One is merely passively recognising existing rights that have previously been created. The same is true of extracted minerals or the fruits of property. If I own land, I own the oil under it. But this cannot, without invoking fiction, be extended to tracing through substitutes. Ownership of a pig can explain ownership of the piglets but does not explain why P can be said to own the horse that D has obtained in substitution for the pig stolen from P.[10]

Rotherham uses a more metaphysical approach. Responding to the conventional view that tracing is merely an evidentiary process, he states:

> We normally think of the institution of property in terms of the vindication of existing rights. For this reason we are inclined to infer that tracing is not essentially remedial. Thus, jurists conclude that tracing is a way of "establishing" the plaintiff's property rights, in the sense of adducing facts to

---

8 G Virgo, *The Principles of the Law of Restitution*, (OUP, Oxford, 1999). See, also, extra-judicially Lord Millett 'Proprietary Restitution' Chapter 12, in eds S Degeling and J Edelman *Equity in Commercial Law*, (Lawbook Co, 2005, Sydney).

9 R B Grantham and CEF Rickett, 'Property and Unjust Enrichment: Categorical Truths or Unnecessary Complexity' [1997] NZLR 668.

10 A Burrows, 'Proprietary Restitution: Unmasking Unjust Enrichment' (2001) 117 LQR 412 at 417-8.

> *prove* that he or she has particular rights. However, tracing also entails to the plaintiff "establishing" rights, in the sense of *creating* a new legal relation. Tracing is the exercise by the plaintiff of a power to change his or her legal position and, in doing so, that of the defendant. This process allows the plaintiff to enjoy those rights conventionally associated with the institution of ownership.[11]

Rotherham also refers to a further problem inherent in the property approach to tracing, described by Birks as the 'geometric multiplication of the plaintiff's property'.[12] When a plaintiff beneficiary seeks to recover their stolen property, the plaintiff may choose to proceed against the holder of the original asset (if that person is not a bona fide purchaser) and claim the return of the asset in specie, or the plaintiff may choose to claim against the holder of the substitute, and recover the substitute.

As Rotherham says:

> Because A cannot be simultaneously the owner of both the original thing and the substitute, it cannot plausibly be argued that, prior to A's election to trace, a full-blown proprietary interest arises in respect of the proceeds of the sale held by B. Prior to tracing, B has a defeasible title to the proceeds of the sale – a title that he is liable to lose if A elects to claim title to those proceeds.[13]

The necessary consequence of this analysis is that there is a creation of rights.

The property approach has no real answer to this criticism. It glosses over the way that interests can, and do, jump from one object to another. Assertions that title is retained are conclusory. Fundamentally, the supporters of the property approach either deny the fiction, or accept it as necessary.

Essentially the fiction exists to serve one point. That is: the law regards property rights as enforceable against third parties. Unless tracing rules permit the ongoing recognition of property rights despite changes in form, the concept of property is attacked at its very core.

Unjust Enrichment Analysis

The alternative analysis says that *Foskett* is a case about unjust enrichment, which provides some guidance for when proprietary remedies can be granted to reverse

---

11 C Rotherham 'The Metaphysics of Tracing: Substituted Title and Property Rhetoric' [1996] 34 *Osgoode Hall Law Journal* 321 at 328-9.

12 P Birks *An Introduction to the Law of Restitution* rev (OUP, Oxford, 1989) at 394.

13 Rotherham, *Proprietary Remedies*, at 330.

unjust enrichment. The strongest advocates of the unjust enrichment analysis as applied to *Foskett* have been Professors Birks and Burrows.[14]

Under this framework, tracing has a role to play at the remedial level. The argument is established as follows:

1. The misappropriation of trust property by a trustee functions as the unjust factor of ignorance or powerlessness.[15]
2. The recipient of trust property transferred where the beneficiary is ignorant of the transfer is strictly liable to make restitution.
3. Liability is usually personal but can be proprietary.
4. Whether the remedy is proprietary will depend upon the ability of the plaintiff to trace.
5. If the property can be traced into the third-party's hands, and is still in their possession, this justifies the grant of the proprietary remedy to reverse unjust enrichment.
6. In this case the proprietary interest is purely remedial: it is a new proprietary interest to reverse unjust enrichment. [16]
7. Because unjust enrichment is the basis of the claim, rather than property, normal restitution defences apply.

The unjust enrichment approach differs from the property approach in two critically important ways.

No need for fictions

The traditional property rule requires an inchoate hovering property interest capable of jumping from one item of property to another. The conceptual difficulties with this approach were discussed above. The unjust enrichment approach, on the other hand, is prepared to accept that any proprietary remedy is created in this scenario in order to reverse unjust enrichment. This analysis avoids the need for the fiction inherent in the proprietary analysis. Rather than tracing being used to identify property rights, it

---

14 See, for example, P Birks 'Property, Unjust Enrichment and Tracing' (2001) 54 CLP 231, Burrows *Proprietary Restitution* and *The Law of Restitution*.

15 Burrows, 'Proprietary Restitution' n 10 above.

16 Cf the view of Swadling who argues that although the traced interest must be new, the fact that it is a new interest does not mean that it is an interest generated by unjust enrichment. Swadling 'Property and Unjust Enrichment' Ch 11 in *Property Problems From Genes to Pension Funds*, JW Harris (ed), (Kluwer Law International, London, 1997) at 132. He argues that there is doubt as to whether the interest is generated by unjust enrichment, as the recipient is never enriched.

is used here to identify value. The defendant's continued possession of the value justifies the imposition of the proprietary right.

As the grant of the property right is remedial only, this analysis does not encounter the conceptual difficulties that arise from the property approach. The question of where the proprietary rights subsist prior to judgment does not arise. Instead, the plaintiff has a power to assert a proprietary right.[17] This is seen to be preferable to the property approach which cannot explain the fluidity of property rights.[18]

Availability of defences

This second distinction is critical. The unjust enrichment approach permits the defendant volunteer to claim a defence of change of position in the appropriate case. However, the property approach permits no such defence. The significance of this distinction can best be illustrated by the following example.

M is a trustee and misappropriates $20,000 in trust funds by using them to pay the premiums on his life insurance policy. When he dies, his insurer pays out $1 m to his nominated beneficiaries: his children. Upon receiving the proceeds into a bank account, M's children decide to spend their own pre-existing money on an expensive holiday. The insurance proceeds are untouched, but the defendants have changed their position by spending their own money on a holiday. If not for the receipt of the insurance proceeds they would not have taken the holiday.

As discussed above, the property approach asks only whether the defendant continues to retain the traceable substitute for its property. If so, the plaintiff will have a successful proprietary claim against the volunteer. The volunteer's change of position is irrelevant to the property analysis. In the example, the volunteers will remain liable to return the proportion of the $1 million that represents the plaintiffs' misappropriated property.

By comparison, using the unjust enrichment analysis, the defendant volunteers would claim that their good faith change of position renders it unfair that they now be

---

17 See Rotherham, *Proprietary Remedies* at 93-4.

18 See, for example, S. Worthington, *Proprietary Interests in Commercial Transactions*, (Clarendon, Oxford, 1996) who states at 175: '*before* the plaintiff makes a claim and elects which property to pursue, it seems unreasonable to assume that the plaintiff has an equitable ownership in *both* [assets]. If a strict analysis is applied, it must be conceded that the tracing rules really give the plaintiff the power to crystallize a proprietary interest, rather than a continuing proprietary interest in a changing and multiplying class of assets'.

required to make restitution to the plaintiffs. [19] Although the misappropriated property can be identified in the defendants' hands, the defendants' expenditure of their own money enables them to resist the restitutionary claim.[20]

The Dispute: Again

The unjust enrichment analysis of *Foskett* marginalizes one of property law's core concepts: the proposition that property rights are maintainable against third parties. It is part of the fundamental nature of equitable title that it can be enforced against the world, until it arrives in the hands of a bona fide purchaser for value without notice. The unjust enrichment approach involves an unprecedented suggestion that property rights will be subject to personal defences: in this case the defence of change of position.

Taken at a broad conceptual level, this is a debate about the power of equitable property rights within the legal system. Should equitable property rights subsist against all but bona fide purchasers for value? In the long run, the debate is not furthered by dogmatic doctrine-driven statements. To assert that 'property rights subsist against third parties, thus the property analysis is right' does not answer the real question. So, for example, Grantham's and Rickett's assertion that we must protect property misses the point. They state:

> Property rights are a significant matter in the common law and represent one of the fundamental building blocks of the Anglo-American legal tradition. ... property rights have a powerful normative force that attracts a level of protection that borders on the absolute. One's right to do as one pleases on or with one's property is constrained only at the margins and protection from interference is available against even the most innocent. From such a perspective the idea that a plaintiff's property rights should be extinguished, to be replaced by rights born of unjust enrichment, merely because the subject matter of the right has changed form, would be a contradiction quite out of keeping with the otherwise generous protection afforded to rights to property.[21]

---

[19] *Lipkin Gorman (a firm) v Karpnale Ltd*; [1991] 2 AC 548; *David Securities Pty Ltd v Commonwealth Bank of Australia* (1992) 175 CLR 353.

[20] The defence operates pro tanto, so a complete defence would only be available if the holiday was of equal value to the enrichment.

[21] Ross Grantham and Charles Rickett 'Tracing and Property Rights: The Categorical Truth' (2000) 63 MLR 905 at 911.

This only serves to repeat the standard mantra: this is the law of property; the property law is right; anything that contradicts property is wrong.[22] It fails to explain *why* the formula is right. To break the impasse and reach the solution we have to ask the right questions. Therefore, we should be asking what reasons exist to support the maintenance of equitable title against third parties. Is there justification for refusing to permit personal defences to impinge upon the claim?

In the next part, economic analysis will be applied to the debate to provide some suggestions for whether the primacy of equitable property rights can be supported.

Economic Analysis

The purpose of this part is to use a normative economic analysis to see if it provides any guidance for preferring one approach over the other.

First, we need to consider the consequences and incentives produced by each of the competing approaches. Consequences and incentives must be considered at the level of individual transactions, but also at the broad social level. If a rule produces inefficiencies or undesirable results on a broad scale, it may not be appropriate.

Having examined the operation of the existing rules, it may appear that from a normative perspective that they are equally balanced. If this were to be the case, broad policy objectives may be used to decide upon a preference.

A. Areas of Commonality

We can easily envisage a scenario where the property approach and the unjust enrichment approach reach the same conclusions regarding liability.

As *Foskett* stood on its facts, there is no dispute from either camp about the outcome. Trust property was transferred to volunteers. It was still in their possession and had to be returned. No change of position had occurred.

There is a clear normative rationale for this common solution. The volunteers paid nothing for the property, and have received the benefit of its full value. If they are required to return it they will suffer no actual loss, and the beneficiaries will also

[22] As Rotherham says: 'the use of metaphysical justification – the argument that something simply *is* property… stifles any comprehensive consideration of how the interest in question *ought* to be protected'. Rotherham, *Proprietary Remedies*, at 118.

suffer no loss. This rule is efficient, as it ensures that no party suffers a loss. It is the 'least cost' option.[23]

B. The Disputed Area

The example offered in Part 1 (a modified version of the *Foskett* facts) represents the problem case. In the problem case, a decision in favour of one party necessarily involves the other party suffering a loss. As the defendants have acted in reliance on the wealth being their own, they will suffer loss if they are required to return the property still in their possession. Equally, the beneficiary will suffer loss if property is not returned. The competing analyses represent the opposite poles in a loss shifting exercise.[24]

If we are to go back and take a normative approach to the question, it is clear that any solution involves the balancing of competing priorities. The purpose of this part is to examine the competing concepts, and to see what approach is suggested when the priorities are considered openly.

Security of receipts

The rationale most commonly accepted for the change of position defence is that it enables the recipient to act on the basis that wealth believed to be his own is his to spend, or otherwise deal with. Birks argued that all recipients, even donees, need to be able to act on the basis that a receipt is theirs to deal with, unless there is reason to believe otherwise.[25]

---

[23] See also A Duggan 'Constructive Trusts from a Law and Economics Perspective' 55 *UTLJ* (2005) 217 at 237 where he explains that a rule requiring a recipient of a mistaken payment to return the money is efficient. He says 'in cases where P tells D about P's mistake before D spends the money,… giving the money back is a virtually costless way of avoiding P's loss'. See also S Levmore, 'Explaining Restitution' 71 (1985) *Va LR* 65.

[24] A separate line of argument could be developed that the settlor should bear the risk. This would involve major change to the underlying law of trusts, and the traditional role of the settlor. For reasons of space this more controversial argument is not pursued in this paper.

[25] P Birks 'Change of Position' The Nature of the Defence and its Relationship to Other Restitutionary Defences' in M McInnes *Restitution: Developments in Unjust Enrichment* (LBC Information Services, Sydney, 1996) at 50 – 54. Note that other views are that the defence is an individualistic one, which prevents unfairness to the defendant. Thus, for example, Lord Goff's statement in *Lipkin Gorman*; see also Kit Barker 'After Change of Position: Good Faith Exchange in the Modern Law of Restitution', Ch 7 *Laundering and Tracing*, ed P Birks, (Clarendon Press, Oxford, 1995). Barker argues that change of position is 'an individualistic

Recent English case law has implicitly supported this view, by providing a wide statement of the availability of the defence. Thus, if a recipient is no longer enriched, for example because the value was stolen or lost, the plaintiff may be unable to recover from the plaintiff. The defence is enabled if there is a causal relationship between the loss and the enrichment, on a 'but for' test.[26]

The security of receipts rationale has been applied by others. Gardner has used it to suggest that the change of position defence creates an efficient rule, as follows:

> a person who enjoys, and perceives himself to enjoy, broad security in his wealth will be more ready to spend it (other things being equal) than one who feels it necessary to look over his shoulder before doing so, so as to check that it is really his to spend – in our context, that it cannot be reclaimed from him by a trust. The former state of affairs thus allows the market to approach full efficiency more closely than the latter, and from an economic point of view is therefore to be preferred.[27]

Gardner's discussion was confined to personal claims, and he noted that the bona fide purchase defence operates to protect a purchaser's security of receipt against property claims. He did not extend his analysis to ask whether his argument forces a reconsideration of the property-based defences.

Others, however, have argued that where the proprietary claim is dependent upon tracing,[28] the change of position defence should operate to protect the defendant. Nolan's argument is that unless the defendant has the personal defence available, the:

> proprietary claim would, often as not, subvert the operation of the defence on the personal claim: by asserting a proprietary claim against a solvent defendant, the plaintiff would get what a defence denies him in a personal claim against the same defendant which arises out of the same facts.[29]

---

defence with moral foundations' at 196 and the bona fide purchase defence is the one which 'deploys a broad policy of transactional security in exchange dealings' at 192.

26 *Scottish Equitable plc v Derby* [2001] 3 All ER 818 at 827, per Robert Walker LJ; *Jones v Commerzbank AG* [2003] EWCA Civ 1663.

27 S Gardner, 'Knowing Assistance and Knowing Receipt: Taking Stock' [1996] *Law Quarterly Review* 56, 91. This discussion was related to the action for knowing receipt, and whether it should be reframed as a strict liability action. Only the operation of the defence in relation to personal claims was under consideration.

28 Nolan calls this a restitutionary proprietary claim. This is distinct, in his view, from the scenario where the plaintiff is asserting a pre-existing equitable title. The latter scenario gives rise to a proprietary claim. See R Nolan, 'Change of Position', Ch 6 in *Laundering and Tracing* n 25 above, at 177-9.

29 Ibid at 178-9.

The argument has much support in the literature, and there is a body of academics who support it.[30] Nonetheless the analysis is still open to criticism. In order to achieve consistency between personal and proprietary actions, a choice is made to prefer the framework applicable to the personal action, rather than asking which of the responses is the better one.[31]

This is the question that must be asked now.

What happens if we use the change of position defence to protect the defendant against proprietary claims? For our analysis it will not matter whether the defence is applied to only 'restitutionary proprietary' claims, or also to 'pure proprietary' claims. The position of the volunteer recipient is the same in both scenarios. Essentially the question is whether the defence provides a better normative rule than the bona fide purchase defence. Is it more efficient to allow the change of position, or to rely on the normal property rules?

Can we identify an efficient rule?

First we need to set out the consequences of each scenario. Then they can be analysed.

Scenario 1: Allowing the defence in the problem case:

- The plaintiffs will suffer a loss of $20,000.[32]
- The donee recipients will be protected to the extent of their change of position.

Scenario 2: Denying the defence in the problem case

- The plaintiffs will recover in full.
- The volunteer will suffer loss to the extent of their reliance on the faith of the receipt.

---

30 This view is also supported by Burrows, *The Law of Restitution*, n 1 above, at p 527; Birks: 'Overview: Tracing, Claiming and Defences', ch 11 in *Laundering and Tracing*, n 25 above, at 319; Virgo '*The Principles of the Law of Restitution*, n 8 above, at 729.

31 The alternative approach to prefer the proprietary position could just as easily be taken. See, for example, K Barker, 'After Change of Position', Ch 7 in *Laundering and Tracing*, n 25 above. In order to solve the problems adverted to by Nolan, Barker argues at 194-5 that the defence of bona fide purchase should be available against personal and proprietary claims.

32 In fact, as the House of Lords held that the beneficiaries were entitled to a proportionate amount of the insurance payout, their indirect loss is greater than $20,000.

Scenario 1 gives the plaintiff the incentive to prevent the fraud. The scenario 1 rule will be more efficient if the plaintiffs can more easily discover the fraud than the defendants. It will be less efficient if the volunteers could have discovered the fraud. The recipients are not only 'more ready to spend' the amount (as Gardner argues), but are actively encouraged by the law to do so, without making any inquiry into the gift's source. If the recipients have a simple means of ascertaining the legitimacy of a receipt the scenario 1 rule does not encourage them to take that step.

The scenario 2 rule will be more efficient if the volunteer defendants are better able to discover the fraud.

So far, the analysis raises more questions than answers. Can the plaintiff prevent the fraud? Can the defendant prevent the fraud? What does it cost each party to do that?

What is the ability of each party to identify and prevent the fraud?

Some brief observations can be made about the volunteer defendant before considering the plaintiff's ability to prevent the fraud.

What is the normal response of a recipient of a gift? If we assume that the donee knows and trusts the donor, she is unlikely to make inquiries into the source of the gift. If the donor is not personally known to the donee, but the source of funds is apparently legitimate, as in *Foskett* , it is also unlikely that the donee will inquire into the gift's source.

It would also seem safe to assume that most donations of gifts do not come from improper sources. If this is correct, then in most cases, a lack of inquiry on the part of the volunteer recipient will in fact be the appropriate response.

Furthermore, in many cases, the volunteer will be unable to probe deeply into the provenance of the gift. In *Foskett*, it would have been impossible for the beneficiaries to identify that the plaintiffs' money had been used to purchase the insurance premium that led to the payout to them. Accordingly, a rule which requires the volunteer to make inquiry in every case, will in many cases involve the volunteer in wasting the money spent on inquiry, as inquiry will usually be fruitless.

However, there will be scenarios where it is appropriate for the volunteer to make inquiries. Where the recipient has grounds to suspect that a gift is not legitimate, inquiries are appropriate.

What conclusions can be drawn from this? In the normal case, refusing the defence will not change the volunteer's conduct. The volunteer will normally not make inquiry into the source of the gift, and there will normally be no need for inquiry. However, from an economic perspective she would still be well-off *if she had* made

inquiries. The gift cost her nothing. Therefore, in the case where inquiry is appropriate, it would seem acceptable to place the burden of such inquiry on the volunteer.

Ultimately though, this only says that the volunteer should bear some risk. Both the change of position defence and the bona fide purchase rule contain limiters to deal with the suspicious scenario.

The next stage of the analysis requires consideration of the position of the plaintiff.

A helpful starting point can be found in an analysis undertaken by Menachem Mautner in a similar context. Mautner has considered the efficiency of the American Uniform Commercial Code (UCC) rules regarding the liability of third parties who purchase property from agents. Although his scenario is not exactly the same, it is sufficiently similar to be useful. Mautner's discussion centres on the scenario of a principal who has entrusted property to an agent, who improperly sells the property to a third party.

In determining whether the UCC rules are efficient, his starting point was the same as ours. He said:

> In cases in which one of the two competing parties could have clearly prevented the occurrence of the conflict ex ante by incurring expenses relatively smaller than the value of the interests at stake, … taking into account the probability of the occurrence of a conflict, priority should be accorded to the other party.[33]

Mautner then considers whether the plaintiff, who has entrusted property to the agent, could have prevented the fraud. Mautner's analysis is almost entirely directed to discussion of the entrustment rules in commercial cases. His commercial case example is where:

> *A*, the manufacturer of a product, appoints *B* to be his selling agent in the market in which *B* is located. *A* delivers goods to *B* on consignment. *B* sells the goods to *C* and absconds with the proceeds. Alternatively *B* pledges the goods to C to secure a loan made by *C* to *B.* [34]

Mautner suggests that the entrusting party is better able to prevent the fraud than the recipient.[35] His argument can be summarized as follows:

---

33 M Mautner 'Eternal Triangles of the Law' (1991) 90 *Mich LR* 95 at 100.

34 Ibid at 129.

35 It is generally the view of the commercial law that 'a person who in relation to the handling of his property reposes confidence in another thereby assumes, in relation to third parties

1. The entruster-entrustee relationship is 'usually intended to be stable and continuous' and 'it is easier for the entrustee to bear the first-starter cost of gathering information for verifying the honesty of the entrustee' than for the recipient.[36]

2. In an 'institutionalized' entruster-entrustee relationship, 'the entruster gains an additional opportunity both to acquaint himself further with the conduct of the entrustee and to foresee potential dishonesty on his part'.[37]

Mautner then considers whether these considerations apply in a non-commercial context. He offers an example of non-commercial entrustment as follows:

> *A*, a university professor, plans to go abroad on sabbatical for a year. *A* entrusts his beloved painting to *B*, his friend, to keep for him for the year. *B* sells the painting to *C*. Alternatively, *B* pledges the painting to *C* to secure a loan made by *C* to *B*.[38]

He suggests that:

> at least the first above-mentioned argument applies to non-commercial entrustments as well: in the non-commercial setting, the entrustment will usually take place between persons having a stable, long-term relationship, so that, as between the entruster and the buyer, the former would usually enjoy a clear informational advantage over the latter in terms of his ability to foresee potential misconduct on the part of the entrustee.[39]

---

dealing with such property, the risk of improper behaviour on the part of the party entrusted.' Roy Goode, 'Proprietary Restitutionary Claims', chapter 5 in W Cornish, Nolan, O'Sullivan and Virgo (eds), *Restitution Past, Present and Future* (Hart Publishing, Oxford, 1998) at 72-3. Goode refers to this as a 'well-established principle'.

36 Mautner, n 33 above, at 131.

37 Ibid at 132.

38 Ibid at 129. The fact that Mautner's examples deal with purchasers and our core question relates to volunteers is not a problem, as the question is whether the entruster can prevent the fraud ex ante. The ability of a defendant volunteer or purchaser to identify the fraud is the same.

39 Ibid at 132. He also raises a further argument at 132 , n 133 for why the entruster is better able to prevent the fraud, saying that 'as a general rule, it is easier to provide information "downstream" than it is to ferret out information "upstream". In our context, it is reasonable to assume that the owner-entruster, who already had possession of the disputed goods prior to their entrustment, would be better located to inform potential purchasers of his interest (by engraving or branding) than would the potential purchaser to discover the existence of the owner-entruster'. This additional argument is unhelpful where the property entrusted is not goods which are capable of physical marking.

If we take this analysis as our starting point, we have two main arguments as to how the entruster is in a better position to prevent the fraud. The following discussion will analyse those arguments.

(a) Making inquiries prior to the commencement of the relationship:

The conclusion that the entruster is in a better position to verify the honesty of the entrustee may be true where the parties already have an ongoing relationship. However, Mautner's analysis breaks down at the point where it is assumed that the entruster will make those inquiries. Economic analysis assumes parties act rationally. In fact, many choices – including the choice of whom to trust - are not made on a rational basis.[40] If the entruster already knows the entrustee, the choice of entrustee is likely to be based on an emotional analysis of the entrustee's character. The entruster is simply unlikely to make stringent checks of the entruster's character, even if to do so is relatively simple.

In any event, in cases where no personal prior relationship exists between the entruster and entrustee, (so a rational decision can be expected) it will frequently not be feasible for the entruster to make meaningful inquiries (contrary to Mautner's assertion). The facts of *Foskett* (a commercial case) are themselves illustrative of this proposition.

The beneficiaries under the trust were individual investors. The investment was apparently reputable. If a prospective investor wished to make inquiries as to the reputation, and prior history of the trustees, they would have had to consider both the individuals who controlled the company, and the company itself. To do this they could have searched regulatory records to identify, for example, a history of personal bankruptcy or involvement in corporate insolvencies. It is unlikely, however, that individual investors could have obtained any information regarding prior criminal behaviour. The process would have to be undertaken by every purchaser, and there were over two hundred purchasers who invested in this project. The huge outlay of time and cost that would be involved is unlikely to produce any substantial probability that material information will be obtained.

A rule which requires individual beneficiaries to make these inquiries is not efficient. A more efficient use of resources is to permit beneficiaries to rely on the regulatory process to exclude any obviously unsuitable individuals.

---

[40] It is also clear that people continually underrate the probability of an adverse event happening to them. See Melvin Eisenberg 'The Limits of Cognition and the Limits of Contract' 47 (1995) *Stanford LR* 211.

(b) Monitoring conduct during the relationship:

Mautner's second proposition asserts that 'the entruster gains an additional opportunity both to acquaint himself further with the conduct of the entrustee and to foresee potential dishonesty on his part'. This proposition is able to be refuted, both generally and in relation to the specific example provided. In Mautner's non-commercial scenario, the entrusting professor requires the services of the entrustee because he, the professor, is about to go overseas. If the professor is overseas, how does he have the opportunity to monitor the entrustee's conduct for dishonesty?

In fact, it is even questionable whether the entruster is generally able to monitor the entrustee's performance sufficiently to identify dishonesty even when the entruster is within the jurisdiction.[41] Most entrustment situations involve fiduciary relationships which encompass more than the bailment scenario envisioned by Mautner. It has been convincingly argued by Easterbrook and Fischel[42] that the fiduciary relationship is inherently one in which it is *not* possible for the entruster effectively to monitor the conduct of the fiduciary, as an essential characteristic of the relationship is 'unusually high costs of ... monitoring'.[43] High monitoring costs result from two main causes: the necessary discretions accorded to the fiduciary and the practical fact that the fiduciary controls the asset. It is difficult for the entruster to monitor the fiduciary's honesty as (1) poor performance may not result from dishonesty, and (2) the fiduciary as controller of the assets is in the best position to hide evidence of fraud in relation to the controlled asset.

This second proposition is important. In fact, not only will the fiduciary be in the position to hide the evidence of fraud, but it would also seem that the fiduciary will be likely to hide evidence of his/her dishonesty. If there is active misappropriation (rather than loss caused by negligence or third party factors) it is unrealistic to suggest that the fiduciary will ordinarily confess the conduct. Once we acknowledge that the fiduciary who misappropriates the principal's assets is also likely to take active steps to hide the evidence of misappropriation, the chance of the entruster adequately monitoring the entrustee's behaviour appears small.

---

41 Irrespective of whether the case is a 'commercial' one or not.

42 FH Easterbrook and DR Fischel, 'Contract and Fiduciary Duty' (1993) 36 *Journal of Law and Economics* 425.

43 Ibid at 427. See also R Cooter and BJ Freedman 'The Fiduciary Relationship: Its Economic Character and Legal Consequences' (1991) 66 NYULR 1045, at 1046-7, where they state: 'because asset management necessarily involves risk and uncertainty, the specific behaviour of the fiduciary cannot be dictated in advance. Moreover, constant monitoring of the fiduciary's behaviour, which would protect the beneficiary, often is prohibitively costly'.

Dishonesty may be relatively simple to spot in Mautner's non-commercial example, but it is suggested that this will not usually be the case.

(c) Identity of entruster:

Mautner's example is confined to voluntary entrustment by competent adults. Mautner's analysis does not work well in a conventional trust scenario, where the entruster is not an adult, but a minor. Apart from any other factors which render monitoring difficult, the minor is certainly not capable of supervising the fiduciary to any sensible degree. In addition, the minor does not choose the fiduciary, and does not have the opportunity to gather information to verify the fiduciary's honesty.

Conclusions so far

At this point it has been demonstrated that Mautner's assertions that the entruster is able to take steps ex ante to prevent the fraud are open to serious question. The analysis above suggests that in the usual entrustment case it is very difficult for the entruster to identify dishonesty both prior to and during the entrustment relationship.

Nonetheless, there would also still be extreme cases where, with hindsight, the risk of fraud was so clear that no reasonable person would have chosen the agent/fiduciary. In that situation, should the plaintiff bear the risk of such poor choice?

Perhaps the answer to this depends on why the plaintiff chose so poorly. Does a poor choice of fiduciary indicate that the principal took the risk of the fiduciary's misconduct? Not necessarily. If the agent hid their evidence of past dishonesty, and the plaintiff trusted the agent, she did not deliberately run an increased risk of fraud.[44] Further, penalizing her for a poor choice of person to trust will not make it easier for her to make a better choice in the future. If, on the other hand, she had the evidence of past dishonesty available to her and chose to entrust her property to the fiduciary despite that evidence, there is more reason to make her bear the loss.

---

[44] Cf the analysis of American academic Emily Sherwin in her article 'Constructive Trusts in Bankruptcy' [1989] *University of Illinois Law Review,* 297 at 356 where she argues that in consensual fiduciary relationships, 'in entering the arrangement, the beneficiary takes the risk that the fiduciary not only will be dishonest but also will be insolvent when the beneficiary asserts a claim'.

The analysis has also demonstrated that there is frequently no easy way for the recipient to identify the fraud. Nonetheless, there are situations in which fraud will be more obvious.[45]

In those cases, it is not inappropriate to require the volunteer to bear the risk of loss, as the cost of any inquiries merely reduces the size of the windfall gain coming to her. Unlike the entruster, she has not earned her wealth, and the removal of the windfall carries no disincentive. This perhaps tips the scales slightly in favour of volunteer liability. The cost to the volunteer of making inquiry, in the necessary case, is cheaper than the cost to the plaintiff of making inquiry.

Other economic considerations

If neither party has a clear advantage over the other, Mautner offers some further arguments. He says:

> In all other cases in which no party enjoys a clear advantage over the other in terms of the ability to prevent the conflict, priority should be granted to the party likely to suffer the greatest loss ex post if he is denied priority and the other party prevails. Additionally, priority rules should be shaped in such a manner as to minimize the parties' resort to the court system and the administrative costs involved in litigation.[46]

I will address the second proposition first.

One possible solution is to adopt a relative or comparative fault liability test, to ensure that the circumstances of each case were carefully analysed to determine whether the loss should fall on the entruster or the recipient. However, such a rule involves substantial adjudication costs in the individual case, and creates a high level of uncertainty in the law. [47]

---

45 Although these are still likely to be infrequent if we factor in the likelihood of the entrustee attempting to hide the proof of her fraudulent behaviour.

46 Mautner, n 33 above at 100.

47 Similarly see the discussion in D Fox, 'Constructive Notice and Knowing Receipt: An Economic Analysis' (1998) 57 Camb LJ 391 at 401 on the costs of possible liability tests for knowing receipt. The discussion compares the likely costs of a 'bright line' rule of liability compared to an 'open-textured standard of liability'. The Privy Council rejected the uncertainty of a relative fault test in *Dextra Bank and Trust Ltd v Bank of Jamaica* [2002] 1 All ER (Comm) 193. See however the analysis by H Dagan 'Restitution and Unjust Enrichment: Mistake' (2001) 79 *Tex L Rev* 1795, who argues that a comparative fault rule along the lines adopted in US tort law is efficient, although he highlights the high adjudication costs of such a rule.

Although this approach would ensure justice in the individual case, the Privy Council's clear rejection of relative fault in *Dextra Bank and Trust Ltd v Bank of Jamaica*[48] renders it unlikely that such a rule will be adopted here. Therefore, this paper will continue with the search for a single liability rule.

Arguably, a rule which definitively prefers the plaintiff over the volunteer (the property approach) is cheaper to administer than a rule requiring consideration of whether a recipient has changed position (the unjust enrichment approach). The change of position rule requires judicial examination of whether, on the facts, the volunteer did acts constituting change of position; and whether in law the acts fit within the test. In our problem case the bona fide purchase test is simple to apply. The rule enables the plaintiff to recover without any need to consider the volunteer's position.

What about the first proposition: that priority should go to the party who will suffer the greatest loss ex post if denied a remedy?

To test this proposition we need to go back to our core example. As discussed earlier, there is no dispute that on the original facts of *Foskett* the volunteer should return the funds. Economic analysis reaches the same answer. The plaintiff, if unable to recover against the volunteer who has changed position, suffers the loss of the full value of the misappropriated funds: in this case $20,000. The volunteer however, suffers no loss if she is required to return the funds. She retains all her own funds, and being required to return the gift only returns her to her original position.

The position is different in the problem case (using the modified facts of *Foskett*) once we assume that the volunteer recipient has changed position as a result of the receipt. The plaintiff beneficiary will suffer the loss of $20,000 if the defence is available to the volunteer who has changed position. What of the volunteer? Until she has spent money consequent upon receipt, she has suffered no loss. However, once she starts to spend money believing in the irreversibility of the gift, any requirement to repay the gift will result in her suffering a loss.

If she has only partially changed position her loss will not be as great as the plaintiff's loss. If, however, she has spent the equivalent amount of the receipt, she will suffer an equal loss to the plaintiff, as she is now worse off to the value of the amount she spent (and cannot recoup).[49]

---

48 [2002] 1 All ER (Comm) 193.

49 The facts of *Foskett* actually raise the possibility of the volunteer incurring significantly greater loss, as the volunteer believed in the correctness of a £1 million payment, as compared to the exact amount flowing from the beneficiaries' money.

If the potential losses to be suffered by the plaintiff and defendant are the same, then a cost-based economic analysis confined to the parties is inconclusive. If the answers are as finely balanced as the above analysis suggests, it is appropriate to consider whether any broad policy imperatives assist in deciding the question one way or the other.

Policy: vulnerability of equitable plaintiffs

One argument in favour of recovery by the plaintiff is that the plaintiff is vulnerable to the abuse by the fiduciary, and is deserving of the law's special protection. This is in fact one of the bases of equitable title - the personal obligation in relation to property is the source of the beneficiary's title.[50]

It could of course be queried whether the 'tenderness' towards such plaintiffs needs to continue in the modern day. The trust is no longer solely a family device to protect the interests of children who are too small to protect themselves. Is it saying too much to suggest that plaintiff beneficiaries of a fiduciary duty are a 'vulnerable class'?

The analysis of Easterbrook and Fischel of fiduciary relationships is important again here. As discussed above, the conclusions drawn are that the fiduciary relationship involves high monitoring costs for the represented party, and a high level of trust on their part. The represented party is unable to cost-effectively monitor the activities of the fiduciary without undertaking the job themselves. This applies whether the plaintiff is a child or a shareholder in a corporation. The identity of the plaintiff is irrelevant. It is the nature of the relationship between the plaintiff and the fiduciary which is critical.

An alternative reason is offered by Emily Sherwin. She argues that the property approach can be justified by reasons of utility. She states that trusts and fiduciary relationships provide a useful social purpose and that preferential treatment of entrusters is necessary to ensure that these relationships continue to be attractive. She argues that 'the assurance of a … remedy that will give the beneficiary a prior right to recover the property or its traceable products… may have a real impact on her initial decision to enter the arrangement.'[51]

---

50 In fact, the trustee's personal obligation as the source of equitable property rights goes some way to explain the fiction. If the beneficiary's title is directly derived from the personal obligation owed by the trustee to the beneficiary, that obligation subsists notwithstanding any change in form of trust property, at least while it is still the trustee's possession.

51 Emily Sherwin 'Constructive Trusts in Bankruptcy' [1989] *University of Illinois Law Review,* 297 at 356.

An Alternative Solution

Both the property approach and the unjust enrichment approach involve one party suffering a loss. We need not accept a rule that requires one innocent party to suffer if alternative rules can be identified.

Mautner had a third proposition not discussed earlier. He said:

> The 'entruster can protect himself against misconduct… by insisting upon a guarantee of the entrustee's liabilities toward the entruster by such means as a cash deposit, guarantee or security interest.[52]

This part will analyse a bond proposal in more depth. It will be demonstrated that although a version of the bond process may have some attractions, the position is not as simple as Mautner suggests, particularly in relation to commercial entrustments. Other alternatives are possible and this part will consider whether an insurance regime or compensation scheme would also serve a satisfactory purpose.

Essentially however, if a viable alternative model can be found, then the dispute between the property and unjust enrichment approaches is easy to solve. If the interests of the entruster are protected through an alternative model, the change of position defence can be made available to the volunteer recipient.

A bond framework

Initially, consideration would need to be given to how a bond system would work. Important factors are: how is the system to be implemented, what relationships are to be governed by the system, and how would the bond be set?

Implementation issues

Will the system work if entrustees are left to protect themselves by 'insisting' upon a bond? Given the earlier analysis of this problem, it appears unlikely that an entruster who reposes actual trust in the entrustee will insist upon a bond from the entrustee.[53] The existence of trust will itself alter the entruster's behaviour. Therefore it would seem to be necessary for any bond requirement to be imposed by legislation.

---

[52] Mautner, n 33 above at 132.

[53] This is particularly true in the non-commercial case. It is not likely that the entrusting professor in Mautner's own non-commercial example would say 'I trust you and want you to mind my painting, but give me a bond anyway'.

If a bond is a required, rather than agreed, feature of entrustment relationships, there must be an agency with the power to set the bond. In non-commercial cases, it could be appropriate to give jurisdiction to a court or tribunal. A model for this already exists in some jurisdictions in relation to the appointment of administrators of deceased estates.[54] It is also common in American for courts to have the authority to impose a bond as part of the appointment process for conservators of a minor or incompetent's property. [55]

In more commercial cases, for example where entities take public money for investment, jurisdiction to set and administer the bond could be given to the relevant regulatory body which supervises public investments. In Australia, securities dealers were required to provide a security bond of $20,000 under the Corporations Act prior to the introduction of the financial services regime. The administration of the bond requirement has been the work of the ASC and now ASIC. This requirement is ongoing on a transitional basis,[56] until the Government finalises the compensation arrangements under the new s 912B of the Corporations Act 2001 for the financial services sector.

Serious consideration would need to be given to the types of relationships governed by the bond requirement. If the bond requirement is limited to parties who are already required to submit to the jurisdiction of the court[57] or regulatory authority,[58] the cost increase will be relatively small, as the relevant agency/court and the parties have already invested time and cost in the process. Adding a bond component to the process is unlikely to be disruptive.

---

54 For example: in Australia, administration bonds are required in New South Wales: section 64 of the *Wills, Probate and Administration Act 1898* (NSW) and in Tasmania: section 25 *Administration and Probate Act 1935*. In Ireland, s 34 of the *Succession Act 1965* requires a bond to be given by 'every person to whom a grant of administration is made'. See also s 14 *Administration of Estates Act 1959*

55 There are too many provisions to warrant mentioning all of them. The exact terms of bonds vary across States. Some States, such as Arizona (see Arizona Revised Statutes s 14-5411), Alaska (Code of Alaska s 26-2A-139), Kansas (KSA s 59-3069) make a bond compulsory, although the Court may be able to waive or reduce the bond. In other States, the Court may order a bond, without being required to: see, for example, District of Columbia Code s 21-2058; Hawaii Revised Statutes, s 560:5-415; Idaho Code s 15-5-411. In Maine, the Court is not required to set a bond until the value of the estate concerned is $25,000, although it may exercise a discretion to set a bond where the estate is less than that: (Maine Revised Statutes, s 5-411A.

56 Currently until 31 December 2006: see ASIC IR 06-21.

57 For example, when a person is applying to become administrator of an estate.

58 Eg entities governed by corporate regulators. In Australia this is ASIC.

However, this solution would mean that entrustment cases that do not currently require prior court/authority approval will continue to fall outside the bond requirements. This would mean, for example, that private trusts would not be within the bond regime.[59] If the bond process were to be expanded to cover as many fiduciary/entrustment relationships as possible, the relevant authority will need to be provided with more resources to cope with the increased workload or long delays would occur. Another possibility would be to require a bond for only the relationships in which the entruster is most vulnerable (for example minors).

How would the bond work?

Even in relationships which are to be governed by a bond requirement, a further choice must be made. A bond could be compulsory in all regulated relationships, or it could be set by the exercise of discretion.[60] A compulsory bond could be dealt with at an administrative level, and the process does not require the consideration of individual cases. However, cases where the entrustment is for a short period or the assets entrusted are of low value may not warrant the costs of the bond process.

A bond process which is discretionary is more expensive to administer. The decision maker must consider each individual application to determine if a bond is required. It is also inevitable that discretion will occasionally be exercised wrongly, with the consequence that wrongdoing occurs in a case where it was deemed not necessary to set a bond. Cases have arisen in America in which a Court exercised its discretion not to set a bond, or to set a small bond, and the fiduciary misappropriated the funds under their control.[61] In such a scenario, the costs of the bond process are wasted.

One possible solution would be to establish a system with the bond being compulsory once a certain value of assets are under entrustment. The asset value at which the relationship becomes regulated would be necessarily be set rather arbitrarily, but this would be a way to manage the process.

---

59 In Australia, another unregulated area is self-managed funds. The financial assistance provisions in Part 23 *Superannuation Industry (Supervision) Act 1993* do not extend to self-managed funds.

60 Or the decision maker could have discretion to waive a bond.

61 A common feature of these cases is that the fiduciary was a family member, typically a parent, and the plaintiff was a minor. See, for example, *Reed v Valley Federal Savings and Loan Company* 655 SW 2d 259 (CA Texas, 1983). The guardianship estate was worth approximately $35,000 but a bond of only $5000 was set by the court. The guardian dissipated the assets of the estate.

Setting the amount of the bond

Once a model is established, decisions must be made regarding the size of any bond required. Existing private bond models demonstrate various possibilities,[62] but have the common feature that the bond is at least the value of assets under entrustment. This makes sense, as the bond will not act as a deterrent to misappropriation unless the bond is at least equal to the value of assets under control. In Australia, the $20,000 bond set under the old Part 7 of the Corporations Law has been recognized to be insufficient to compensate investors for loss.[63]

Even if the bond is set at a sufficiently high level to deter fraud, some other issues remain. In particular, how is the bond to be secured?

If the entrustee is required to put up her own assets to meet the bond, this should function as a deterrent. However, this model could be a problem if the entrustee does not have sufficient assets, or if the entrustee's assets are jointly owned with others. Additionally, if the entrustee's main asset is a matrimonial home, and is used for security, the home could not be sold without the security being released. It would be understandable if the entrustee's spouse were reluctant to use the home as security for the entrustment relationship.

These considerations could be addressed by a model which did not require the fiduciary to post the entire bond personally. Instead, the fiduciary could be permitted

---

62 In New South Wales, for example, the bond is normally to be equal to the value of property of the deceased: section 65 *Wills, Probate and Administration Act 1898* (NSW). In Ireland the bond is to be 'double the amount at which the estate of the deceased is sworn': s 34(2)(a) *Succession Act 1965*. American models differ. There are multiple models. The one which most often appears requires 'the amount of the aggregate capital value of the property of the estate in the conservator's control, plus one year's estimated income, and minus the value of securities' deposited with the Court and land which the conservator cannot sell without Court authority. This model is used, for example, in Alaska (see Code of Alaska, s 26-2A-139(a)); Arizona: (Arizona Revised Statutes s 14-5411); Colorado: (CRS 15-14-415); District of Columbia: (s 21-2058) and Hawaii (s560:5-415). The calculation required in Kansas is for '125% of the combined value of the tangible and intangible personal property in the conservatee's estate and the total of any annual income from any sources which the conservator may be expected to receive on behalf of the conservatee, minus any reasonably expected expenses': KSA s 59-3069 (b).

63 See 'Compensation for Loss in the Financial Services Sector - Issues and Options', Department of Treasury, 6 September 2002, para 168; Australian Securities Commission Licensing Review Report Investment Advisory Services "Good Advice", November 1995, p 40 and P Latimer, 'Compensation of Investors for Failure of a Securities Industry Licensee' (1997) 14 *C&SLJ* 495 at 498, summarising the ASC view.

to find sureties. Problems exist with this model too. Once it is not the fiduciary's own assets being applied to the bond, the deterrent feature of the model disappears. The fiduciary no longer bears the cost of their misappropriation. Instead, if the fiduciary commits wrongdoing, the effect of this model is simply to shift the burden of loss from the fiduciary's wrongdoing to a new third party: the surety. Commercial sureties could be found, but the entruster would still need to guarantee the funds to the provider of the bond. [64]

A bond is also unsuitable for public/commercial relationships, such as the public investment that featured in *Foskett v McKeown*. There is no way that a bond process could adequately protect against the potential large scale misappropriation in such a case. As discussed above, this was recognized as a problem with the security bonds for Australian securities dealers.

Is insurance a better response?

Particularly where substantial protection is required, insurance companies would seem better able to deal with the uncertainties inherent in the bond process. Insurance companies are used to assessing risks of adverse events and setting premiums at an appropriate level.

Is it feasible to require compulsory insurance cover?

In Australia, the question of an appropriate compensation regime within the financial services sector is under consideration by the Department of Treasury. Section 912B of the Corporations Act requires financial services licensees to have 'compensation arrangements' in place. The Position Paper released by the Department of Treasury preferred an insurance model over both security bonds and a statutory compensation scheme.[65]

Appropriate insurance would need to include fidelity insurance as a requirement rather than an optional extension. As dishonesty is normally excepted from standard professional indemnity insurance, a full compensation scheme through insurance requires fidelity insurance. Treasury has not indicated whether fidelity insurance will be mandatory, and the availability in the market of such insurance was not clear.

---

64 In Treasury's Issues and Options Paper, the view was taken at para 168 that 'a substantial bond may prove an impediment to those wanting to start in the financial services sector in a small way'.

65 In England, there is a financial services compensation scheme for consumers and small business. In Australia the National Guarantee Fund operates for market transactions.

In the commercial sector (such as financial services) where annual licences are required to be issued, the requirement for insurance could be incorporated into the approval process with only a moderate level of difficulty. There would be a certain amount of additional administrative work and cost in checking that insurance has been provided. For the regulator, this will be less than the work of administering a bond process.

In the private sector, it is harder to establish an insurance requirement. If, for example, the establishment of the entrustment requires legal advice, the adviser can inform of the requirement for insurance. However, in entrustment relationships which are established informally, it is much harder to identify an appropriate insurance regime.

What are the problems with an insurance system?

The most obvious question must surely be 'who pays the premiums?' A fiduciary acting gratuitously could not be expected to pay, and a professional fiduciary will charge the costs of insurance on to client. Ultimately, therefore, the cost of insurance will be paid for by the represented party.

In fact, if the cost of insurance is borne by the represented party, either directly or indirectly, this means that the represented party is paying to ensure that her choice of fiduciary is a good one. This method actually requires the plaintiff entruster to bear the risk of her fiduciary's fraud, through the payment of insurance premiums. It ensures that the fund is protected in the event of misappropriation, but does not place the risk of loss onto any innocent party.

More problematic could be the impossibility of ensuring that insurance is maintained. Treasury noted that 'a requirement that insurers automatically alert ASIC to the cancellation or non-renewal of professional indemnity/fidelity cover… would assist' and that there is no current obligation to provide ASIC with a certificate of currency.[66] If the compensation method is to be offered via insurance, this obligation seems necessary.

An insurance model also encounters the moral hazard problem. Moral hazard describes the problem where the dishonest fiduciary is not deterred from dishonest conduct, because they know that the loss is borne by an insurance company, and not the plaintiff. Moral hazard also arises in the security bond scenario where the security for the bond is provided by a corporate surety.

---

[66] Department of Treasury, 'Compensation for Loss in the Financial Services Sector - Position Paper,' 24 December 2003, chapter 3, para 160. Another concern is whether the insurer should be able to cancel the insurance.

Perhaps the risk of moral hazard is one that must be present in order to adequately protect entrusters. Furthermore, the type of misappropriation under discussion will almost inevitably also constitute a crime and the entrustee will be subject to criminal charges. If the entrustee is not deterred by the possibility of criminal charges, it seems unlikely that the moral dilemma would impact substantially upon their decision making processes.

There is, however, a positive feature of insurance, in that parties who wish to comply and put risk management procedures in place will ultimately benefit through reduced premiums.

Ultimately, if insurance is adopted as the primary compensation method for financial services in Australia under s 912B, time will determine if the correct option was chosen.

Statutory compensation schemes or industry fidelity funds

Treasury also considered the possibility of a statutory compensation scheme. Compensation schemes are funded by levies on licensed parties. The National Guarantee Fund operates in Australia, and there are overseas examples for stock exchange operators. Fidelity funds can also exist within certain professions or industries. For example, fidelity funds exist in Australian states to protect solicitors' clients from loss occurring through default or dishonest failure to account.

In England, the response to the need for investor compensation in the financial services sector was to establish the Financial Services Compensation Scheme (FSCS). The scheme is funded by levies, and provides compensation to consumers and small business in the event of insolvency of a provider.[67]

Costs of statutory compensation scheme/fidelity fund

Treasury determined that 'the administration of a statutory fund would involve significant costs'.[68] It also considered that there was insufficient evidence of loss in the financial services sector to warrant the costs of implementing a compensation scheme, although it indicated that the question of such a scheme could be reconsidered if more data regarding losses became available in the future.

---

67 See the FSCS homepage: www.fscs.org.uk for further information.

68 Position Paper, Chapter 3, at para 178, the anticipated costs were the cost of developing a scheme; the cost to industry of additional returns necessary to set levies; the cost of levies; the administration of the scheme and paying claims.

In England, the costs of establishing the FSCS were recovered from levies, and the scheme's management expenses are funded by levies on an ongoing basis.[69]

Advantages and Disadvantages

Many of the effects of this scheme are the same as insurance model. As with insurance, this model creates a moral hazard issue, as the operator knows that innocent investors will not suffer from their behaviour. However, it does provide protection to entrusters. Further, the costs of the system are borne by entrusters, which appears appropriate.

One substantial administrative benefit that occurs if a statutory scheme is implemented, is that the problem of maintaining insurance cover ceases to be a problem. Entities are levied on a yearly basis, and no further checking or reporting of their status is necessary. This eliminates one of the more difficult administrative problems that arises with an insurance model. It could also be anticipated that cost efficiencies could occur if all processes are conducted by associated government organizations.[70]

However, unlike an insurance system there is no inherent incentive for regulated entities to improve their services as this will give them no reduction in levies.[71]

Some recommendations

The analysis demonstrates that the issues are different in private fiduciary relationships compared to public/commercial entrustment scenarios. The level of regulation and questions of scale in the commercial scenarios make insurance or fidelity funds more feasible than bonds. Although a statutory compensation scheme may be expensive to establish, it would eliminate any concerns regarding the insurance market, and the capacity of parties to obtain (and maintain) insurance.

However, in the private arena, the lack of regulation makes any change significantly costly to establish. Both fidelity funds and insurance models appear infeasible in an arena where no registration, licence or ongoing professional qualification is required. The bond process is the most feasible for private entrustments, but would entail

---

69 FSCS also generates income from recoveries against parties in default. Financial information is available in the FSCS Annual Reports, available on their web site.

70 For example, in England, the Financial Services Authority (the regulatory authority) invoices regulated licensees for levies under the Financial Services Compensation Scheme.

71 Although if there are no or few claims, the fund may be self-funding after a period of time.

substantial costs to implement. Perhaps a workable model would involve compulsory bonds only in situations where property was entrusted on behalf of minors.

Conclusions

It has been demonstrated that there is no conclusive result from law and economics analysis to justify shifting from the property approach to the unjust enrichment approach. Neither the beneficiary nor the innocent volunteer is clearly the least cost avoider who can take precautions to identify and prevent the fiduciary fraud. Further, the loss to be borne by either party ex post is the same. The analysis also discussed how broader policy arguments tend towards maintaining the property approach.

Ultimately, it is suggested that for many relationships, the answer lies elsewhere, through the use of compensation models which ensure that an entruster's interests are protected in the event of the entrustee's fraud. If the entrustee's interests are protected through a true compensation model, there is no reason to deny the change of position defence to the volunteer.

If there is no appropriate compensation model, which may be the case with private entrustments, the property rule should remain.

THE BARRIERS TO THE ENFORCEMENT OF FOREIGN JUDGMENTS AS OPPOSED TO THOSE OF FOREIGN ARBITRAL AWARDS

Juliane Oelmann[*]

Introduction

International trade and commerce has significantly increased during the last decades. Along with this development comes a boost in international litigation. Thus, several jurisdictions can become involved in the judiciary process. Hence, becoming familiar with the relevant procedures in international litigation is more relevant than ever before.[1] Especially the question of where and under what conditions a foreign judgment can be enforced in a second forum is meaningful in this context. Therefore, this article will focus on that particular issue. Precisely, the article analyses the departure grounds existing for the courts of a second forum to refuse enforcing a foreign judgment. The major attention will be drawn to the refusal ground of public policy.

After an analysis on when this barrier may operate, a comparison with the enforcement process of foreign arbitral awards, also with reference to the departure ground of public policy, will follow.

Concerning the enforcement of foreign judgments, the Council Regulation (EC) No 44/ 2001, known as the Brussels I Regulation, will be referred to. This regulation essentially replaced the Brussels Convention on Jurisdiction and the Enforcement of Judgments in Civil and Commercial Matters (1968)[2] for all Member States of the European Union except for Denmark. For the latter the Convention of 1968 remains in effect.

---

[*] Juliane Oelmann, First State Exam at University of Bonn, Referendarin, LLM (Bond).

[1] Behr, Volker, 'Enforcement of United States Money Judgments in Germany', (1993-1994) 13 *Journal of Law and Commerce*, 211, 211.

[2] Hereinafter referred to as Brussels Convention.

Although there are some structural changes coinciding with the transformation of the Brussels Convention into a regulation,[3] most provisions are closely related, and the main structure as well as the fundamental principles and goals of these two instruments remain identical. The preliminary considerations of the Brussels I Regulation confirm this by expressing the desire for continuity between the Brussels Convention and the new Regulation. The demanded continuity also implies the necessity to interpret the provisions identically. Hence, the case law rendered under the Brussels Convention remains very important in terms of interpreting the Brussels I Regulation.[4] Therefore, when this comment refers to the Brussels Convention those considerations are as much relevant for the Brussels I Regulation.

With reference to the enforcement process of foreign arbitral awards the note bases upon the United Nations Convention on the Recognition and Enforcement of Foreign Arbitral Awards (1958).[5]

Contextually with this comment it is of particular interest for Australia as a Member State of the Hague Conference on Private International Law to mention the Hague Convention on Choice of Court Agreements. It was concluded in June 2005.

The Convention represents a culmination of more than a decade of negotiations concerning jurisdiction, recognition, and enforcement of judgments in civil and commercial matters. It is designed to promote international trade and investment through enhanced judicial co-operation.[6]

So far, Australia has been dealing with the enforcement of foreign judgments in the *Foreign Judgments Act 1991* (Cth).[7] The Act establishes a statutory scheme under which judgments of non-Australian courts can be enforced in Australia. Despite its advantages, there are disadvantages inherent in the Act. It is limited to money judgments, to superior courts (with a few exceptions), and to specified countries.[8] The Schedule of the Foreign Judgments Regulations 1992 (Cth) sets out those courts and

---

3 Kennett, Wendy, 'Current Developments Private International Law The Brussels I Regulation', (2001) 50 *International and Comparative Law Quarterly*, 725, 725 f.

4 Pontier, Jannet A., Burg, Edwige, 'EU Principles on Jurisdiction and Recognition of Judgments in Civil and Commercial Matters according to the case law of the European Court of Justice', 2004, p. 1f.

5 Hereafter referred to as New York Convention.

6 Brand, Ronald A., 'Introductory Note to the 2005 Hague Convention on Choice of Court Agreements', (2005) 44 *International Legal Materials*, 1291.

7 Hereafter referred to as Act.

8 Downes, Kylie, Hodge, Michael, 'Enforcing foreign judgments in Australia and Australian judgments in foreign jurisdictions', PROCTOR 24 (2) 2004, 27.

countries. Under the Convention these disadvantages do not occur to the same extent. The only resemblance is that the Convention is confined to the Contracting States.

> With relation to rules concerning recognition and enforcement of foreign judgments the Convention contains special provisions. Those resemble the ones under the Brussels I Regulation and the New York Convention. E.g., under Article 9 (e) recognition and enforcement can be denied if such would be 'manifestly incompatible with the public policy of the requested State'.

The public policy clause has always been a point of some concern. Thus, it might prevent some states from becoming parties to the Convention. However, due to the existence of these parallel provisions, the concerns of the parties to the Hague Conference could be relieved. This could be achieved by looking at the experience with the public policy clause under the Brussels I Regulation and its precursor. Firstly, the interpretation of public policy could follow the ideas under the Regulation. Thus, the latter could serve as a guideline. Secondly, the case law dealing with public policy under the Regulation demonstrates that the concerns are not as major as they might be perceived as.

The Necessity of Enforcement in a Second Forum

The necessity to enforce a judgment in a country other than the one where the judgment was delivered can arise from numerous reasons.

One reason why the court of a second forum might have to pay attention to a foreign judgment is that a defendant who succeeded in the original action in a foreign state could bring forward the foreign judgment in his favour to defend himself against a new action initiated by the original unsuccessful plaintiff in the enforcing forum.[9] Another reason is that an original judgment in favour of the plaintiff can be utilised to found a new action in the enforcing forum for the purpose of enforcement.[10] The literature summarises these reasons as using the foreign judgment either as a 'sword or a shield',[11] meaning either as an offensive or a defensive device.

Another important consideration leading to enforcement in a second forum is the question where the judgment debtor has assets. There will not be a problem if he has assets in the forum of the original court. However, if this is not the case the judgment will only be effective if the defendant either voluntarily complies with it or if it is

---

9 Caffrey, Bradford A., 'Enforcement of Foreign Judgments', 1985, p. 149.

10 Ibid.

11 Caffrey, above n 9, p. 150.

enforceable in the country where he has assets, which logically will be a foreign country.[12]

Prerequisites to the Enforcement of a Foreign Judgment

Enforcing a foreign judgment in a second forum becomes relevant because there is no automatic enforcement process.[13] Rather than that, the party seeking enforcement has to satisfy certain requirements. Those are set out in the Brussels I Regulation.

Under Article 38 paragraph 1 of the Brussels I Regulation the interested party has to lodge an application to declare the judgment enforceable. As per Article 39, the application must be submitted to a competent court or authority, which is listed in the Annex to the Regulation. Article 40 paragraph 2 requires the applying party to provide an address for services of process within the area of jurisdiction of the court applied to. According to Article 40 paragraph 3 the applicant must attach certain documents to the application. These documents, mentioned in Article 53, are a certified copy of the judgment and a certificate as per Article 54.

In conclusion, there are not many procedural obstacles to overcome in the process of the application to enforce a foreign judgment. Therefore, difficulties in the procedure can only result from other provisions of the Brussels I Regulation. The examination of these provisions will follow in a later part of this article.

Departure Grounds Under the Brussels I Regulation

Article 41 of the Brussels I Regulation embodies the rule. According to this rule, a foreign judgment shall be declared enforceable immediately on the completion of the formalities mentioned in Article 53, prohibiting any review of the judgment under Article 34 and Article 35. The aforementioned provisions contain the departure grounds for the courts to refuse enforcement. Consequently, checking the judgment under those articles is only possible when either of the parties lodges an appeal to the competent authority.[14] This is set out in Article 45 paragraph 1 of the Regulation.

---

12 Pryles, Michael, Waincymer, Jeff, Davies, Martin, 'International Trade Law', 2nd edition, 2004, p. 462.

13 Pryles, above n 12, p. 520.

14 Kennett, above n 3, 734.

Concluding from this, refusing the enforcement is to be handled as the exemption under the Brussels I Regulation[15].

However, Article 45 states that if an appeal is lodged the court has to refuse or revoke a declaration of enforceability on one of the grounds listed in Articles 34 and 35. Article 45 refers to those provisions under which the recognition of a foreign judgment can be refused and declares them applicable to the enforcement of a foreign judgment, too. Article 34 contains three departure grounds and Article 35 paragraph 1 establishes a further one. The list of the departure grounds within the sphere of the Regulation is exhaustive, meaning that no other reason will be accepted, no matter how profound it may be.[16]

One commonality of all departure grounds of the Regulation lies in the fact that they do not provide any discretion for the courts. Hence, if a court comes to the decision that the elements of the provisions are satisfied it has no choice but to refuse or revoke a declaration of enforceability. The wording of the Articles 34 and 35 implies this by saying 'shall not be enforced'.[17]

Another common feature is that either the judgment debtor, or the enforcing court sua sponte can set the departure ground in motion.[18]

The Public Policy Exception

Article 34 paragraph 1 of the Brussels I Regulation contains the public policy clause. According to the provision a court of the second forum must refuse the enforcement if such would be contrary to the public policy in the enforcing country.

Defining Public Policy

In order to understand when the public policy exception can be effectively used as a defensive device, it is essential to define its exact meaning. This however is where the problem begins. In fact, the public policy exception is considered as the least well-

15 Ibid.

16 Briggs, Adrian, 'Civil Jurisdiction and Judgments', 3rd edition, 2002, p. 440, para. 7.12.

17 Kaye, Peter, 'Civil Jurisdiction and Enforcement of Foreign Judgments', 1987, p. 1436.

18 Kreindler, Richard H., 'Transnational Litigation: A Basic Primer', 1998, p. 234.

defined departure ground.[19] Because of that, some legal writers and authorities fear an excessive abuse of the clause.[20]

Concerning an attempt of defining the meaning despite the arising difficulties, the first thing one needs to do is abandon the belief that there is a uniform or global understanding of public policy. This is not the case. Therefore, it is even more challenging to attribute a precise meaning to the provision. Owing to a lack of a universal definition, each jurisdiction must be analysed individually with regard to how it interprets the term.[21]

Another reason why it is difficult to approach the term of public policy is its character of a value phrase.[22] Defining value phrases is on the border to impossibility. Firstly, the priority of values differs from jurisdiction to jurisdiction. What might be of great interest and importance in one jurisdiction might be irrelevant in another one. Secondly, a value term implies not being static or permanent. It is rather subject to sociological and economic changes deriving from internal or external influences. Hence, the foundation of a public policy rule alters from time to time. Something that was considered inappropriate in the past might become generally accepted in the future or in the present and vice versa. Owing to the changes the courts are forced to redefine the term of public policy each time a party tries to invoke it.[23]

However, there is one feature to the clause, which all Member States of the Brussels I Regulation and outsider states with similar devices in an international treaty agree upon. That is that the public policy provision is to be handled as an escape clause to review only such cases that touch upon severe national concern. In other words, the clause should only operate in exceptional cases.[24] Because of that, there is a general agreement under which circumstances the reservation clause cannot be raised successfully. That is the case when there is only a variance –be it minor or major- between the law of the initial forum and that of the forum where enforcement is sought. To frame it differently, if the court of the second forum would obtain another

---

19 Pittman, Jonathan H., 'The Public Policy Exception to the Recognition of Foreign Judgments', (1989) 22 *Vanderbilt Journal of Transnational Law*, 969, 970.

20 Minehan, Karen, 'The Public Policy Exception to the Enforcement of Foreign Judgments: Necessary or Nemesis?', (1995-1996) 18 *Loyola of Los Angeles International and Comparative Law Review*, 795, 796.

21 Kreindler, above n 18, p. 239.

22 Caffrey, above n 9, p. 212.

23 Caffrey, above n 9, p. 213.

24 Reuland, Robert C., 'The Recognition of Judgments in the European Community: The Twenty-Fifth Anniversary of the Brussels Convention', (1992-1993) 14 *Michigan Journal of International Law*, 559, 591.

judgment than the court in the initial forum did, this does not suffice to invoke the provision.[25] Consequently, only a fundamental infringement can cause the pro-enforcement rule to give way to the exception.

Several jurisdictions attempted to provide a guideline of what such a fundamental infringement would have to be like.

This means that, under German law a court has to refuse enforcement in extreme cases in which a breach of the fundamental values of German law or of German state policy is threatened.[26] The German Code of Civil Procedure, Zivilprozessordnung, embodies this rule in section 328 number 4. It means that the enforcement of a foreign judgment may not touch upon the basic principles of German law or when enforcement would be unbearable.[27] Such a situation would especially arise if the enforcement would violate basic rights.[28]

In the United States a court has to deny enforcement of a foreign judgment if such enforcement 'injures the public health, the public morals, the public confidence in the purity of the administration of the law, or undermines that sense of security for individual rights, whether of personal liberty or of private property, which any citizen ought to feel'.[29] Under English law a foreign judgment must 'offend against English concepts of liberty and humanity', such as a judgment that was supposed to be used to enforce a contract dealing with slavery and where the transaction upheld under the foreign judgment is prejudicial to the interests of the United Kingdom or its relations with friendly countries.[30]

The courts of the countries use different words in their definitions. Consequently, there is no identical definition. However, all these 'definitions' reveal many similarities. The most striking commonality is that all jurisdictions seem to require a breach so profoundly in its scope that it would totally undermine basic rights in case of an enforcement of the judgment.

---

25 Fine, J. David, 'Defences Against Recognition or Enforcement of Interstate or Foreign Judgments', (1987) 61 *Australian Law Journal*, 350, 361; Pittman, above n 16, 970.

26 Behr, above n 1, 223.

27 Thomas, Heinz, Putzo, Hans, 'Zivilprozessordnung Kommentar', 27. Auflage, 2005, p. 537.

28 Above n 24, p. 538.

29 *Somportex Ltd. v. Philadelphia Chewing Gum Corp.*, 318 F. Supp. 161, 169 (E.D. Pa. 1970).

30 *Foster v. Driscoll* [1929] 1 K.B. 470.

Procedural and Substantive Public Policy

Whether public policy will be affected when a foreign judgment is sought to be enforced in a second country can be judged from two different angles. On the one hand, there is procedural public policy and on the other hand, there is material or substantive public policy. The former aims at regulating the processes used by the rendering forum, starting with the initial service and finishing with delivering a judgment and the opportunity of an appeal. Thus, procedural public policy wants to ensure due process. It arises from the varying civil procedure systems of the Member States. In contrast to that material public policy does not examine the procedure followed in the foreign country itself. It rather focuses on the effects of enforcing the foreign judgment in the second forum. A breach of material public policy occurs when the court of the country of origin has applied a law, which establishes such a great affront to principles of law of the forum where enforcement is sought that would make it impossible to enforce the judgment there.[31]

Case Law Dealing with the Public Policy Exception

As was seen in the antecedent passages defining the notions of public policy remains mainly to the national courts. Hence, it is very likely that each court will apply the reservation clause differently. Then there is a growing fear that the clause could take on a life of its own and turn into a catchall provision instead of serving the purpose of a 'ground of last resort'[32] only.[33]

Whether this fear is justified will be examined now. The article will therefore examine the question of how free the Member States of the European Union really are in defining the term of public policy. In order to do so, it is essential to consider the case law of the European Court of Justice. The court dealt with the aforementioned question in the case of *Krombach v. Bamberski.* There, the court profoundly discussed the problems that are likely to arise under Article 27 paragraph 1 of the Brussels Convention, now Article 34 paragraph 1 of the Brussels I Regulation.

---

31 Newton, Justin, 'The Uniform Interpretation of the Brussels and Lugano Conventions', 2002, pp. 376 f.

32 Pittman, above n 19, Jonathan H., 970.

33 Vest, Loudon Lindsay, 'Cross-Border Judgments And The Public Policy Exception: Solving The Foreign Judgment Quandary By Way of Tribal Courts', (2004-2005) 153 *University of Pennsylvania Law Review*, 797, 799.

Facts of Dieter Krombach V. André Bamberski

Dr Krombach, the defendant, is a German physician. In 1982 he gave an injection to the 14-year-old French daughter of his girlfriend. Both had stayed at his home in Germany. The girl, Kalinka Bamberski, died on the day after she had received the injection. Hence, criminal proceedings against Dr Krombach were initiated in Germany. However, due to a lack of evidence these proceedings were discontinued. This is why the victim's father, Mr Bamberski, the plaintiff, started criminal proceedings against the defendant in France. Krombach was committed for trial in the Paris criminal court on a charge of voluntary homicide. In the course of these proceedings the plaintiff also claimed damages in an adhesion action for pain and suffering. The French Criminal Court claimed it had jurisdiction over the dispute because of the victim's French nationality. The court ordered the defendant to appear in court, which he did not do. Instead, his two lawyers tried to appear on his behalf. However, the court ruled that under Article 630 of the Code of Criminal Procedure that because the accused had failed to appear in person no counsel could represent him. Hence, the written statements of the defendant's lawyers were not admitted in the trial. Eventually, the court delivered a judgment in absentia against Krombach pronouncing a criminal sentence and at the same time awarding damages to the plaintiff. After these judgments had been rendered Mr Bamberski applied to the German District Court, the Landgericht, to enforce the judgment for damages in accordance with the Brussels Convention. The court granted the application. In response to that Krombach appealed to the Oberlandesgericht, the Appellate Court, and ultimately to the Bundesgerichtshof, the Federal Supreme Court. He asserted that the French judgment against him was rendered in violation of Article 27 paragraph 1 of the Brussels Convention, thus being contrary to the German public policy if enforced in Germany. The Federal Supreme Court was sympathetic to the defendant. Yet, the court was not certain about the impact of the Brussels Convention. It therefore asked the European Court of Justice for a preliminary ruling on the interpretation of Article 27 paragraph 1.

Issues of the Krombach Case

With the submission of Germany's Federal Supreme Court pending at the European Court of Justice the judges in Luxembourg had to address the following questions:

Firstly, it was relevant to determine whether a court of the state in which enforcement is sought can take into account under Article 27 paragraph 1 of the Brussels Convention regarding a defendant domiciled in that state the fact that the court of the state of origin based its jurisdiction solely on the victim's nationality.

The second question was whether the state in which enforcement is sought can take into account under the public policy clause regarding a defendant domiciled in that state the fact that the court of the state of origin refused to allow the defendant to have his defence presented unless he appeared in person.

The court had to reply to these questions in the glare of the relationship governing the Brussels Convention and the ECC Treaty. Since the former was based on Article 220 of the Treaty (now Article 293 ECC), the European Court of Justice has consistently held that basic rights are an essential part of the general rules of law whose observance the court has to supervise. In order to do so, the court refers to the constitutions of the Member States as well as to guidelines provided by international treaties for the protection of human rights on which the Member States have collaborated or which they have signed. Therefore, the European Convention on Human Rights is significant in this context.[34] Another important aspect for the court to take into account is the aim of the Brussels Convention. One of its major objectives is to facilitate to the greatest possible extent the free movement of judgments and a very quick enforcement process within the Member States.[35] The competence of the European Court of Justice to interpret the Convention derives from the adoption of Protocol 3 on June 3, 1971.[36]

Decision of the Krombach Case

In general, the court decided that the Contracting States remain free in principle to determine the contents of public policy according to their own national concepts. This statement implies that a domestic standard may be applied. Simultaneously, the court ruled that the boundaries of that concept are a matter of interpretation of the Brussels Convention.[37] Therefore, although the European Court of Justice must not determine the details of the public policy clause, it is still necessary for the court to review the limits within which the national courts may take recourse to the public policy reservation when trying to refuse the enforcement of a judgment delivered in another member state.[38]

With reference to the first question the court came to the conclusion that the court of the state in which enforcement is sought cannot take into account under Article 27 paragraph 1 of the Brussels Convention the fact that the court of the state of origin

---

34 Lowenfeld, Andreas F., 'International Litigation and Arbitration', 2nd edition, 2002, p. 459.

35 Lowenfeld, above n 34, 458.

36 Minehan, above n 20, 808.

37 *Krombach v. Bamberski*, C–7/98, ECR 2000, I-1935, para. 22.

38 *Krombach v. Bamberski*, C–7/98, ECR 2000, I-1935, para. 23.

based its jurisdiction solely on the nationality of the victim of a criminal offence.[39] Article 28 of the Brussels Convention confirms this finding.

As for the second question, the European Court of Justice responded that the national courts are allowed to take recourse to the public policy clause if the state of origin refused to allow the defendant to have his defence presented unless he appeared in person.[40]

Contextually with the second issue of *Krombach v. Bamberski* Articles 29 and 34 of the Brussels Convention become relevant. Those provisions prohibit any review of the foreign judgment as to its substance. Therefore, a court of the state in which enforcement is sought may not refuse enforcement solely on the ground that there are discrepancies in the legal rules of the Contracting States. To invoke the clause it is not sufficient that the court of the second forum might have come to a different decision. Because of that, a referral to the public policy provision is only possible when the enforcement of the foreign judgment would be 'at variance to an unacceptable degree'[41] with the legal rules, the enforcing court is bound by. Setting the bar high to satisfy the elements of the departure ground once more demonstrates its exceptional character. The cases in which the clause operates successfully relate to situations in which the guaranteed rights of the state of origin and of the Brussels Convention did not suffice for the protection from a grave breach of a party's fundamental rights. The main emphasis is placed on constitutional rights in this context.[42] One of those basic rights is to provide for a fair process. The latter includes the entitlement to present an effective legal defence in court.[43] The accused does not forfeit entitlement to this right just because he or she does not appear in person.[44]

When asking for a preliminary ruling on the Krombach case the German Federal Supreme Court was most uncertain about the interpretation of Article II of the Protocol of the Brussels Convention. In that provision the right of an effective defence without appearing in person is only recognised for those persons who are prosecuted for an unintentionally committed offence. Hence, the right to be defended without appearing personally shall not apply to persons prosecuted for offences, which are serious enough to deny that right.[45] Without hesitation, the European Court concluded

---

39 *Krombach v. Bamberski*, C–7/98, ECR 2000, I-1935, para. 34.

40 *Krombach v. Bamberski*, C–7/98, ECR 2000, I-1935, para. 45.

41 *Krombach v. Bamberski*, C–7/98, ECR 2000, I-1935, para. 37.

42 Behr, above n 1, 223.

43 *Krombach v. Bamberski*, C–7/98, ECR 2000, I-1935, para. 44.

44 *Poitrimol v. France*, European Court of Human Rights Series A No 277 – A.

45 Rinkau, C- 57/80, ECR 1391, para. 12.

from its previous cases that the right to a fair trial and consequently the right of a fair defence is guaranteed in all proceedings no matter whether the offence was committed intentionally or unintentionally.[46]

Ultimately, the European Court of Justice concluded that in case the French judgment was to be enforced in Germany it would be contrary to Article 103 Grundgesetz, the German Constitution. Consequently, the enforcement would be contrary to Germany's public policy. Therefore, the German courts were allowed to refuse enforcement, thus indemnifying the defendant from paying damages to the plaintiff.

Effects of the Krombach Judgment

In line with the Krombach case comes a crucial lesson with regard to the interpretation of the Brussels Convention and its successor.

The Krombach decision reveals circumstances under which recourse to the public policy clause is possible and when it is not. Thus, it fills the value phrase with a more precise and comprehensible meaning.

Firstly, it was said that the clause could not be invoked solely on the ground that jurisdiction was based on a victim's nationality. A case following the Krombach judgment confirmed this major finding.[47]

Secondly, the judgment presents the grounds on which the provision can be raised successfully. Although the attempt to define a breach of public policy still sets out general terms by demanding a 'manifest breach of a rule of law regarded as essential'[48] in the enforcing state, it is more than a hollow value phrase. The reason for that lies in the fact that with the Krombach case the European Court of Justice incorporated the content of Article 6 of the European Convention on Human Rights into the interpretation of the departure ground in question. This step implies that the requirements of procedural fairness under the Brussels Convention are now congruent with those of the European Human Rights Convention.[49]

As a result of this incorporation, it can be observed that the court's decision enhances an increasingly European approach regarding the interpretation of the Brussels

---

46 *Krombach v. Bamberski*, C–7/98, ECR 2000, I-1935, para. 42.

47 *Maxicar v. Renault II*, C-38/98, ECR 2000, I-2973.

48 *Krombach v. Bamberski*, C–7/98, ECR 2000, I-1935, para. 37.

49 Muir Watt, Horatia, 'Evidence of an Emergent European Legal Culture: Public Policy Requirements of Procedural Fairness Under the Brussels and Lugano Conventions', (2001) 36 *Texas International Law Journal*, 539, 549.

Convention. Therefore, it leaves less room for the national judges to define the notions of public policy. This is why one legal scholar suggests that the judgment reflects a significant shift in which a European standpoint is taking over, and forcing the national values to retreat and giving way to European parameters.[50] How far this European approach of interpretation is beneficial and whether the Member States will adhere to the guidelines set out by the European Court of Justice remains to be seen.[51]

A disadvantage of imposing European values is that it undermines to some extent the sovereignty of the Member States. This is especially true with regard to the public policy term since it seemed to be the last bastion without interference from superior European authorities. Instead of considering exclusively their own national interests and values, the local judges now have to refer to the European ones. Those do not necessarily correspond with the national ones. In the worst case, they might even be contrary to them. However, it is also important to consider the European approach from a different perspective. One has to keep in mind that the European Court only set boundaries to the interpretation of public policy. Within those limits the national courts are still free to attribute a precise national meaning to the value phrase. Thus, despite the superior guidelines set out by the European Court of Justice the emphasis with relation to public policy is still placed on domestic public policy. Therefore, when deciding whether the enforcement of a foreign judgment would be contrary to the enforcing court's public policy, the courts have to consider the guidelines set out by the European Court. Nevertheless, as an ultimate safeguard all European jurisdictions may refuse the enforcement of the foreign judgment if it runs counter to the national (domestic) public policy[52]. However, the probability of European values being contrary to those of the Member States is not dangerously high. The reason being is that –as was mentioned above- the European Court seeks to protect fundamental rights of the Member States and its citizens. Therefore, the emphasis is put on constitutional rights. All Member States of the European Union want to ensure those rights to its citizens. Despite minor differences in the exact scope of those fundamental rights, the basic notions are very close.

Furthermore, there is only way to guarantee a uniform application of the public policy clause. This can be achieved when superior guidelines have to be taken into account in every national consideration concerning the content of public policy. This argument relates to another one, which can be put forward in favour of a European interpretation of public policy. That is that the ruling of Luxembourg has to be discussed in the glare of the Convention's objectives. Only when the Member States

50 Muir Watt, above n 49, 543.

51 Muir Watt, above n 49, 546, 552.

52 Behr, above n 1, 221.

apply the Convention uniformly the goal to facilitate to the greatest possible extent the free movement of judgments can be enhanced.

The decision has also conjured quite a debate concerning the competence of the European Court to impose European values on the interpretation. Prior to the decision, the majority view was that the European Court of Justice did not have competence to define the notions of public policy. This competence was rather exclusively allocated to the national courts.[53] This view was also expressed by a judgment of Germany's Federal Supreme Court. It had ruled that public policy was not a subject for the European Court of Justice to deal with. Therefore, the German court refused to allow a reference to be made on that subject.[54] In addition, two of the Advocate Generals of the European Court of Justice were convinced that defining the notions of public policy was a task solely for the national judges.[55] However, Luxembourg ignored these opinions. The court stated that a national definition of public policy would advance an excessively huge concept of the refusal ground. Since this is to operate in exceptional cases only an excessive use had to be prevented. That aim could only be obtained by making the Member States comply with some European values.[56] Otherwise, the clause would lose its exceptional character.

One of the beneficial aspects of the Krombach judgment lies in the fact that it contributed to the clarification of the relationship between Article 27 paragraph 1 and paragraph 2 of the Brussels Convention. The national courts assumed over a long period that there were no other cases than those immanent to Article 27 paragraph 2 of the Convention dealing with procedural fairness. Consequently, it was thought Article 27 paragraph 1 could only envisage those cases in which enforcement was refused for other reasons than for a violation of due process. The decision in question however reverses this supposition by explicitly extending paragraph 1 to procedural fairness –with the incorporation of Article 6 of the European Convention on Human Rights- instead of letting this clause operate only in cases where a breach of substantive public policy would occur.[57]

However, even after Krombach the potential residual role of the public policy clause has not been clarified yet. In specific, it remains to determine if recourse to Article 27 paragraph 1 is prohibited if the elements of another ground for refusal were not fully

---

53 *Hoffmann v. Krieg*, C-145/86, ECR 1988, 645, para. 58; Kroppholler, Jan, 'Europäisches Zivilprozessrecht: Kommentar zum EuGVÜ und Lugano Übereinkommen', 6. Auflage, 1998, p. 343.

54 BGH 26.9.1979 BGHZ 167, 171.

55 *Krombach v. Bamberski*, C–7/98, ECR 2000, I-1935, para. 24.

56 *Maxicar v. Renault II*, C-38/98, ECR 2000, I-2973, para. 28.

57 Muir Watt, above n 47, 550.

satisfied. The European Court of Justice dealt with another case in which the relationship between the public policy reservation and other departure grounds was at stake. In that decision the court held that recourse to public policy is excluded when the issue must be resolved on the basis of another, respectively a more specific provision in the Brussels Convention.[58] This view was confirmed in the case of *Hendrikman v. Magenta*.[59] However, the aforementioned cases did not consider the potential residual role of the departure ground of public policy. It remains unanswered if the 'mutual exclusivity' criterion as laid down in *Hoffmann v. Krieg* means that whenever a legal situation could be captured by another provision but its elements are not fully satisfied a recourse to public policy is precluded?[60] If this view was confirmed it would be easier to operate the public policy clause as an exceptional provision. Thus, the fear of an abusive use of the clause could be avoided.[61] Since the European Court itself had not decided on the residual role of public policy itself yet, one of the court's Advocate Generals expressed a tendency. According to his opinion even in cases where there is a specific rule this fact does not necessarily preclude a successful invoking of the public policy clause after the failure of the more specific provision.[62] Whether the court will follow this proposal has to be watched in future judgments.

In conclusion, the Krombach judgment has achieved the aim of defining pubic policy more precisely than it was done before. Thus, the uniformity in the application of the departure ground will increase. On the other hand, some questions – especially the potential residual role- still remain unanswered.

The Public Policy Clause and its Consistency with the Principle of the Free Movement of Judgments

One of the fundamental objectives governing the Brussels I Regulation is to simplify the formalities for the enforcement of foreign judgments. Precisely, the Brussels I Regulation aims at promoting the free movement of judgments within the

58 *Hoffmann v. Krieg* ,C-145/86, ECR 1988, 645, para. 21.

59 *Hendrikman v. Magenta,* C-78/95, ECR 1996, I-4943, para. 23.

60 Newton, above n 31, 400.

61 Ibid.

62 *Krombach v. Bamberski*, C–7/98, ECR 2000, I-1935, para. 27.

community.[63] The preamble of the Regulation indicates this by implementing Article 293 of the EEC.[64]

The aim underlying the principle of enhancing a free circulation of foreign judgments is to promote the 'legal protection of persons established in the European Community'. Thus, the judgment creditor will receive a fair and great chance that a judgment rendered to him in the initial forum shall be enforceable in another state in circumstances where this becomes necessary.

In order to guarantee a full implementation of the principle of the free movement of judgments the absence of two different types of hurdles is required. Firstly, there should not be any departure grounds from the general rule to enforce another state's judgment. Secondly, the Brussels I Regulation should not contain any procedural barriers. It can be inferred from these considerations that the principle of free movement of judgments can be split up into two sub-principles. The first one is the principle of full respect for judgments delivered in a member state. The second one is the principle of a swift and simple procedure for the enforcement of a foreign judgment.[65]

According to the first sub-principle, a judgment should have the same authority and effectiveness in the country where enforcement is sought as it has in the country of origin.[66] However, although this principle is a fundamental aim of the Brussels I Regulation, its full implementation is impossible. The reason for that lies in the Regulation itself. As was demonstrated above, it contains several grounds to deviate from the pro-enforcement rule.[67] The reason for providing those barriers lies in the following: the departure grounds are generated by principles and interests, which are embodied in the Regulation as much as the principle of the free circulation of judgments is. In other words, the Regulation contains more than one fundamental principle. With the accumulation of so many objectives, under certain circumstances a collision of these different principles is likely to occur.[68] Then it becomes necessary to find the right balance between those conflicting values, to try to reconcile them, and to allow maximal development for each of them.[69] However, giving effect to all of the

---

63 *Unibank v. Christensen*, C-260/97, ECR 1999, I-3715.

64 Lane, Suriyakumari, 'Free Movement of Judgments within the EEC', (1986) 35 *International and Comparative Law Quarterly*, 629, 629.

65 Pontier, above n 4, p. 28.

66 *Hoffmann v. Krieg*, C 145/86, ECR 1988, 645, para. 10f..

67 Pontier, above n 4, p. 30.

68 Pontier, above n 4, p. 31.

69 *Solo Kleinmotoren v. Boch*, C 414/92, ECR 1994, I-2237, para. 20.

principles to the greatest extent is sometimes impossible. The reason is the mutual exclusivity of some of the principles. If the aims are mutually exclusive, then it is indispensable to give priority to one principle over another one. It is commonly accepted that there are values whose observance and protection are more important than upholding the free circulation of judgments. Hence, the Regulation's objective to promote to a greatest possible extent a free movement sometimes has to supersede to higher valued policies implied in other provisions. One of those priorities can be the adequate protection of the defendant or the adequate protection of superior national interests.[70]

At this stage the public policy exception comes into play. As was seen above, the provision aims at ensuring the protection of fundamental interests of the Contracting States' legal orders. This is considered to be so important that it was given priority and that is has found explicit expression in the Regulation.

However, although the Regulation attributes different weight to opposing principles this does not mean that one principle must always be totally neglected in the glare of another one. If it is possible, both aims should be upheld. This can be inferred from the Krombach judgment, which demands a narrow interpretation of the public policy clause.[71] On the one hand, the exception is given priority over the principle of a free circulation of foreign judgments. On the other hand, the public policy reservation is to operate in exceptional cases only.[72] However, this exceptional character of the clause does not imply that the refusal grounds should be interpreted in a way, which would make them fully redundant to the other principles. Therefore, construing them to an extent that would deprive them of their overall legal effect is unacceptable as well.[73]

The second sub-principle derived from the objective of the free movement of judgments means that a party seeking enforcement should not be hampered by procedural obstacles. Hence, the ideal situation to promote this principle would be to have an automatic recognition and enforcement process. Since this is not the case, the procedure has to be at least rapid and easy.[74] However, the Brussels I Regulation is no danger to ensure the implementation of this principle. As was set out in the beginning of this note, there are no great formal hurdles to overcome to enforce a foreign judgment.[75]

---

70 Kaye, above n 17, p. 1437.

71 *SoloKleinmotoren v. Boch* C 414/92, ECR 1994, para. 20.

72 *Hoffmann v. Krieg,* C 145/ 86, ECR 1988, 645, para. 21.

73 Pontier, above n 4, p. 35 f.

74 Lane, above n 64, 639.

75 Lane, above n 64, 640.

In conclusion, the principle of the free movement of judgments and its sub-principles are firmly established in the Brussels I Regulation. Nevertheless, their nature is not absolute because the Regulation contains competitive or conflicting principles. Also, the Regulation provides provisions to deviate from it. Still, it has a major impact on the interpretation of the departure grounds, especially regarding the public policy exception.[76] Therefore, although the public policy clause does not to a full extent support the free circulation goal, this clause must be maintained for it ensures the protection of fundamental values. Here, as so often in law, it is essential to find the right balance between contravening interests. The guidelines suggested by the European Court for the interpretation of public policy are helpful for the formation of this balance.

Enforcement of Foreign Arbitral Awards Under the New York Convention

As was pointed out in the introduction, international trade has massively expanded during the last decades. Coinciding with this expansion was an increase in the number of international commercial disputes. The business world, being aware of the complications that still exist to obtain and enforce a foreign judgment, therefore turned to other grounds of conflict resolution. The most popular method in this respect is international commercial arbitration. It is believed, that by taking recourse to this dispute resolution mechanism it will be easier to enforce a foreign arbitral award in a second forum than it would be to enforce a foreign judgment in a forum different from the delivering forum.[77]

Therefore, it will be examined here whether this 'belief' is correct regarding the court's ability to refuse the enforcement of a foreign arbitral award under the public policy reservation of the New York Convention.

Owing to the comparative nature of this note, the author will point out the similarities and commonalities between the departure ground of public policy in the Brussels I Regulation and the New York Convention first and afterwards discuss the differences.

---

[76] Pontier, above n 4, p. 44.

[77] Contini, Paolo, 'International Commercial Arbitration The United Nations Convention on the Recognition and Enforcement of Foreign Arbitral Awards', (1959) 8 *American Journal of Comparative Law*, 283, 283.

Commonalities and Similarities

Both, the Brussels I Regulation and the New York Convention contain departure grounds to refuse the enforcement of foreign judgments respectively awards. The refusal grounds are set out in Article V of the New York Convention. Under Article V (2) (b) the courts may deny enforcement of a foreign arbitral award if such would be contrary to the public policy of the enforcing country. The list of departure grounds under the New York Convention is exhaustive as well, meaning that the courts can only deny enforcement if the elements of one of the provisions are satisfied.[78] Also, either the defendant or the court ex officio may invoke the public policy reservation.[79] Another commonality is that under the New York Convention the departure ground is split up into the two categories of procedural and substantive public policy.[80] Furthermore, under both instruments the public policy clause has a dynamic character, which is subjected to sociological and economic changes. Just because it is embodied in a different Convention cannot change the term's nature of a value phrase.[81]

As much as the Brussels I Regulation forbids any review to the substance under the public policy clause, the New York Convention must not serve as an open door for a judicial review of the arbitral award either.[82]

A little bit more controversial is whether the public policy exception is to be interpreted narrowly as under the Brussels I Regulation? This question has been subject of a debate among different legal authorities. The problem results from the fact that the legislative history of the New Convention neither provides clear guidelines in favour, nor against a narrow construction of the public policy clause.[83] The precursors of the New York Convention contained the departure ground of public policy, too. However, the wording was not exactly the same. Rather than mentioning only public policy – as the New York Convention does - its precursors extended the clause. They

---

78 Reisman, W. Michael, Craig, W. Laurence, Park, William, Paulsson, Jan, 'International Commercial Arbitration Cases, Materials and Notes on the Resolution of International Business Disputes, 1997, p. 1261.

79 Craig, W. Laurence, Park. William W., Paulsson, Jan, 'International Chamber of Commerce Arbitration', 3rd edition, 2000. p. 684.

80 Lew, Julian D M, Mistelis, Loukas A, Kröll, Stefan M, 'Comparative International Commercial Arbitration', 2003, p. 723.

81 Ibid.

82 Lowenfeld, above n 34, p. 352.

83 Berglin, Hakan, 'The Application in the United States Courts of the Public Policy Provision of the Convention on the Recognition and Enforcement of Foreign Arbitral Awards', (1985-1986) 4 *Dickinson Journal of International Law*, 167, 168.

added the terms contrary to 'principles of law' or violative of 'fundamental principles of the law'.[84]

Therefore, one opinion suggests a narrow interpretation of the provision. This hypothesis allegedly finds support in the pro-enforcement bias underlying the Convention.[85] Also, an expansive interpretation of the departure ground would oppose the Convention's basic effort to abolish the pre-existing barriers to the enforcement of foreign arbitral awards.[86]

Another authority, however, concludes exactly the opposite from the omission of the words cited above. According to that opinion, leaving those terms out is an indicator favouring a broad interpretation of the clause.[87]

To resolve this controversy, a look at the history of the Convention as a whole is instructive. Especially the general pro-enforcement bias points towards a narrow interpretation of public policy. If the courts interpreted the public policy provision broadly, the realisation of the Conventions' aims would be much more difficult. The judges could, according to their subjective preferences, classify a minor 'offence' as contrary to public policy. This would impede the enforcement of foreign arbitral awards instead of facilitating it. In addition to that, considerations of reciprocity advise the courts to handle the public policy reservation with great care.[88] Reciprocity in this context means that the decision of foreign courts should only be enforced if the foreign country was likely to enforce the judgments rendered by the second forum as well.[89] Achieving this reciprocity is more likely if the foreign countries will mutually rely on the other state's award. That is only possible with a narrow interpretation of the clause.

Hence, it is advisable to interpret the clause narrowly and reserve it for the gravest infringements only.

---

84 Lowenfeld, above n 82.

85 Contini, above n 77, 304.

86 Cole, Richard A., 'The Public Policy Exception to the New York Convention on the Recognition and Enforcement of Arbitral Awards', (1985-1986) 1 *Ohio State Journal on Dispute Resolution*, 365, 366.

87 Quigley, Leonard V., 'Accession by the United States to the United Nations Convention on the Recognition and Enforcement of Foreign Arbitral Awards', (1960-1961) 70 *Yale Law Journal*, 1049, 1077f.

88 Lowenfeld, above n 34, p. 353.

89 Behr, above n 1, 221.

Differences

Despite the aforementioned commonalities also a few differences are observable with regard to the public policy clause under the Brussels I Regulation and the New York Convention.

Acceptance of the New York Convention

As a starting point it is worth mentioning that the New York Convention is one of the most widely accepted multinational conventions. Its basic thrust is to liberalise and facilitate to the greatest possible extent the procedures to enforce foreign arbitral awards. A large number of countries are parties to the Convention.[90] This fact makes the instrument highly satisfactory to award creditors seeking enforcement.[91] Although the Brussels I Regulation pursues the same objective, only the Member States of the European Union are bound by the Regulation. Therefore, the instrument is of great use when it comes to the enforcement of judgments within the Member States. If however, a judgment from a non-member state needs to be enforced, the Regulation will not operate, thus making the enforcement process more complicated. Therefore, one advantage of the New York Convention lies in its global application and acceptance.

Discretion

The most obvious distinctive feature can be inferred from the choice of words of the public policy provision. Whereas the Regulation does not give any discretion to the courts, the New York Convention allows the courts to exercise discretion. This is implied by using the words 'may refuse'. Thus, if a court chooses to enforce a foreign award although one of the departure grounds is established, this does not result in a violation of the treaty.[92] Providing a court with discretion enables it to be much more flexible in the process of decision-making. Instead of applying inflexible rules of law to a situation where the application of these rules might lead to an unfair result, the court can consider circumstances, which are not described in the provision in question. The consideration of those circumstances might be very important for the finding of a judgment that is individually just and fair. The discretion under the New

---

90 Harnik, Hans, 'Recognition and Enforcement of Foreign Arbitral Awards', (1983) 31 *American Journal of Comparative Law*, 703, 703.

91 Lew, above n 80, p. 693 f.

92 Lowenfeld, above n 34, p. 362.

York Convention is therefore a great benefit in the enforcement process of foreign arbitral awards as opposed to the enforcement of foreign judgments under the Brussels I Regulation.

Domestic Versus International Public Policy

In the *Overseas v. Rakta*[93] case, the court had to face the question whether public policy under the New York Convention is equitable to or if it is reflected in domestic governmental policies. In other words, the court had to answer if public policy under the New York Convention has the same meaning as public policy in a national sense.[94]

The court concluded in its judgment that there is a difference between domestic and international public policy. The latter demands a narrow interpretation of the provision in question.[95] In contrast, when it comes to domestic law, the public policy elements may be construed broadly. The reason for a narrow construction is that a wide interpretation favouring a protection of national interests would run counter to the Convention's objectives.[96] One of the most important objectives is to facilitate the enforcement process of foreign arbitral awards. To stress that, the courts also pointed out that the defence should not be used to safeguard national interests but rather to focus on supranational values. Hence, considerations with an international character were thought to outweigh the benefits, which were intended by some national statutory provisions. A court decided so in the case of *Scherk v. Alberto Culver Co.*[97] Eventually, enforcement should only be denied if such would 'violate the forum state's most basic notions of morality and justice'[98]. This international standard was repetitively emphasised in the *Fotochrome Inc. v. Copol Co. Ltd*[99]case. However, although it was ruled in the *Overseas* case that a violation of domestic public policy will usually not establish the defence, under certain circumstances this general rule might be invalid. That could be the case when the violated national law is based upon

---

93 *Overseas v. Rakta*, 508 F.2d 969, (2nd Circuit 1974)

94 Junker, Joel R.,'The Public Policy Exception to Recognition and Enforcement of Foreign Arbitral Awards', (1977) 7 *California Western International Law Journal*, 228, 239.

95 *Overseas v. Rakta*, 508 F.2d 969, 974, (2nd Circuit 1974).

96 *Overseas v. Rakta*, 508 F.2d, 969, 974, (2nd Circuit 1974).

97 *Scherk v. Alberto Culver,* 417 U.S. 506 (1974).

98 *Overseas v. Rakta*, 508 F.2d, 969, (2nd Circuit 1974).

99 *Fotochrome Inc. v. Copol Co. Ltd*, 517 F.2d, 512 (2nd Circuit 1975).

the aforementioned basic notions of morality and justice. Then, the public policy exception, as set out in the New York Convention, is likely to apply.[100]

Defining that enforcement of a foreign arbitral award is contrary to public policy if the basic notions of morality and justice are violated, was a short-term relief of dealing with the public policy clause. The reason for that is that this concept does not offer any specific guidance. In contrast, it is very vague and sets out a subjective standard which calls for interpretation by the national judges.[101] In particular, it remains unsure in which circumstances the public policy clause can be used effectively as a defence.[102]

Concluding from the previous passages, in most of the cases that had to be decided a conflict between the pro-enforcement bias of the Convention and the national public policies arose. However, the evolving case law on the application of the public policy clause has hardly presented a challenge on how far the courts are willing to go in terms of enforcing foreign arbitral awards. Thus, although quite a few cases dealt with the public policy clause, there is still a lack of certain guidelines as to when the clause will operate.[103]

The narrow interpretation of the public policy defence conjures up another problem, too. Originally, the provision's intention was to serve as a 'catchall' for those cases in which the fundamental notions of morality and justice have not yet been captured by a more specific departure ground.[104] Regarding this intention, one authority alleges that the United States' courts failed to comply with it. Rather than letting the public policy clause operate as a catch all provision, it is claimed that the narrow interpretation of the clause leads to its insignificance or even total meaninglessness. Thus, according to this opinion, there would hardly be a difference if the defence was non-existent under the New York Convention.[105] According to this opinion, the defence is an expression of ultimate national power. Apparently, however, the courts so far feared exercising this power, thus enforcing foreign arbitral awards in the vast number of cases. Also, the conceptual weaknesses of the provision will make it almost unpredictable for arbitrators, lawyers, the courts, and for the business world to anticipate the limits of fairness in international commercial arbitration.[106] It is therefore concluded that the narrow interpretation might result in a decreasing use of

100 Junker, above n 94, 241.

101 Cole, above n 86, 377.

102 Cole, above n 86, 365, 375.

103 Berglin, above n 83, 169.

104 Junker, above n 94, 245.

105 Ibid.

106 Ibid.

international commercial arbitration. Thus, attorneys might advise their clients to refer to the 'safety' of the national judiciary. This, however, is contrary to the intention to take the burden from the courts and to promote arbitration as a method of alternative dispute settlement.[107]

Defining Public Policy under the New York Convention

As was seen above, a precise definition of public policy under the New York Convention is still missing. That makes it impossible to predict if the clause will operate in an individual case. That in return is problematic in so far as the public policy clause is the most general departure ground and a party is likely to try invoking it when any specific provisions have failed.

Several legal writers have therefore used this unpleasant situation as an incentive to attempt defining the clause's meaning more precisely and avoid any uncertainty as to when it will operate.

One author intends to achieve certainty by endowing the clause with an objective standard. Thus, the national judges shall be prevented from applying their different subjective standards to the provision and in return guarantee a uniform application of the clause. The suggested standard sets out positive terms to describe when enforcement would be contrary to public policy. With the infringement of those terms, the courts can refuse the enforcement of the foreign arbitral award.[108] Precisely, the public policy reservation shall operate if the enforcement of the award would violate federal law.[109] Consequently, the striking feature of this definition is the violation of federal law. The scope of federal law should cover all congressional enactments, as well as rules and regulations emerging from those enactments. Case law, however, should not form a part of federal law.[110]

By emphasising a violation of federal law, the logical consequence must be that a breach of a state's law does not suffice to trigger the departure ground. However, this does not imply a detriment to the state law. The reason being is that the circumstances under which the state courts used to deny enforcement are expressed in the other

---

107 Junker, above n 94, 246 f.

108 Barry, Robert A.J., 'Application of the Public Policy Exception to the Enforcement of Foreign Arbitral Awards under the New York Convention: A Modest Proposal', (1978) 51 *Temple Law Review Quarterly*, 832, 843, 844.

109 Barry, above n 108, 845.

110 Barry, above n 108, 846 f.

grounds for refusal of the Convention.[111] Violation in the glare of this definition means that the action, which forms the basis of the award must establish a violation of public policy.[112]

Another opinion suggesting a more precise definition of public policy derives its idea from the definition under German law. The latter defines public policy rather broadly. Under German law an award would be contrary to public policy if the enforcement would offend the good morals or the main objectives, and also, if it would require a party to engage in an illegal action. The public policy clause in Germany is applied within the framework of certain circumstances described in the law. E.g., the German Code of Civil Procedure, Zivilprozessordnung, names certain situations under section 328 in which the enforcement of a foreign arbitral award would be contrary to public policy. This is a contrast to referring to the vague concept of morality and justice exclusively.

For those who might find a statutory definition undesirable, the author suggests another option to define public policy more specifically. He proposes that parliament should indicate which laws contain fundamental principles whose violation would be contrary to public policy.[113]

Yet another opinion proposes a clearer definition of public policy. The authority suggests that the clause should operate if the enforcement would violate the most basic notions of morality and justice of the second forum. Additionally, the courts should be allowed to deny enforcement if it would be detrimental to the interests of third parties or of the public.[114] The latter part of the definition relates to the idea that in some legal areas, rules are created to protect interests, which are beyond the fair resolution of a dispute between two parties. In these situations, public policy should be used to override an arbitration agreement because the arbitration will usually only consider the interests of the involved parties.[115]

All of the proposals refer to the national law. Therefore, these suggestions seem to counter run the standards of the courts. They had demanded to give public policy under the New York Convention a supranational meaning rather than focusing on a domestic standard. As soon as the definitions refer to a national standard, a global and uniform application of the Convention cannot be guaranteed. This is because every

111 Barry, above n 108, 848.

112 Ibid.

113 Junker, above n 94, 246ff.

114 Cole, above n 86, 383.

115 Cole, above n 86, 382.

country has its own federal law, thus its own meaning of public policy. Yet, in accordance with the proposals, certainty as to when the clause can be invoked should have priority over a uniform application. Therefore, with relation to the pro arbitration bias it is comprehensible to let the international standard supersede the more important factor of certainty.

Despite the potential disadvantage of abandoning an international standard, the first and the second proposal reveal significant benefits. Firstly, they provide specific guidelines as to when the public policy provision can be raised successfully. The reason for that is that the first definition only takes recourse to objective criteria. Thus, the circumstances under which the clause will operate do no longer depend on the subjective interpretation of the national judges.[116] This in return will increase the certainty when the clause can be invoked successfully.

The benefit of having a definition as under German law is a high predictability as to when the clause will work. This definition also confines the circumstances strictly enough in order to prevent the provision from becoming a loophole. Finally, it is important to mention that despite a broader interpretation of public policy under German law, arbitration flourishes in Germany. Thus, the broader interpretation is not detrimental to the use of arbitration.[117]

With relation to the third proposal, the following is relevant: it differs from the other proposals because it still adheres to the concept of basic notions of morality and justice. By referring to this standard, the factor of uncertainty remains. By adding the second part to the definition, precision is increased. Nevertheless, the judges still have to determine which public or third party interests are intended to be protected. This might be interpreted differently from judge to judge again. Hence, this definition does neither promote a uniform application, nor does it increase certainty. Consequently, the adaptation of this standard is unlikely to bring along a great change to the current problems with the departure ground in question.

Conclusion

The remarks above have revealed similarities as well as differences regarding the application of the public policy clause under the Brussels I Regulation and under the New York Convention.

---

116 Barry, above n 108, 850.

117 Junker, above n 94, 249.

Concerning the choice of using international litigation or international arbitration the following is relevant: the fact that under the New York Convention the courts are endowed with discretion makes arbitration a more attractive dispute resolution mechanism. This discretion is another feature of flexibility in arbitration, which is lacking to a certain extent in international litigation. Therefore, with regard to this factor, the assumption that enforcement under the New York Convention is easier than under the Brussels I Regulation is correct. The assumption is also confirmed by the world-wide acceptance and application of the New York Convention as opposed to the Regulation.

The other conspicuous feature is the difference concerning the focus on domestic public policy under the Brussels I Regulation as opposed to international public policy under the New York Convention. With regard to this distinction, it is instructive to return to the competence of the European Court in terms of interpreting the Brussels I Regulation.

It was seen that the national courts may address the European Court of Justice when uncertainties with the application of the Regulation arise. Because of this power, the Court was able to impose a certain European standard as to the construction of the public policy clause. It sets out boundaries and the national courts are forced to adhere to these limitations. Within those limits however, the national courts are free to rule according to their domestic concept of public policy. Therefore, despite that there are certain superior guidelines within the European Union, the national courts still have a lot of authority and power. The guidelines help facilitating a uniform application of the clause under the different jurisdictions of the Member States.

Concerning the New York Convention, there is no superior organ the national courts can address in cases of uncertainty in the interpretation of the public policy clause. Rather than that, it only depends on the Contracting States and their view of what public policy should mean under the Convention. When an international standard is suggested without any precise guidance as to what this standard must look like, then it is very likely that all courts will interpret the clause differently and ultimately return to the national concepts anyway. Therefore, in order to achieve a uniform application it might be justified for the Contracting States of the New York Convention to seek a definition, which is more precise and to allow the states to refer to a national standard. However, the bar must be kept at quite a high level in order to prevent undermining the Conventions' objectives. That way, it can at least be said with certainty as to when the clause will operate. In addition, one should rely on the national courts' respect to handle the clause in exceptional circumstances only, thus promoting the objectives of the New York Convention.

STATUTORY UNCONSCIONABILITY AND GUARANTEES

Janine Pascoe[*]

Abstract

Those who follow developments in the law relating to guarantees have been confronted in recent times with an ever increasing array of statutory provisions designed to deal with unconscionable conduct. It seems that allegations of unconscionable conduct, in its various forms, on the part of the creditor have become part of the standard defence of those guarantors seeking to set aside the guarantee. It may not be too far fetched to suggest that the enactment of a raft of unconscionability provisions across a range of statutes at both state and federal levels has left many perplexed. Many of the provisions are similar, but not identical. This article examines the scope and application of overlapping provisions in the ASIC Act and TPA which govern unconscionability with respect to financial services and considers their role in relation to contracts of guarantee. It further considers what role the regulator, ASIC, has to play in setting normative standards in a systemic way for banks and financiers when taking guarantees.

Introduction

Cases involving guarantees have proved a fertile ground for the courts to consider the parameters of unconscionability, both in equity and under statute. Vulnerable guarantors continue to enter into sureties which may be both improvident and unfair. Recent cases have involved wives,[1] as well as elderly parents,[2] siblings,[3] de facto spouses[4] and others in a close relationship to the guarantor.[5]

---

* BA, LLB (Hons) PhD (Monash) Senior Lecturer, Dept of Business Law & Taxation, Monash University.

1 *Burrawong Investments Pty Ltd v Lindsay* [2002] QSC 82.

2 *Watt v State Bank of New South Wales*.

3 *Equuscorp Pty Ltd v Worts* [2000] VSC 179.

4 *Liu v Adamson* [2003] NSWSC 74; *State Bank of NSW v Hibbert* [2000] NSWSC 628.

In earlier times, before the intervention of statute, plaintiffs could seek relief at common law under the *Amadio*[6] principles, or, in the special case of wives, under the rule in *Yerkey v Jones*.[7] The common law rules apply to guarantees given to support both consumer and business borrowings. As this article will explain, some of the recent provisions codifying unconscionability appear to be replacing the traditional *Amadio* principles with broader notions of unfair exploitation.

Statutory remedies to set aside unconscionable contracts, including guarantees, are of relatively recent origin. While there are advantages[8] in pursuing an action for statutory unconscionability, the enactment of a raft of unconscionability provisions across a range of statutes at both state and federal levels has left many perplexed. This article explains the coverage and interaction of the provisions in the *Trade Practices Act 1974* (Cth) ('TPA') and *Australian Securities and Investments Commission Act 2001* (Cth) ('ASIC Act') targeting unconscionable conduct in relation to 'financial services' and considers how those provisions apply in the guarantee context. The Uniform Consumer Credit Code ('UCCC')[9] and State and Territory legislation also provide relief from unjust or unconscionable consumer contracts.[10] However, the focus of this paper is on statutory unconscionability at the Commonwealth level, with particular attention given to provisions affecting guarantors of small business debts. Those provisions cover transactions which occur across State/ Territory boundaries. There is scope for the better resourced Commonwealth regulators, the Australian Securities and Investments Commission ('ASIC') and the Australian Competition and Consumer Commission ('ACCC'), to deal with contraventions of national significance and to take action where conduct involves major detriment to consumers and small business.

---

5 *Kranz v National Australia Bank* [2003] VCSCA 92 (brothers-in-law); *ANZ Banking Group Ltd v Alirezai* [2004] QCA 6 (close friends).

6 *Commercial Bank of Australia Ltd v Amadio* (1983) 151 CLR 447 ('*Amadio*').

7 1939) 63 CLR 649, affirmed in *Garcia v National Australia Bank Ltd* (1998) 194 CLR 395 ('*Garcia*').

8 In particular, since 2001 plaintiffs may pursue a remedy in damages for breach of the TPA/ASIC Act unconscionability provisions.

9 The UCCC regulates guarantees given to support debts of a wholly or predominantly, household, personal or domestic nature: ss 6 (5) and 50.

10 See, for example, s 70(1)-72 UCCC and s 7 (1) *Contracts Review Act 1980* (NSW) *('*CRA*')* which provide relief from contracts which are termed 'unjust' rather than 'unconscionable'. The State and Territory Fair Trading Acts regulate 'unconscionable' consumer transactions. The Victorian and Tasmanian FTAs also catch unconscionability in small business dealings: *Fair Trading Act 1999* (Vic) s 8A; and *Fair Trading Act 1990* (Tas) s 15A.

As a result of financial services reforms[11] providing for some conduct with respect to financial services which would previously have come under the TPA, to be governed by the ASIC Act, the two acts contain three sets of provisions which are substantially similar, but not identical. Those provision deal with unconscionability within the meaning of the unwritten law, consumer-type unconscionability and small business unconscionability. The respective application of those three types of unconscionable conduct provisions to contracts of guarantee is determined by considering particular exclusions and definitions under the TPA and ASIC Act.

This article highlights some limitations of the various provisions dealing with unconscionable conduct in financial services. Provisions dealing with the 'unwritten law' have not been given an expansive interpretation, but have been confined within the traditional boundaries recognised in equity. The relatively recent introduction of prescriptive provisions dealing with consumer and small business unconscionability has generally been regarded as denoting an extension in the law[12] but the degree to which this is the case is still uncertain. The consumer unconscionability provisions have a limited role to play in the guarantee context, given that most problems have concerned guarantees given to support small business debts.[13] There have not been many appellate decisions concerning the small business unconscionability provisions, so although clearly their terms are of a broader ambit than the equitable doctrine, their ultimate scope is yet to be determined. Those actions which have been successful have involved only 'the most egregious behaviour'.[14] The scarcity of cases under these

---

11 The *Financial Services Reform Act 2001* (Cth) introduced ss 12CA-12CC dealing with unconscionable conduct into the ASIC Act.

12 See, for example, Bryan Horrigan, 'Unconscionability breaks new Ground-Avoiding and Litigating Unfair Client Conduct After the ACCC Test Cases and Financial Services Reform' [2002] *Deakin Law Review* 4; The Hon Paul de Jersey AC, Chief Justice Queensland Court of Appeal, 'Good Faith, Commercial Morality and the Courts', 22nd Annual Banking and Financial Services Law Conference, Cairns, 6 September 2005. (http://www.courts.qld.gov.au/publications/articles/speeches/2005).

13 See, for example, *Yerkey v Jones* (1939) 63 CLR 649 and in more recent times *Garcia v National Australia Bank Ltd* (1998) 194 CLR 395; *Bylander v Multilink Investments Pty Ltd* [2001] NSWSCA 53 (Unreported, New South Wales Court of Appeal, Handley, Giles and Heydon JJA, 14 March 2001); *State Bank of New South Wales v Chia* [2000] 50 NSWLR 587; *Burrawong Investments Pty Ltd v Lindsay* [2002] QSC 82.

14 Fair Trading Coalition, 'Trade Practices Act Reform-Policy Paper', 4 August 2003 (available at www.mtaa.com.au/news.ftcpolicypaper.pdf). See also, Robert Gardini, 'The Dawson Review-A Small Business Perspective' (2003) 9(1) *UNSWLJ Forum*.

provisions may be due to a transactional cap of $3 million,[15] which is presently under review.[16]

Frequently, complaints by guarantors against financiers go to the inherent unfairness and one-sidedness of the guarantee. In recent times commentators[17] and consumer and reform bodies[18] have pointed to the need to consider regulatory approaches designed to deal with contracts which are unfair or harsh in terms of their outcomes, arguing that common law and statutory prohibitions against unconscionable conduct are inadequate in this regard.

Background to the TPA Unconscionability Provisions

The starting point of the analysis in this article is to examine the TPA provisions from which mirror provisions enacted in legislation elsewhere have originated. Part IVA of the TPA, initially comprising ss 51AA and 51AB, was introduced into the Act by the *Trade Practices Legislation Amendment Act 1992*(Cth). Section 51AB is the previous s 52A[19] renumbered. It was the first statutory provision prohibiting unconscionable conduct to be inserted into the TPA. It is confined to conduct relating to the provision of goods or services 'ordinarily acquired for personal, domestic or household use or consumption'.[20] Although the Act does not define 'unconscionable conduct', s 51AB

---

15 Section 51AC (10).

16 The Federal Government supports raising the threshold to $10 million. See Joint Media Statement, Hon P Costello and Hon F Bailey, 'Government Progressing Trade Practices Act Reforms', Media Release No. 013/2005 (www.treasurer.gov.au/tsr/content/pressreleases/2005). The increase was included in the *Trade Practices Legislation Amendment (Small Business Protection) Bill 2005*. The Bill has not yet been enacted. It has been renamed the Trade Practices (Legislation Amendment) Bill (No 2) 2006.

17 Nicola Howell, 'An Update of Unfair Contracts legislation-Examining the Need for Nationally Consistent regulation of Unfair Terms in Consumer contracts', Paper presented at 15th Annual Credit Conference, 29 September 2005 (www.gu.edu.au/centre/cccl/content_publications.html).

18 Standing Committee of Officials of Consumer Affairs ('SCOCA') National Working Party Discussion Paper, *Unfair Contract Terms*, January 2004 [2.1] (available at http://www.consumer.gov.au/html/publications.htm). A list of submissions to the SCOCA Working Party on Unfair Contract terms may be found at www.fairtrading.qld.gov.au).

19 Section 52A was inserted in 1986.

20 Section 51AB (5). For a discussion of the meaning of services 'ordinarily acquired for personal, domestic or household use or consumption' see Paul Latimer, 'The definition of

(2) enumerates a non-exhaustive list of guiding facts, familiar to many, which guide the court in determining whether the conduct is 'in all the circumstances' unconscionable. Incorporation of these factors means that in its terms s 51AB is wider than unconscionability under *Amadio*.[21] The equitable doctrine is confined to procedural unfairness; that is, unfairness in the bargaining process.[22] In terms of the *Amadio* doctrine, this is established by showing that in the process of taking the guarantee the creditor had knowledge of the existence of a special disability suffered by the guarantor and took advantage of that disability.

Some of the factors listed in s 51AB (2) suggest that the statute may permit relief from contracts which are unfair in their terms or operation, despite the absence of unfairness in the bargaining process. That is, not all of the factors are directed towards conduct surrounding the taking of the guarantee. For example, the court may have regard to such matters as:

- Whether the consumer was required to comply with conditions which were not reasonably necessary for the legitimate interests of the supplier; and
- The amount for which, and the circumstances under which, the consumer could have acquired goods or services from another party.[23]

Transactions involving unconscionable conduct in the business context are caught by ss 51AA and 51AC. Section 51AA provides that 'a corporation must not, in trade and commerce, engage in conduct that is unconscionable within the meaning of the unwritten law, from time to time, of the States and Territories'. The insertion of s 51AA was recommended by the then Trade Practices Commission to enable commercial transactions to be governed by principles of unconscionability, while allowing access to remedies under the TPA. [24]

'personal domestic or household use or consumption' in the TPA' (1998) 6 *Current Commercial Law* 45.

21 *ACCC v Oceana Commercial Pty Ltd* [2003] FCA 1516 [336]; *ACCC v Simply No-Knead (Franchising) Pty Ltd* (2000) 104 FCR 255, 265.

22 For a discussion of the failure of both the common law and statutory provisions in addressing substantive unfairness issues, see SCOCA National Working Party Discussion Paper, above n 18.

23 Sections 51AB (2) (b) and (e).

24 Explanatory Memorandum to the *Trade Practices Legislation Amendment Act 1992* (Cth), clause 41.

Section 51AC was introduced in 1998[25] and specifically prohibits one business (the 'business supplier') from dealing unconscionably with another business (the 'business consumer'). The provision was intended to assist small business and excludes transactions greater than $3 million or those in which the target business is a listed pubic company. If a proposed increase of the threshold to $10 million goes ahead, [26] the provision will cover most guarantees of small business debts. The increase may well relegate s 51AA to only a residual operation. Section 51AC provides a list of factors similar to, but more extensive than those outlined in s 51AB. On the face of it, several of those factors also appear to catch instances of substantive injustice.

Although there is no definition of unconscionable conduct in ss 51AB and 51AC of the TPA, its meaning appears to be coloured by the inclusion of the words 'in all the circumstances' in those provisions. For example, in *Hurley v McDonald's Australia Ltd* the Full Court of the Federal Court held:

> Before sections 51AA, 51AB or 51AC will be applicable, there must be some circumstances other than the mere terms of the contract itself that would render reliance on the terms of the contract 'unfair' or 'unreasonable'….[27]

The complaints of many aggrieved guarantors go to inherent unfairness in the terms of the guarantee itself. Standard form guarantees typically used by financial institutions, even those which comply with disclosure regimes under legislation and industry codes, may contain provisions which are unjust in themselves. A good example is the inclusion of clauses imposing a continuing and unlimited liability upon vulnerable guarantors. Known as 'all moneys' or 'all accounts' guarantees, [28] these may frequently be unexpected or misunderstood by the guarantor.[29] This kind of unfairness can be more appropriately remedied by laws which specially target unfair

---

25 *Trade Practices Amendment (Fair Trading) Act 1998*. See generally, Janine Pascoe, 'The Effect of the Federal Government's Small Business Package on Guarantees of Business Debts' (1998) 12 *Commercial Law Quarterly* 17.

26 See above n 16.

27 (2001) 22 ATPR ¶41-741, 40, 586.

28 Most guarantees currently used in business lending have a provision to the effect that they cover 'all moneys which are now or may hereafter from time to time be owing or payable on any account to the creditor by the security provider'. Clause 28.13 of the Code of Banking Practice (revised August 2003, modified May 2004) permits all moneys' guarantees, albeit with some limitations. See, Janine Pascoe, 'Guarantees and All Moneys Clauses' (2004) 4 (2) *QUT Law and Justice Journal* 245.

29 See, for example, *Commonwealth Bank of Australia v Horkings* [2000] VSCA 244; *Garcia v National Australia Bank Ltd* (1998) 194 CLR 395; *State Bank of New South Wales v Chia* (2000) 50 NSWLR 587; *ANZ Banking Group Ltd v Karam* [2005] NSWCA 344.

contract terms, and reforms in this area are still pending in Australia.[30] Given that the principles of unconscionability continue to play a role in litigation involving allegations of improvident and unfair guarantees it is important to consider whether the boundaries of statutory unconscionability are expanding, as argued in the academic commentary,[31] or whether the provisions are of little practical utility to guarantors as argued by consumer bodies.[32]

Unconscionable Guarantees-ASIC Act or TPA?

Determining whether conduct in relation to a guarantee falls within the unconscionability provisions of the ASIC Act or Part 1VA TPA involves a circuitous consideration of various definitions and exclusions set out in those Acts. Section 4(1) of the TPA includes within the definition of 'services' a contract between a banker and a customer and 'any contract for or in relation to the lending of moneys'. But this definition can only include 'services' that are not specifically excluded 'financial services'. This is because the TPA excludes from it operation conduct that would otherwise fall within Part 1VA ss 51AA or 51AB, in so far as it is conduct relating to 'financial services'.[33] Section 51AC, on the other hand, is not excluded from operation in relation to 'financial services', with the result that the ASIC Act consumer and general unconscionability provisions operate exclusively with respect to financial services, but there is an overlap between the TPA and ASIC Act small business unconscionability provisions. This issue of overlap, however, can be dealt with under the rule of statutory construction that a later specific Act, in this case the ASIC Act, impliedly overrides an earlier Act, such as the TPA which has a general operation.

---

30 See the SCOCA Discussion Paper, above, n 18.

31 There is ample literature explaining this trend in recent times. See, for example, Gail Pearson, 'The Ambit of Unconscionable Conduct in relation to Financial Services' (2005) 23 *Companies and Securities Law Journal* 105, 107-8; Bryan Horrigan, 'The Expansion of Fairness-based Business regulation-Unconscionability, Good Faith and the Law's Conscience' (2004) *Australian Business Law Review* 159.

32 Consumer Credit Legal Service (Vic), Submission to SCOCA Working Party on Unfair Contract Terms, 11 March 2004.; Consumers' Federation of Australia, Submission to SCOCA Working Party on Unfair Contract Terms, March 2004; Nicola Howell, above n 17.

33 Section 51AAB. Similarly, the consumer protection provisions of Part V TPA do not apply to 'financial services': s 51AF.

Definition of 'Financial Services'

The starting point in determining which statute applies, therefore, is to consider whether conduct involving the taking of a guarantee or giving advice about obligations under a guarantee can be considered to be a 'financial service' within the meaning ascribed by both the TPA and ASIC Acts.

'Financial service' is defined in s 4(1) of the TPA effectively to have the same meaning as in the ASIC Act. According to s 12 BAB of the ASIC Act a 'financial service' encompasses a 'financial product' such that a person provides a financial service when they deal in a financial product or provide financial product advice.[34] In addition, 'financial product' is defined in s 12BAA ASIC Act. That provision explains that a 'financial product' encompasses a 'credit facility' within the meaning of the regulations.[35] Under Regulation 2B (1) (h) of the ASIC Act a 'credit facility' includes a 'guarantee of an obligation under a credit contract'.

The result of these interlinking provisions, insofar as the interplay between the TPA and ASIC Acts is concerned, is that:

- A person who deals in or who gives advice about a guarantee of obligations under a credit contract provides a financial service.[36]
- Contracts of guarantee are therefore considered 'financial services' which are generally regulated by the ASIC Act in respect of conduct that is unconscionable.
- The ASIC Act provisions regulate the three types of unconscionability i.e. within the meaning of the unwritten law, consumer unconscionability and small business unconscionability.[37]
- Sections 51AA and 51AB TPA governing the unwritten law and consumer unconscionability respectively, do not apply to 'financial services' and therefore do not apply to guarantees.

---

34 Section 12BAB (a) and (b) ASIC Act.

35 An amendment expanding the definition of 'financial services' to include credit facilities came into effect on 11 March 2002. See s 12BAA (7) (k).

36 Sections 12BAB (1) (a) and 12BAB (1) (b). Section 12BAB (7)-(10) outlines the meaning of 'dealing' in a financial product. Section 12BAB (5) outlines the meaning of providing financial product advice.

37 Sections 12CA, 12CB and 12CC ASIC Act.

- Section 51AC TPA is not excluded from operation in relation to 'financial services', but as a matter of statutory construction the ASIC Act takes precedence.

Given the relatively recent introduction of ASIC Act unconscionability provisions covering 'financial services', the case law on the earlier mirror provisions in the TPA can be regarded as a useful guide to the interpretation on the ASIC Act sections, despite some minor differences in drafting. Dicta in a number of first instance decisions indicates that the meaning of provisions governing unconscionability and misleading and deceptive conduct in the two Acts should not differ in any significant respect.[38]

Guarantees and Unconscionability within the Meaning of the Unwritten Law

Section 12CA (1) of the ASIC Act applies the wording of s 51AA TPA to 'financial services'. It provides:

> A corporation must not in trade or commerce engage in conduct in relation to financial services if the conduct is unconscionable within the meaning of the unwritten law, from time to time, of the States and Territories.

Until there is a body of case law dealing with s 12CA, jurisprudence on the mirror provision in s 51AA will necessarily guide its meaning. According to the Explanatory Memorandum[39] accompanying the insertion of s 51AA into the TPA, this provision embodies the specific equitable concept of unconscionable conduct as recognized by the High Court in *Commercial Bank of Australia Ltd v Amadio*.[40] It undoubtedly also encompasses the 'special wives' equity' established in *Yerkey v Jones*[41] and reaffirmed

---

38 See, for example, *GPG (Australia Trading) Pty Ltd v GIO Holdings* (2001) 117 FCR 23, 68 (Gyles J); *Vitarni Pty Ltd v Macquarie Bank* [2002] NSWSC 978 [68]; *Commonwealth Bank v Spira* [2002] NSWSC 905 [132] and *Reiffel v ACN 075 839 226 Ltd* [2003] FCA 194, [60], [65]. For a fuller analysis of the various cases suggesting the comparability of interpretation of overlapping TPA and ASIC Act provisions see Gail Pearson, above n 31. Pearson notes, at 107-8, that statements indicating the comparability of the two sets of provisions were made in the context of the ASIC Act prior to the introduction of the *Financial Services Reform Act 2001* (Cth).

39 Explanatory Memorandum accompanying the Trade Practices Legislation Amendment Bill 1992 (Cth) [41]. The EM also refers to the High Court's decisions in *Blomley v Ryan*(1956) 99 CLR 362.

40 (1983)151 CLR 447.

41 (1939) 63 CLR 649.

by the High Court in *Garcia v National Australia Bank Ltd*,[42] given that the majority in that case, concluded that the old rule protecting guarantor wives was a specific sub species of a wider genus of unconscionability, although separate and distinct from the particular unconscionability principles outlined in *Amadio*.[43] Like the *Amadio* doctrine, the special wives' equity offers relief only from procedural unfairness. It is the failure to inform fully a volunteer wife, who has reposed trust and confidence in her husband, of the nature and effect of the guarantee that gives the rise to the equity to set it aside. No equity arises by virtue of the wife having entered into an improvident or unfair transaction.

Whether the unwritten law also incorporates other broader notions of unconscionability which extend to specific equitable doctrines such as estoppel, unilateral mistake, and relief from forfeiture is still open for debate.[44] It is worth noting that the High Court majority in *Garcia* referred to a passage from the judgment of Mason J in *Amadio* where his Honour said: 'It goes almost without saying that it is impossible to describe definitively all the situations in which relief will be granted on the ground of unconscionable conduct'.[45] The words 'within the meaning of the unwritten law from time to time' reinforce this conclusion, as they clearly envisage that the parameters of unconscionability are susceptible to variation and refinement from time to time, and are not restricted to the law as it stood at the time the section was enacted.[46]

---

42 (1998) 194 CLR 395

43 (1998) 194 CLR 395, 408. See Horrigan, above, n 31, 174 who accepts that the unwritten law covers both *Amadio* and the special wives' equity: 'The flavour of the language in *Garcia* suggests an acceptance by the High Court of different strands of unconscionable conduct in this area of the law'.

44 The Full Federal Court in *ACCC v Samton Holdings* (2002) 117 FCR 301, 318, set out five categories of unconscionable conduct which would support the grant of relief on principles set out in specific equitable doctrines. See also Gino Dal Pont, 'The Varying Shades of "Unconscionable Conduct"-Same Term Different Meaning' (2000) 19 *Australian Bar Review* 135. Horrigan, above n 31, 175, who notes that the meaning of s 51AA requires clarification by the High Court.

45 (1998) 194 CLR 395, 408.

46 In *ACCC v Samton Holdings Pty Ltd* (2002) 117 FCR 301, 319, the Full Federal Court held that the use of these words in the statute overcomes the Explanatory Memorandum's suggested limitation of s 51AA to specific *Amadio* type unconscionability.

Limitations of Unconscionability under the Unwritten Law

Guarantees cases such as *Amadio* and *Garcia* have provided a paradigm for judges to explore the limits of unconscionability, in both its generic and specific forms. By its express reference to *Amadio*, the Explanatory Memorandum accompanying the introduction of s 51AA TPA makes clear that the general statutory provisions incorporating the judge-made law apply to guarantees. What remains unclear is the scope of the unwritten law. This is important to determine in the guarantee context, because while s 12CC of the ASIC Act remains subject to the $3 million threshold the unwritten law will continue to be relevant to disputes in the financial services sector, including, inter alia, disputes involving guarantees.

The Amadio doctrine-Situational disadvantage?

Section 12CA provides a statutory enactment of the *Amadio* principles, enabling a contract to be set aside where:

> one party was in a position of special disadvantage[47] vis-à-vis the other party and in the circumstances the other party knew or ought to have known of that special disadvantage; and
>
> - the other party has taken advantage of that situation.[48]

The *Amadio* doctrine focuses on the circumstances surrounding the taking of the guarantee and the 'unwritten law' within the meaning on the scope of s 12CA is accordingly confined to procedural unfairness. It does not cover substantive unfairness which occurs through the imposition of harsh contractual terms, although manifest inadequacy of consideration or some other kind of harsh outcome may well be evidence of unconscionable conduct.[49]

---

47 The classic statement of Fullagar J in *Blomley v Ryan* (1956) 99 CLR 362 , 405 referred to 'poverty or need of any kind, sickness, age, sex infirmity of body or mind, drunkenness, illiteracy or lack of education, lack of assistance or explanation where assistance or explanation is necessary'.

48 (1983) 151 CLR 447, 462 (Mason J).

49 See, for example, *Blomley v Ryan* (1954) 99 CLR 362. A more recent case where substantive factors were given some weight is *Familiar Pty Ltd v Samarkos* (1994) 115 FLR 443 (Supreme Court of Northern Territory) where an administration fee of $500,000 was charged for a loan of an equivalent amount. Ultimately, however, the court's decision was based on the fact that the defendant took unconscientious advantage of the plaintiff's particular financial vulnerability, in this case, his desperate financial need.

In *ACCC v CG Berbatis Holdings Pty Ltd* (*'Berbatis'*) [50] the High Court confined s 51AA TPA within narrow boundaries. The majority, referring to the *Amadio* test, emphasised that a mere inequality in bargaining position would not suffice to establish unconscionable conduct. The weaker party would have to suffer from a disadvantage or disabling condition that was 'special' in the sense that it seriously impaired their capacity to make a judgment about their best interests.[51] It will not avail for the weaker party to argue that they have been commercially pressured into making a 'hard bargain' as a result of an inevitable disparity in bargaining position.

At first instance in *Berbatis*[52] French J had argued that special disability might also encompass circumstances, which he termed 'situational disadvantage', arising from a position of acute financial or legal disadvantage as well a special disadvantage in the traditional sense of some inherent personal difficulty.[53] The High Court, while dismissing the ACCC's claim on the facts, did not determinatively rule that 'special disability' might not be based on situational factors, however the thrust of the majority judgments favoured the traditional *Amadio* approach.

More recently, in *ACCC v Oceana Commercial Pty Ltd*[54] Keifel J, considering the scope of s 51AA TPA, observed:

> One party to a commercial transaction will often know much more than the other. A difference in bargaining power, even a substantial difference, does not amount to the 'special disadvantage' of which Mason J spoke in *Amadio* and as was further discussed in *Berbatis*… The parties in *Berbatis*, as Gleeson CJ observed, were at a distinct disadvantage in seeking an extension or renewal of their lease but there was nothing 'special' about that. Critically, they did not suffer from a lack of ability to judge or protect their financial interests. [55]

The *Berbatis* approach was also followed recently by the New South Wales Court of Appeal in *ANZ Banking Group Ltd v Karam*,[56] a guarantee case concerning alleged unconscionability under the general law and the *Contracts Review Act 1980* (NSW). The decision confirms that a finding of unconscionable conduct will not be supported in the absence of a special disability or disadvantage on the part of the weaker party. An interesting issue which arose in the *Karam* case concerned the interaction of the

---

50 (2003) 214 CLR 51.

51 (2003) 214 CLR 51, 77-78 (Gummow and Hayne JJ), at 11 (Gleeson CJ). Kirby J dissented.

52 *ACCC v C G Berbatis Holdings Pty Ltd* [2000] ATPR ¶41-778, 41,196–7.

53 Above n 47.

54 [2003] FCA 1516.

55 Ibid [329].

56 [2005] NSWCA 344.

doctrines of unconscionability and economic duress. Like *Berbatis*, the *Karam* case involved parties entering a contract under circumstances of acute financial difficulties, of which the other party, in this case the bank, was aware. The trial judge made a finding of unconscionability based on his conclusion that illegitimate economic pressure, amounting to economic duress, had been imposed by the bank. However, the NSW Court of Appeal rejected the finding of illegitimate pressure, concluding that a claim to relief on the basis of pressure which is not in itself unlawful should be determined under the equitable doctrines of undue influence, unconscionability and relevant statutory remedies and not under economic duress. [57] Although the Karams faced perilous financial circumstances, this was not of the bank's making and it could not be said that there was anything unlawful or unconscionable about the bank's conduct. Therefore, it seems that the general law doctrine of unconscionability remains well constrained within its traditional boundaries.

The question arises whether this narrow interpretation of s 51AA will carry across to the small business unconscionability provisions. For the time being though, it is appropriate to conclude that unconscionability under the unwritten law does not include conduct, which, without more, is merely unfair, unreasonable or which amounts to a 'hard bargain'.

If provisions codifying the unwritten law covered only *Amadio* type unconscionability, this would significantly curtail their effectiveness in the guarantee context, given the large number of cases brought by women who have guaranteed their husbands' debts. Proof that the lender knew, or ought to have known, of the guarantor's disability and unfairly sought to take advantage of that disability is very difficult to establish. Prior to the affirmation of the special wives' equity in *Garcia* there was a long list of failed unconscionability cases where the wife failed to prove actual or constructive knowledge of a special disability on the part of the lender.[58] The *Garcia* case itself involved a failed *Amadio* defence. However, there are good reasons, explained above, and accepted in the recent case of *ACCC v Samton Holdings Pty Ltd*,[59] for regarding the special wives' equity as part of the unwritten law for the purposes of s 12CA.

---

57 Ibid [95]-[96].

58 See, for example, *Akins v National Australia Bank Ltd* (1994) 34 NSWLR 155*; National Australia Bank v Garcia* (1996) 39 NSWLR 577; *European Asian of Australia Ltd v Kurland* (1985) 8 NSWLR 192; *Warburton v Whiteley* (1989) NSW Conv R 55-453*; ANZ Banking Group Ltd v McGee* (1994) ASC 56-278.

59 (2002) FCR 301, 309-10.

The Garcia doctrine –Extending to relationships other than marriage?

The rule in *Garcia* although a narrow one, applying only to marital relationships where the wife is substantially regarded as a volunteer, nevertheless has the significant advantage of enabling a wife to resist enforcement of a guarantee on the grounds that she lacks full understanding of its effect. The rationale for the rule, as emphasised by the majority, was to be found in the special nature of the marriage relationship which is based on trust and confidence between the spouses.[60] The majority in *Garcia* left open the question of whether the *Yerkey* principles might apply in the future to 'long term and publicly declared relationships short of marriage between members of the same or opposite sex'.[61]

Cases since *Garcia* have adopted an inconsistent approach to the extension of the special wives' equity to other committed relationships based on trust and confidence. There have been several conflicting first instance decisions and the appellate decisions suggesting an extension of the doctrine to other relationships have been based on obiter comments. Despite some signals to the contrary, the courts have generally been slow to extend the wives' equity principles beyond the narrow boundaries of legal husband/wife relationships. In *Watt v State bank of NSW*[62] the ACT Court of Appeal held that the law in Australia had not yet developed to the stage where the *Garcia* principles extended to parent/child relationships. In *State Bank of NSW v Hibbert*[63] Bryson J declined to extend the principles in *Garcia* to de facto relationships. However, in the more recent case of *Liu v Adamson*[64] Master Macready in the same court preferred the view that de facto relationships were covered by the *Garcia* principles.

In the Victorian Court of Appeal case of *Kranz v National Australia Bank*[65] Charles JA, with whom Winneke P and Eames JA concurred, disagreed with the necessity for confining the principles in *Garcia* to only 'the most intimate of family relationships'. He held that a surety could rely upon a *Garcia* defence whenever the bank was aware of a relationship which put the bank on inquiry. [66] Although only an obiter opinion, as the plaintiff failed to show the bank knew or ought to have known of the special relationship of trust and confidence between him and the debtor, it indicated a softening of attitude at the appellate level, leaving the door open a little further for

60 (1998) 194 CLR 395, 403 (Gaudron, McHugh, Gummow and Haynes JJ).

61 Ibid 404.

62 [2003] ACTA 7.

63 [2000] NSWSC 628.

64 [2003] NSWSC 74.

65 [2003] VSCA 92.

66 Ibid [24].

future cases. Unfortunately, from the point of view of certainty of the law, the High Court refused to grant special leave to appeal in the *Kranz* case.

Subsequently, in the Queensland Court of Appeal case of *ANZ Banking Group Ltd v Alirezai*[67] the majority judges agreed that the *Garcia* principle applied whenever there was a relationship of trust and confidence between the debtor and surety. Like the *Kranz* case, the evidence in *Alirezai* failed to disclose that the bank was put on inquiry. Until the High Court determines the issue it is not clear whether the doctrine is still narrowly confined to matrimonial relationships. An application for special leave to appeal the *Alirezai* decision has been lodged with High Court. As both *Kranz* and *Alizerai* demonstrated, presently in cases outside of marriage, sureties will face major hurdles in proving that the financier had notice of a special relationship of trust and confidence between debtor and surety.

Ultimately, whilst there is scope for movement in interpretation of the elements of equitable doctrine, this has not happened to any significant degree in cases applying either *Amadio* or *Garcia* principles. Under the unwritten law, notice of special disability or of a relationship of trust and confidence plays a crucial role. As the post *Garcia* litigation demonstrates, this requirement seems to be a significant obstacle in the extension of the *Garcia* principles beyond spousal relationships. The straightforward test adopted in the UK,[68] whereby a financier is put on notice in any case where the relationship between the debtor and surety is non-commercial, would bring greater certainty to this area of the law. However, in Australia the equitable principles in relation to guarantees differ markedly from the direction of the courts in the UK[69] and there has been no indication that the UK test will gain approval in Australia. In any event, the development of extensive disclosure regimes, at least for mainstream financiers,[70] make it much harder for a case of unconscionability to be made out on the basis of lack of understanding of the nature and effect of guarantee.

Conclusions on the 'Unwritten Law'

Recent cases indicate the limitations of the 'unwritten law'. First, it is concerned with procedural rather than substantive unfairness. Even then, unfairness in the bargaining

---

67 [2004] QCA 6.

68 *Royal Bank of Scotland v Etridge (no 2)* [2002] 2 AC 773 , 814 (Lord Nicholls).

69 The High Court in *Garcia* (1998) 194 CLR 395 at explicitly rejected the House of Lord's approach in *Barclays Bank v O'Brien* [1994] 1 AC 180, to setting aside spousal guarantees.

70 See the detailed precontractual requirements under Clause 28.4 of the revised Code of Banking Practice, which apply to guarantees given to support loans to individuals and small business. As at 31 December 2005, 13 banks had subscribed to the Code.

process is viewed in narrow terms, as the *Berbatis* decision indicates. This is in line with the traditional reluctance of courts to rewrite contracts.[71] There has been a reluctance to extend equity's assistance to cases which do not fall squarely into the principles outlined by the High Court in *Amadio* and *Garcia*. Secondly , by its very nature the 'unwritten law' develops on an incremental basis, which, as the post *Garcia* litigation shows, can produce uncertainty about the ambit of unconscionability in the guarantee context. This limits the usefulness of s 12CA.

Given these limitations, it seems likely that in the future there will be an expanded role for s 12CC ASIC Act, governing small business unconscionability, particularly if the proposal for raising of the transactional threshold from $3 million to $10 million takes effect. The majority of recent cases involving disputed guarantees concern guarantees supporting small business debts incurred by spouses and others in a close relationship to the guarantor.

Guarantees and Consumer Unconscionability

Section 12CB(2) generally mirrors s 51AB(2) TPA in setting out the following non-exhaustive factors to which a court may have regard in determining whether conduct was unconscionable:

- the relevant bargaining positions of the parties (para. (a));
- whether, as a result of the conduct the consumer was required to comply with conditions that were not reasonably necessary for the protection of the legitimate interests of the corporation (para. (b));
- whether the consumer was able to understand any documents relating to the supply of goods or services (para. (c));
- whether any undue influence or unfair tactics were used (para. (d)); and
- the amount for which and the circumstances under which the consumer could have obtained the goods or services from an alternative supplier (para. (e)).

The list of factors enumerated in various consumer provisions, including s 12CB ASIC Act, s 51AB TPA and s 70 UCCC include some which suggest that the sections may go

---

[71] As long ago as *Maynard v Mosely* (1676) 3 Swan 651, 655, Lord Nottingham noted that 'Equity cases are replete with reminders that equity mends no man's bargain'. The words of Lord Hodson in *White and Carter (Councils) Ltd v McGregor* [1962] AC 413, 445 that 'It is trite law that equity will not rewrite an improvident contract where there is no disability on either side' foreshadow the development of unconscionability principles in Australia.

beyond procedural unconscionability.[72] However, the reference to 'all the circumstances' which is incorporated into all of these provisions has been generally interpreted as directing the court's attention to matters surrounding the making of the contract and not to whether the contract terms may themselves be unfair.[73]

The consumer provisions have a limited role to play in assisting disaffected guarantors given that the vast majority of third party guarantees are given to support small business borrowing[74] and that most guarantee litigation has concerned guarantees given to support small business debts.[75] The case of *Begbie v State Bank of New South Wales*[76] highlights this limitation. In that case the plaintiff's claim to have a mortgage and guarantee set aside under s 52A TPA (the predecessor of s 51AB) failed because the primary loan facility was not a service 'of a kind ordinarily acquired for personal, domestic or household use'. The case involved the common scenario of a vulnerable female guarantor acting as a surety to support the business debts of her partner. Drummond J noted that the borrowing of funds, even of a substantial amount, sufficient to enable the borrower to buy a private residence would be a service of the kind covered by the consumer unconscionability provisions.[77] Consumer guarantees are specifically targeted by the UCCC,[78] which operates under the aegis of state government consumer agencies. Section 70 (1) UCCC prohibits 'unjust' contracts, with a list of factors guiding the meaning of 'unjust' set out in s 70(2). Like their ASIC Act/TPA counterparts, in practice these factors have been regarded as targeting procedural unfairness.[79]

---

72 For example, s 12CB (b) and (e).

73 See discussion in SCOCA Discussion Paper, *Unfair Contract Terms*, January 2004 [2.1.1] above n 18.

74 New South Wales Law Reform Commission, Report on the Practice of Third Party Guarantees in New South Wales, *Darling please sign this form*, Research Report II, October 2003 p xiii.

75 See, for example, *Yerkey v Jones* (1939) 63 CLR 649 and more recently *Garcia v National Australia Bank Ltd* (1998) 194 CLR 395; *Bylander v Multilink Investments Pty Ltd* [2001] NSWSC 53; *State Bank of New South Wales v Chia* (2000) 50 NSWLR 587; *Burrawong Investments Pty Ltd v Lindsay* [2002] QSC 82.

76 (1994) ATPR ¶41-288.

77 Ibid 41, 898.

78 Sections 50-57.

79 See SCOCA Discussion Paper, above n 18 [2.1.3].

Guarantees and Small Business Unconscionability

Both ss 12CC of the ASIC Act and 51AC TPA regulate unconscionable conduct involving financial services.[80] As explained earlier 'financial services' are defined in s 4 (1) TPA to effectively have the same meaning as in the ASIC Act. The ASIC Act's definition of 'financial services' now captures conduct involving the giving of advice about guarantees and conduct involving the dealing in a guarantee of obligations under a credit contract.[81] Although the current ACCC guidelines[82] on unconscionable conduct affecting small businesses make no reference to guarantees, ss 12CC ASIC Act and 51AC TPA potentially have a flexible and far-reaching application in the guarantee context.

Section 12CC generally mirrors s 51AC in that it excludes transactions above a $3 million threshold where the 'victim' is a listed public company. There are slight differences in the wording of the two sections.[83] Both s 12CC and s 51AC TPA are closely modeled on the domestic unconscionability provisions, albeit with a more extensive list of factors set out in s 12CC (2). The enumerated factors cover both procedural and substantive unfairness. These factors include those listed in s 12CB (2) as well as the following:

---

80 Section 51AAB TPA does not exclude s 51AC from applying to 'financial services'.

81 See n 36 above. This overcomes an earlier problem of statutory construction which arose when s 51AB was introduced into the TPA. The drafting of s 51AB (and later 51AC), and of the various factors listed in those two sections, is directed towards two party situations, not a three party situation such as the giving of a guarantee to support financial accommodation provided by a lender to a principal debtor. If s 51AC is restricted to bi-partite dealings between the 'supplier' and 'business consumer' then guarantors of small business debts are still governed by s 51AA which incorporates the *Amadio* principles and was given a narrow interpretation by the High Court in *Berbatis*. Although there were strong arguments for suggesting that the TPA domestic and small business unconscionability provisions applied to tripartite situations such as those involving third party sureties, the position was not entirely clear. See Mark Sneddon, 'Banking and Finance: Guarantees and the Trade Practices Act, Section 51AB' (1994) 22 *Australian Business Law Review* 368, who argues that the provisions did extend to transactions involving guarantees.

82 *A Small Business Guide to Unconscionable Conduct*, Australian Competition and Consumer Commission, May 2005 (www.accc.gov.au).

83 For example, s 12CC ASIC to unconscionable transactions involving a 'supplier' and 'service recipient', while s 51AC TPA refers `to a 'supplier' and 'business consumer'. For a discussion of the minor differences in the wording of the two provisions see Pearson above, n 31, at 121-12.

- the extent to which the supplier/recipient's conduct was consistent with its dealings with other small business parties (paras (2)(f) and (3)(f));
- the requirements of any applicable industry code (paras (3)(g) and (4)(g));
- the requirements of other industry codes which the other party might reasonably have expected to apply (paras (2)(h) and (4)(h));
- whether the supplier/acquirer unreasonably failed to disclose any relevant intended conduct and associated risks to the small business party (paras (2)(i) and (3)(i));
- the extent to which the supplier/acquirer was willing to negotiate the terms and conditions of any contract (paras (2)(j) and(3)(j)); and
- the extent to which the parties acted in good faith (paras (2) (k) and (3) (k)).

The relationship between the general law of unconscionability and the provisions governing consumer and small business unconscionability is still being formulated by the courts and is yet to be considered by the High Court. However, the language of the small business provisions and in particular the inclusion of detailed indicia of unconscionability, clearly give them a wider application than the 'unwritten law' which covers traditional notions of unconscionability as articulated in *Amadio* and *Garcia*. This construction is supported by the weight of opinion, both judicial and academic.[84]

Unlike the 'unwritten law' provisions the small business provisions do not depend on a finding of a 'special disability', but they do require a finding of conduct which is 'in all the circumstances' unconscionable. What is still to be resolved is whether the finding of the High Court in *Berbatis* that conduct which can be described as a 'hard bargain' is not unconscionable applies to the small business provisions.[85] The answer seems to depend on the threshold test defining unconscionable. A low threshold test for unconscionability would catch conduct which is 'unfair'. In *Garry Rogers Motors (Aust) Pty Ltd v Subaru (Aust) Pty Ltd*, Finkelstein J stated in respect of s 51AC TPA: 'I take as the measure of unconscionability, conduct that might be described as unfair'.[86]

---

84 *ACCC v Oceana Commercial Pty Ltd* [2003] FCA 1516 [336]; *Securities & Investments Commission v National Exchange Pty Ltd* [2005] FCAFC 226, [29]-[30]; Gail Pearson, above n 31, 123-124; The Hon Paul de Jersey AC, above n 12, 4 and Bryan Horrigan, above n 31, 162.

85 As suggested by Robert Baxt and Elizabeth Bennett, 'Unconscionability and the Berbatis case', June 2003, available at Allens Arthur Robinson website (www.aar.co.au).

86 (1999) 21 ATPR ¶41-703, 43, 016.

A higher level test would require something in addition to 'unfairness', such as conduct offending against good conscience or involving lack of morality. It seems that in practice the courts have opted for the higher threshold test. In *Hurley v McDonald's Australia Ltd*,[87] the Full Federal Court took as its starting point the Shorter Oxford Dictionary definition of 'unconscionable', i.e., 'showing no regard for conscience; irreconcilable with what is right or reasonable'. Beyond this, the Full Court further observed:

> For conduct to be regarded as unconscionable, serious misconduct or something clearly unfair or unreasonable, must be demonstrated...The various synonyms used in relation to the term' unconscionable' import a pejorative moral judgment.[88]

In *Australian Securities & Investments Commission v National Exchange Pty Ltd*[89] the Full Federal Court also accepted that unconscionability is a concept that requires a high level of moral obloquy. Consequently, few of the test cases brought to establish the ambit of the statutory unconscionability provisions have been successful. The cases have generally affirmed the traditional role of the courts in upholding contractual bargains. A bargain entered into under pressure that is 'unfair', without more, will not be overturned on unconscionability grounds. Perhaps one of the limitations in the small business provisions is the absence of the word 'unfair' as a marker of unconscionable conduct.[90] It is interesting to note that the 1997 Reid Report, in its recommendations to the Federal Government for enhanced small business protection, cautioned against using the term 'unconscionable' indicating:

> ...[I]t would be better to use a new word, without the legal entailments of 'unconscionability' while avoiding the words 'harsh' or 'oppressive' in the initial clause establishing the standard.[91]

It was recommended that the word 'unfair' be used instead. The Federal Government clearly had reservations about the breadth of the Reid Report's recommendation for what amounted to an unfair trading law in Australia, preferring to adhere to traditional concepts. The Minister said in his Second Reading Speech:

> The bill uses the expression 'unconscionable conduct' in order to build on the existing body of case law which has worked well in relation to consumer

---

87 (2000) 22 ATPR ¶41-741 (Heerey, Drummond and Emmett JJ).

88 Ibid 40, 585.

89 [2005] FCAFC 226 [43].

90 See Robert Gardini, above n 14.

91 Report of the House of Representatives Standing Committee on Industry Science and Technology, *Finding a balance-Towards fair trading in Australia*, May 1997, p 179.

> protection provisions of the Act and which will provide greater certainty to small business in assessing their legal rights and remedies.[92]

The recent Report of the Senate Economics References Committee cautioned against the use of the word 'unfair', noting that introducing the concept of unfairness into s 51AC TPA 'carries the risk of making the section unworkably ambiguous by calling on concepts with an unclear legal meaning'.[93]

Indicia of Unconscionability Under s 12CC ASIC Act

The expanded indicia of unconscionability in s 12CC go beyond traditional indicators of unconscionability under general law principles as incorporated in s 51AA. This conclusion can be supported on three grounds:

- The inclusion of factors which embrace substantive unfairness, including the proposed inclusion of a new 'unilateral variation' clause;
- Special recognition of the notion of good faith; and
- Acknowledgement of the role of industry codes.

It is therefore useful to consider those factors insofar as they are relevant, in particular, in the guarantee context.

Substantive Unfairness Factors

Matters such as the equivalent terms and price offered by competitors[94] and being forced to comply with conditions which go beyond the stronger party's legitimate interests require a consideration of whether the substantive terms of the contract are unfair.[95] Generally, however, consumer bodies have been unimpressed with so-called 'shopping lists' in various pieces of unconscionability legislation that contain terms which are apparently substantive.[96] For example, it has been observed that although s 70(2) of the UCCC contains four factors which canvass issues of substantive unfairness:

---

92 House of Representatives, Hansard, 30 September 1997.

93 The Senate Economics References Committee Report on *The Effectiveness of the Trade Practice Act 1974 in Protecting Small Business*, March 2004, p 36 (available at www.aph.gov.au).

94 Section 12CC (2) (e) and (3) (e) and its counterpart in s 51AC (3) and (4).

95 Section 12CC (2) (b) and (3) (b) and its counterpart in s 51AC (3) and (4).

96 See above n 32.

> In our opinion s 70 of the Code is primarily concerned with procedural unfairness and does not provide a tribunal to reopen a contract simply where a term of a contract is unjust. Academic consideration of the issue is divided and we are unaware of any judicial decisions in which substantive issues **in isolation** [emphasis added] formed the basis for orders reopening an unjust contract.[97]

While it is true that most of the factors enumerated in unconscionability shopping lists deal with procedural matters, there have been a few decisions in which the substantive harshness of a standard term on which the stronger party sought to rely was a major factor, or the sole factor, in the Court's decision.[98] Arguably, many broadly framed 'all moneys' clauses, typically found in third party guarantees and mortgages, could be subject to attack on the basis that they are not reasonably necessary to protect lenders' legitimate interests. Most guarantees currently used in business lending have a provision to the effect that they cover 'all moneys which are now or may hereafter from time to time be owing or payable on any account to the creditor by the security provider'. The widespread use of third party guarantees has been the subject of critical reports by government and law reform bodies.[99] Given the plethora of cases involving spouses and other vulnerable guarantors,[100] coming before the courts in which the lack of awareness or comprehension of an 'all moneys' clause was one of the main grounds for challenging the guarantee, controlling their use is necessary in guarantees given to support small business debts. Therefore, the business unconscionability provisions in the ASIC Act may provide some potential for assisting individual litigants in striking down provisions such as all moneys clauses in guarantees which are

---

97 Victorian Consumer Credit Legal Service, above n 32, 11-12.

98 In the general commercial context, see, for example, *ACCC v Black on White* [2001] FCA 372 (Unreported, Federal Court of Australia, Spender J, 4 April 2001); *John Dorahy's Fitness Centre Pty Ltd v Buchanan* (Unreported, NSWCA 16 December 1996, No 94040386).

99 Martin Committee Report, Australia, House of Representatives Standing Committee on Finance and Public Administration, *A Pocketful of Change: Banking and Deregulation* (AGPS, Canberra 1991); Trade Practices Discussion Paper (1992), *Guarantors: Problems and Perspectives;* Australian Law Reform Commission Report No.69 (1994) Part II, *Equality before the law: women's equality*; *Good Relations, High Risks: financial transactions between family and friends* Report of the Expert Group on Family Financial Vulnerability (1996); Australian Banking Industry Ombudsman Ltd, Report on Relationship Debt (Bulletin No 22, 1999).

100 See, for example, *Commonwealth Bank of Australia v Cohen* (1988) ASC 55-681; *Commonwealth Bank of Australia v ABC Property Planners Pty Ltd*, (Unreported, Supreme Court of NSW, Equity Division, Cohen J, 25 May 1991); *European Asian of Australia Ltd v Lazich* (1987) ASC 55-564; *Commonwealth Bank of Australia v Khouri* [1998] VSC 128 (Unreported, Supreme Court of Victoria, Harper J, 4 December 1998) ; *Garcia v National Australia Bank Ltd* (1998) 194 CLR 395; *State Bank of New South Wales v Chia* (2000) 50 NSWLR 587; *Karam v ANZ Banking Group Ltd* [2001] NSWSC 709.

substantively unfair. However, the provisions do not effect any systemic control over unfair terms so will not necessarily deter financiers continuing to use 'all moneys' clauses in their standard guarantee documents.

Reference to Industry Codes

Courts may take note of the requirements of 'any applicable industry code'[101] or 'any other industry code'[102] if the 'service recipient' acted under the reasonable belief that there would be compliance with that code. An applicable code refers to a prescribed mandatory industry code.[103] In the guarantee context, the incorporation of the 'code factors' into the business unconscionability provisions enables a court to consider best banking practice standards. For example, the Australian Bankers' Association Code of Banking Practice[104] covers banking services generally and also contains detailed provisions which specifically relate to guarantees. It extends to guarantees over domestic as well as small business debts.[105] Adoption of the self-regulatory Code is voluntary and not prescribed under s 51AE TPA. Therefore, it can be considered by the court to be 'any other industry code' for the purposes of the factors in s 12CC.

The Banking Code contains several clauses which are relevant in the context of unconscionability. For example, clause 2.2 provides an undertaking by banks that:

> We will act fairly and reasonably to you in a consistent and ethical manner. In doing so we will consider your conduct, our conduct and the contract between us.

Clause 25.1 states that banks will:

---

101 Section 12CC (2) and (3) (g).

102 Section 12CC (2) and (3)(h).

103 Section 51AE provides that regulations may prescribe a code for the purposes of Part IVB TPA. To date the only prescribed code is the Franchising Code of Conduct.

104 *Code of Banking Practice*, Australian Bankers' Association, 2003 (Revised August 2003, modified May 2004). The Code can be obtained from the Australian Bankers' Association website (www.bankers.asn.au).

105 Clause 40 defines 'small business' any business which has less than 20 employees or, in the case of a manufacturing business, less than 100 employees. It further provides that a banking service which is a 'financial product' or 'financial service' applies to 'retail clients' as described in Chapter 7 of the *Corporations Act 2001* (Cth). The definition of 'retail client' includes a small business, which carries the same meaning for the purposes of both the *Corporations Act* and Banking Code: ss 761G (7) and (12) *Corporations Act*.

> Exercise the care and skill of a diligent and prudent banker in selecting and applying our credit assessment methods and in forming our opinions about your ability to repay….

This commitment seems relevant to the practice of 'asset-based' lending, which has been recogised as an undesirable feature of many guarantee transactions.[106] This refers to a situation the bank has doubts about both the borrower's and guarantor's ability to be able to repay the loan, but rely on security in the case of default.[107] Clause 25.2 states that:

> With your agreement we will try to help you to overcome your financial difficulties with any credit facility you have with us.

Clause 28.4 prescribes a disclosure and conduct regime for banks which have adopted the code.[108] The Code also permits, but imposes restrictions upon, the taking of 'all moneys' securities.[109] These provisions are directed at redressing problems concerning lack of procedural fairness highlighted in high profile guarantee litigation which had reflected unfavourably upon the practices of banks and other financiers in taking guarantees.

It has been considered that the Banking Code would be given considerable weight in informing courts of the standards required by banks and other financiers,[110] particularly given that it has been established and endorsed by industry after widespread

---

106 The New South Wales Law Reform Commission Report on the Practice of Third party Guarantees in New South Wales, *Darling please sign this form*, Research Report II, October 2003 [4.58]. See also Banking and Financial Services Ombudsman, Policies and Procedures Manual, p 24 (www.bfso.org.au) which states that 'No banker should rely on realisation of assets held as security as the primary source of repayment…'.

107 For a critique of the practice of 'asset-based' lending and the limitations of the Banking Code in that context, see Nicola Howell 'Ethical Lending: a corporate social responsibility or a legal responsibility', Paper presented to the 2nd National Consumer Credit Conference, Melbourne, 8 November 2004 (www.gu.edu.au/centre/cccl/content_publications.html).

108 Clause 28.4 (b) and (c) refers to disclosure of demands, excesses and overdrawings in relation to the debtor as well as the provision of certain documents to the guarantor such as credit reports, statements of financial accounts, relevant credit reports and the final letter of offer to the debtor.

109 Clause 28.13.

110 See, eg, Elizabeth Wentworth, *The New Code of Banking Practice and Guarantees::What Lawyers need to Know*, General Counsel to the Banking and Financial Services Ombudsman, Paper presented at the Law Council of Australia/Law Institute of Victoria Financial Services Seminar, 16 October 2002.

consultation.[111] Therefore, it seems likely that the Code has the potential to be relevant to cases brought under s 12CC ASIC Act, even in evaluating the conduct of a financier who had not adopted the Code.[112]

Despite the comprehensive revision of the guarantee provisions of the Code,[113] some aspects of its treatment of all moneys clauses are unsatisfactory and difficult to reconcile with the general obligation of unfairness in Clause 2.2. Clause 28.13 of the Code permits all moneys clauses in a way which significantly undermines the protection to guarantors. Clause 28.13 provides that:

> A guarantee given by you will be unenforceable in relation to a future credit contract unless we have:
>
> (a) given you a copy of the contract document of the future credit contract; and
>
> (b) subsequently obtained your written acceptance of the extension of the guarantee
>
> except to the extent the future credit contract (together with all other existing credit contracts secured by that guarantee) is within a limit previously agreed in writing by you and we have included in the notice we give you… a prominent statement that the guarantee can cover a future credit contract in this way.

The qualification allowing the guarantee to be extended to future contracts in the manner spelt out in clause 28.13 means that guarantees will be enforceable in relation to further advances that a guarantor may not be aware of, or has agreed to at the time, providing that a statement has been given that this can occur and the lending is within the previously agreed limit. Providing an initial warning purporting to cover transactions which may occur many years later is not sufficient protection and the effect of clause 28.13 is not in the spirit of the general requirement to act fairly and reasonably in Clause 2.2, particularly given that unfairness relating to 'all moneys' provisions has been at the heart of many of the leading guarantee cases.[114] Clause 28.13 is a weaker provision than s 54 UCCC, which requires the specific written consent of the guarantor at the time of the future advance. Therefore, despite the fact that a guarantee with an 'all moneys' clause complies with the Banking Code and requires a

---

111 The revised version of the Code, which came into effect in August 2003, adopts the recommendations of the Viney Report: Richard Viney, RTV Consulting Pty Ltd, *Review of the Code of Banking Practice*, Final Report, October 2001.

112 Elizabeth Wentworth, above n 110, 5.

113 Adopting the recommendations of the Viney Report, above n 111.

114 See cases referred to above n 29.

generic warning of the risks involved, it may be still have been presented on a take it or leave it basis, drafted incomprehensibly and without specific warnings of the degree of risk involved or requiring written acceptance at the time of later advances.

The meaning of the words in the Code 'acting fairly and reasonably to you' and whether a court would conclude that a guarantee with an 'all moneys' clause which prima facie complies procedurally with the Banking Code is unconscionable are both uncertain. Therefore, although the inclusion of factors permitting reference to the terms of industry codes seemingly import an 'industry' standard of fairness into the business unconscionability provisions, it is by no means clear what this standard encompasses and therefore whether the standard will be higher than under existing equitable principles. Moreover, the Banking Code, obviously covers a specific section of the financial services sector and not all banks have subscribed to the Code.[115] Other non-bank lenders are not under the same obligation to act 'fairly and reasonably' although, as noted, the requirements of the Code could well be regarded as a benchmark for other credit providers when taking third party guarantees.

Reference to Good Faith

'Good faith' is an elusive concept, but one that is receiving increasing prominence in Australian jurisprudence and academic writing, including a debate as to whether such an obligation should be implied into the contract.[116] The complexities of the debate are beyond the scope of this paper, but some general observations can be made.

The High Court in *Royal Botanic Gardens and Domains Trust v South Sydney City Council*[117] considered, but found it unnecessary to decide, whether such an obligation exists and therefore the matter is unsettled. On the whole, courts have taken a strict approach to the implication of term into a contract. The focus, from the point of view of whether there has been a breach of an implied contractual term, is on the extent to

---

115 As at 2005, 13 banks have subscribed. See Australian Bankers' Association website: www.asn.au.

116 Recent examples include *Alcatel Australia Ltd v Scarcella* [1998] NSWSC 483 (Unreported, New South Wales Court of Appeal, Sheller, Powell and Beazley JJA, 16 July 1998) and *Burger King Corporation v Hungry Jacks Pty Ltd* [2001] NSWCA 187. The concept that a contract imposes a duty of good faith and fair dealing on the parties is established jurisprudence in the United States and has been growing in acceptance in Australia. See Elisabeth Peden, 'Incorporating Terms of Good Faith in Contract Law in Australia' (2001) 23 *Sydney Law Review* 222 and J Paterson, 'Duty of Good faith-Does it have a Place in Contract Law?' (2000) 74 *Law Institute Journal* 47.

117 (2002) ALJR 436, 443.

which the term was obvious, reasonable and equitable, and necessary as a matter of business efficacy.[118]

However, s 51AC merely requires of whether good faith has occurred in contractual dealings, it does not elevate the concept into an implied contractual duty. What does 'acting in good faith' require? Arguably, it can incorporate a number of different concepts. Perhaps, as some argue,[119] it involves moral or ethical considerations. Others have said 'It presumably transcends the implied obligation to cooperate to secure the fundamentals of the contract. Honesty comes to mind, as does not acting capriciously or for an extraneous purpose'.[120] Presumably, good faith differs from reasonableness. The question arises whether acting in a spirit of good faith requires a party to subordinate his or her own interests, or at least not to pursue them aggressively or unreasonably. On this point, Barrett J in *Overlook v Foxtel* [121] held that it 'underwrites the spirit of the contract and supports the integrity of its character, but it does not require a party to subordinate his or her own interests provided pursuing them did not unreasonably interfere with the other party's enjoyment of contractual rights'. It is interesting to contrast this approach with earlier views in the United Kingdom regarding good faith between negotiating parties. The House of Lords in *Walford v Miles*[122] held that a duty to negotiate in good faith would be unworkable in practice, intruding into negotiations which could be broken off at any time. Recently the Victorian Supreme Court considered the implied duty of good faith in commercial contracts in the case of *Meridian Retail v Australian Unity Retail Network.*[123] It was unnecessary for the court to determine whether a term requiring good faith could be implied into the franchise agreement under consideration, but observed generally that the High Court had not yet definitely endorsed an implied term of good faith into commercial contracts. The Court noted that principles under Victorian case law indicated that:

- Courts should be reluctant to imply a term of good faith into commercial contracts as a matter of course;
- If implied such a term might mean nothing more than an obligation to act reasonable in the course of the contract; and

---

118 *BP Refinery (Westernport) Pty Limited v Shire of Hastings* (1977) 180 CLR 266.

119 Robert Baxt and Joel Mahemoff, 'Unconscionable Conduct Under the Trade Practices Act-An Unfair Response by the Government: A Preliminary View' (1998) 26 *Australian Business Law Review* 5.

120 The Hon Paul de Jersey AC, above n 12, 6-7.

121 [2002] NSWSC 17 [65]; [67].

122 [1992] 2 AC 128, 138.

123 [2006] VSC 223.

- The implication of such a term probably operates to protect a vulnerable or disadvantaged person from exploitative conduct which subverts the original purpose of the contract.

Enhanced guarantee disclosure regimes have already made inroads into the traditional principle that guarantees are not contracts *ubirrimae fidei* (that is, of the utmost good faith). However, the relationship of banker and customer has never been a fiduciary one[124] and it is doubtful whether the existence of the 'good faith factor' in statutory unconscionability provisions will impose further requirements upon financiers than already exists under disclosure regimes. Whether the absence of good faith will constitute unconscionable conduct will invariably depend on a consideration of all the circumstances surrounding the particular transaction.

Proposed Insertion of New 'Unilateral Variation' Factor

The presence of a provision giving the stronger party the ability unilaterally to vary any clause in the contract may be a matter going to substantive unfairness, although the financial services industry regards such clauses as necessary and not inherently unfair.[125] Unilateral variation of a credit contract will invariably result in a change in liability under a guarantee given to secure that loan. The *Trade Practices Amendment (Small Business Protection) Bill 2005* [126] proposes to introduce a package of reforms in response to recommendations made by the 2002 Dawson Review[127] and 2003 Senate Inquiry.[128] The Bill includes an amendment to s 51AC TPA providing that the 'shopping list' of factors be expanded to include reference to whether the contract

---

124 Except in exceptional circumstances: *Lloyds Bank v Bundy* [1975] QB 326; *Commonwealth Bank of Australia v Smith* (1991) 102 ALR 453; *National Bank of Australia Ltd v Finding* [2001] 1 Qd R 168.

125 See eg, Submissions of the Australian Association of Permanent Building Societies (7 April 2004) and Australian Bankers' Association (8 April 2004) to the *Unfair Contract Terms* Discussion Paper, referred to above n 18. The financial sector argues that their unilateral right to vary a contract should be retained in order to accommodate regulatory and market changes, to incorporate innovative developments in banking and credit services and to ensure compliance with prudential requirements.

126 See above n 16. Presumably there will be mirror amendments in the ASIC Act covering unconscionability with respect to 'financial services'. The legislation was proposed for introduction in the 2005 Spring Sittings of the Federal Parliament, but did not proceed.

127 2002 Review of the Competition Provisions of the *Trade Practices Act 1974* (the 'Dawson Report').

128 The Senate Economics References Committee Report on *The Effectiveness of the Trade Practice Act 1974 in Protecting Small Business*, March 2004 (available at www.aph.gov.au).

imposed terms allowing for the unilateral variation of any of its terms. Consumer bodies have argued that the right unilaterally to vary a contract may be regarded as unfair and should be dealt with in proposed uniform unfair contracts legislation. On the other hand, participants in the banking and finance sector, have argued for the continued right to unilaterally vary contracts as a matter of necessity, given the predominance of long term contracts in the industry [129] Ultimately, the Federal Government adopted the middle ground in accepting that such clauses should not be banned outright, but may in some cases be an indicator that unconscionable conduct has occurred 'in the bargaining process'.[130] The inclusion of this new factor may mean that it will be harder for financiers to justify why a unilateral variation clause should be included, although presumably reference to the 'legitimate interests' factor will be relevant in standard credit and guarantee contracts. In that context the new factor may significantly alter the ambit of the unconscionability provisions by requiring financiers to justify the inclusion of terms which can be altered unilaterally on the basis that they are necessary for the protection of their legitimate interests.

Impediments Imposed by the Phrase 'In all the Circumstances'

In *ACCC v Simply No Knead (Franchising) Pty Ltd*[131] Sundberg J, commenting on s 51AC TPA, provided a reminder that 'the Court is aided but not controlled by the factors listed in subs (3). Ultimately, although the additional factors in s12CC give it a wider ambit than the 'unwritten law', the statute requires conduct to be unconscionable "in all the circumstances" '. The inclusion of these words has been regarded by consumer bodies,[132] as placing the primary focus of these provisions on procedural fairness and thus as an impediment to allowing redress for contracts with terms which are inherently unfair. The Standing Committee of Officials of Consumer Affairs indicated that courts have been reluctant to find unfairness solely on substantive grounds.[133] This conclusion seems consistent with the view of the Full Federal Court in *Hurley v*

---

129 The Australian Bankers Association, above n 125, p 3, has argued against the proposition that clauses permitting unilateral changes should be regarded as unfair *per se,* stating that the concept of unfair terms needs to be considered with due regard for the operating context of financial services contract, particularly credit contracts.

130 See Press release, Hon P Costello 'Australian Government Response to the Senate Inquiry into the Effectiveness of the Trade Practices Act 1974 in Protecting Small Business' (available at www.treasurer.gov.au/tsr/content/pressreleases/2005).

131 [2000] FCA 1365 [37].

132 See, for example, Victorian Consumer Credit Legal Centre, Submission (11 March 2004) to the *Unfair Contract Terms* Discussion Paper, above n 18, 12.

133 Discussion Paper, *Unfair Contract Terms*, ibid n 20.

McDonald's Australia Ltd.[134] This is of concern so far as banking and finance contracts are concerned, given that the imposition of standard terms in many cases go beyond what is necessary for the protection of a financier's legitimate interests. As this paper has emphasized, the common inclusion of 'all moneys' clauses in guarantees is an example of potential substantive unfairness. Unfortunately, the usual response of a court is to set aside a guarantee if it has been given in circumstances which are procedurally deficient, but it will rarely intervene solely on the grounds that its term are inherently unfair to the guarantor. Nevertheless, it is to be hoped that the Federal Government proceeds with its small business unconscionability reforms in the near future. The raising of the monetary threshold to $ 10 million, together with the addition of the additional factor requiring a court to consider whether a unilateral variation clause is unconscionable in all the circumstances will no doubt be useful reforms, particularly in the financial services context.

Statutory Unconscionability and the Role of the Regulator

Despite some of the developments affecting statutory unconscionability regimes, discussed above, they have generally remained focused on providing remedies for individual litigants adversely affected by procedural irregularities. The costs and hurdles of instigating litigation are obvious. While landmark decisions such as *Amadio* and *Garcia* have flow on effects which may impact generally upon the conduct of financiers when taking guarantees, technically the outcome of litigation is only binding on the parties. However, there is scope for involvement of financial services regulators in regulating unconscionable conduct at a general or systemic level.

Subsequent to 11 March 2002, when s 51AF was inserted into the TPA excluding ss 51 AA and 51AB and Part V from applying to 'financial services', unconscionable conduct with respect to financial services is now the generally the responsibility of ASIC under the 'mirror' provisions of the ASIC Act, ss 12CA, 12CB and 12CC. A point of interest, given that s 51AC is not expressly prohibited from applying to 'financial services', is whether the ACCC or ASIC is the appropriate regulator to deal with unconscionability in the 'financial services' context. The ACCC's role with respect to unconscionability in financial services is preserved by ASIC's delegation of its powers to take action on its behalf to the Chief Executive Officer of the ACCC.[135] The ACCC may commence proceedings or be joined as an applicant in proceedings with ASIC, although the remedies it seeks, post 11 March 2002 will be based on ASIC Act provisions, such as the injunctive power in s 12GD, which is in substance the same in

---

134 See above, n 27.

135 As noted in *ACCC v Oceana Commercial Pty Ltd* [2003] FCA 1516 [314].

its terms as s 80 TPA.[136] However, it can be argued, that statutory construction that the ASIC should have the principal role as regulator relation to financial services.

Test Cases

In recent times the ACCC has initiated test cases[137] in matters involving unconscionability of importance to particular classes of consumers, in particular small business, pushing for a wider interpretation of the small business provisions, in, for example, the *Berbatis* case.[138] ASIC's current case against Citigroup Ltd, for insider trading is coupled with an associated allegation of unconscionable conduct, suggesting that it too is pushing for a wider interpretation of unconscionability which overlaps various kinds of market misconduct in the financial services sector.[139]

The ACCC has sought a range of orders including declarations, corrective advertising and injunctions, restraining future conduct which might affect classes of consumers generally. Some of these cases have involved the financial services sector and although very few cases have involved guarantors they potentially have implications for guarantee cases. On the whole, while the results of recent high profile public interest litigation have been disappointing in terms of the failure to establish unconscionable conduct in particular cases, the ongoing involvement of the regulator in testing the ambit of statutory unconscionability and the available statutory remedies is important. Intervention by the regulator in cases involving the public interest reinforces its proactive role in promoting normative standards of honesty, good faith and fair dealing. The courts have expressly approved of this role, exercising a degree of latitude as to the kind of orders which are appropriate for the regulator to seek in public interest unconscionability cases. For example, in *ASIC v National Exchange Pty Ltd*,[140] ASIC's claim of unconscionable conduct under s 12CC ASIC Act failed. However, the Full Court of the Federal Court dismissed the respondent's

---

136 *ACCC v Commonwealth Bank of Australia* [2003] FCA 1397 [22]. ASIC was joined as an applicant in these proceedings.

137 See, for example, *ACCC v CG Berbatis Holdings Pty Ltd* (2003) 214 CLR 51; *ACCC v Oceana Commercial Pty Ltd* [2003] FCA 1516; *ACCC v Samton Holdings* (2002) 11 FCR 301. See also the appeal by ASIC in *Australian Securities & Investments Commission v National Exchange Pty Ltd* [2005] FCAFC 226.

138 The cases have, inter alia, involved disputes about commercial tenancies, franchising and business financing. Some cases involve overlapping claims of misleading and deceptive conduct: see *Oceana*, ibid.

139 ASIC Media Release 06-096, 'ASIC commences proceedings against Citigroup for conflicts and insider trading breaches'.

140 [2005] FCAFC.

argument that the claims for declaratory and injunctive relief by ASIC were inappropriate because the appeal was an 'attempt by ASIC to obtain an advisory opinion on the law'. In the words of the Full Court:

> It is said that this appeal is an attempt by ASIC to obtain an advisory opinion on the law by means of claims for declaratory relief. Injunctive relief is said to be inappropriate because there is no threat by National Exchange to send out any further offer documents and insofar as the injunction seeks to do no more than compel National Exchange to comply with the law it is both superfluous and oppressive. In our view there is no substance to this contention. There is a live issue as to the meaning and effect of the statutory provisions. [141]

Similarly, in *ACCC v Oceana Commercial Pty Ltd*[142] the Federal court dismissed an unconscionability claim against a bank and dismissed the ACCC's application for injunctions attaching terms and conditions upon the bank in relation to future secured lending. However, Keifel J accepted that declarations and injunctions may be useful and appropriate in public interest litigation where they serve to vindicate legislation aimed at the protection of consumers.[143] Therefore, it seems that while the regulator is actively trying to assert a role in deterring unconscionable conduct in relation to financial services it is constrained by the terms of the unconscionability provisions, which, as the results of the test cases show, catch only the most blatant forms of unconscionable conduct.[144] While the outcome of some of the test cases may have been disappointing from the regulator's perspective they have demonstrated the potential width of its powers to act in appropriate cases.

Financial Sector Codes of Conduct

ASIC's powers to approve and monitor financial services sector codes of conduct[145] also provides scope for it to be proactive in generally promoting practices consistent

---

141 Id bid[51].

142 [2003] FCA 1516.

143 Ibid [343]-[344].

144 See also the arguments of Robert Gardini above n 14.

145 See s 1101A *Corporations Act 2001* (Cth) which provides ASIC with the power to approve and monitor, but not mandate financial services sector codes of conduct and ASIC Policy Statement 183 ASIC Policy Statement 183, Approval of Financial Services Sector Codes of Conduct (available at www.asic.gov.au). Until 2003 ASIC was responsible for monitoring industry compliance with the Code of Banking Practice. The terms of the revised Code, which were issued on 1 August 2003, provide for the establishment of an independent committee to be responsible for monitoring and ensuring compliance with the Code in the future.

with the objectives upon which the statutory prohibition against unconscionability are premised. While voluntary codes address industry specific issues and consumer problems not necessarily covered by legislation, under the statutory criteria for code approval in s 1101A(3) *Corporations Act*, ASIC may only approve a code which is not inconsistent with the *Corporations Act* or any other law of the Commonwealth under which ASIC has regulatory responsibility. Additionally, ASIC must consider whether a relevant code promotes the provision of 'fair, honest and professional services'.[146] These powers to approve industry codes, although not yet exercised to date, do give the regulator the power to assess the provisions of codes to determine the extent to which they are consistent with the ASIC Act unconscionability provisions. The Banking Code's latitude in the treatment of 'all money's clauses in guarantees, discussed above, could potentially come under the scrutiny of ASIC under the powers vested in it under s1101A *Corporations Act*. The Code arguably permits a bank to enforce an 'all moneys' guarantee many years after the initial guarantee had been signed in circumstances which, applying the 'not inconsistent test' may be regarded as unconscionable within the meaning of ss 12CC ASIC Act/ 51AC TPA. It is hard to hard to envisage how such action could be construed as necessary to protect the best interests of the bank.

Use of the Injunctive power

Within the context of the current ASIC/TPA Act regime, it may be possible for some degree of general control over unfair terms to be achieved through the use of the injunctive remedy, although that approach has been used very rarely. Declarations of contraventions may also play a role in generally signally accepted standard of conduct.

Under ss 12GD ASIC Act/ 80 TPA, the regulator and, indeed, any other party, can seek an injunction to restrain contraventions of the Act. An injunction might therefore be sought to restrain continuing use of an abrasive term in a financier's standard contract on the basis that such use would constitute unconscionable conduct. If granted, an injunction of this kind would obviously have a general impact. The New South Wales Law Reform Commission[147] briefly noted that the implementation of a preventative approach can be achieved through powers under the TPA.[148] Section 80 of the TPA

---

146 Policy Statement 183, ibid [47].

147 New South Wales Law Reform Commission, Issues Paper 17 (2000) *Guaranteeing Someone Else's Debts,* April 2000 [2.108]-[2.103].

148 Or similarly, in New South Wales, by the use of s 10 of the *Contracts Review Act 1980*, a provision which has rarely been used but allows the relevant Minister to apply for a court

allows the court to grant an injunction in terms it deems appropriate in relation to contraventions of the Act. The Commission said:

> It is possible that a similar [preventative} approach may be achieved by a broad power to grant injunctions to prevent conduct in breach of provisions under the trade practices legislation relating to unconscionable conduct and misleading and deceptive conduct. The provisions, however, remain untested in this regard. [149]

Since those comments, the ACCC has had recourse to the injunctive powers of the TPA in a guarantee case. The case of *ACCC v National Australia Bank Ltd*[150] indicates the potential scope of a preventative approach utilising the injunctive power in guarantee cases. The case is of significance in illustrating the flexibility of remedies and orders which can be obtained for breach of s 51AA TPA. It is important to note that these kinds of proceedings can be brought against financial institutions and lenders generally, not just banks. The ACCC brought proceedings alleging that the bank had acted unconscionably in obtaining and enforcing a personal guarantee for $200,000 from a wife as security for a business loan to a company of which her husband was a director. The ACCC alleged that when the bank sought the guarantee, it did not explain its nature or effect or advise her that she should obtain independent legal advice. The ACCC also alleged that the bank knew the borrower company was in serious financial difficulty but did not inform the wife. A year later the bank demanded payment of the company's resulting in the sale of the family which was used as security. The Federal Court declared that the bank had acted unconscionably in its dealings with the wife. While she may have made out a *Garcia* defence, the proceedings brought under s 51AA avoided protracted litigation and enabled the case to be resolved by mediation and consent orders.

The Court ordered, by consent, injunctions against the bank and one of its managers to restrain them from obtaining personal or business guarantees in Tasmania without properly explaining the nature of the guarantee and the need to obtain independent legal advice before signing the guarantee. The Court also ordered by consent that the bank include in its internal Lending Manual a statement requiring its entire lending staff throughout Australia to strictly comply with these procedures when obtaining personal consumer or business guarantees. It ordered the bank to circulate to its entire lending staff throughout Australia a bulletin to this effect.

---

order restricting the terms on which a person may enter into a contract of a specified class of contracts if the person 'has embarked, or is likely to embark, on a course of conduct leading to the formation of unjust contracts'.

149 Above n 147, [2.112].

150 Unreported, Federal Court of Australia, 5 June 2001 No T 22 of 2000.

There is thus scope for the regulator, in appropriate cases, not only to obtain a favourable result binding on the immediate parties to the litigation, but also to apply normative standards to similar cases in the future, to obtain injunctions and seek consent orders restraining lenders from using onerous 'all moneys' clauses. In the *National Australia Bank* case the ACCC brought proceedings based on a contravention of s 51AA of the TPA. There is greater scope for attacking 'all moneys' clauses under s 12CC ASIC Act/s 51AC TPA, given the expanded indicia for unconscionability now contained in those provisions.

However, the injunctions in ACCC came about as a result of court sponsored mediation between the parties. There has not been a similar success in the use of the injunctive power in actual litigated disputes. For example, in a number of test cases the regulator has sought to deal with alleged contraventions of provisions that involve major detriment to consumers by seeking resort to the injunctive remedy. In *ACCC v Oceana Commercial Pty Ltd* [151] widely framed injunctions against a bank, restraining it from lending money, except on particular terms and requiring it to provide advice whenever someone informs it that a customer has been misled as to the value of a property, were refused. Keifel J held that the bank had not acted unconscionably and therefore did not find it necessary to determine the issue of the source of the injunctive power and whether it was available to the regulator in the circumstances. However, the court did make declarations of contraventions against a number of other respondents in the case, stating:

> Declarations which simply reflect findings of contraventions are not always warranted but I accept that they may be useful in public interest litigation where they serve to vindicate legislation aimed at protection of consumers. [152]

Injunctions against a bank, in the context of misleading and deceptive conduct, were refused on discretionary grounds in *ACCC v Commonwealth Bank of Australia*[153] despite evidence of behavior by bank which contravened the Act. The court ordered instead declarations and corrective advertising as an appropriate way of addressing misleading and deceptive conduct by the bank.

It seems that the injunctive power has not been widely utilized, and is likely to play, at most, a minor role in regulating unconscionable conduct in the financial sector.

---

151 [2003] FCA 1516.

152 Ibid [343].

153 [2003] FCA 1397 [31].

Implementing Specific Unfair Contract Terms Legislation?

While the codification of different types of unconscionability concerning 'financial services' remains a useful tool for informing financiers of the manner in which their conduct may be held unconscionable, as this paper has outlined, it is not enough to regulate the use of one-sided standard form contracts, commonly used in the financial sector. The issue of unfair standard terms, particularly as used in the financial sector, needs to be addressed more directly than it is under the current unconscionability and consumer protection provisions of the ASIC Act/TPA. It also needs to be addressed in a way that facilitates action by the regulator to deal with unfair terms at a general or systemic level. As this paper has explained, existing prohibitions against unconscionable conduct in equity and in the TPA/ASIC Acts[154] do not adequately deal with ongoing concerns involving substantive unfairness in the guarantee context. Contracts of guarantee, like other types of standard form contracts used in the financial sector, are often characterised by draconian provisions, heavily weighted in favour of the financier and, particularly in the case of 'all moneys' terms, going beyond what is legitimately needed for the financier's protection. There is a strong case to be made for the implementation of unfair contracts legislation at the national level. This legislation could be seen as an important adjunct to existing unconscionability legislation, which, as the results of recent test cases show, catches some, but not all types of overreaching behaviour in the bargaining process. The provisions do not catch transactions which the courts may regard merely as a 'hard bargain' and, more significantly, they generally do not catch contracts tainted by substantive unfairness in their actual terms.

A uniform, national approach to harsh and unconscionable standard form contracts is needed. The problem of unfair terms is being more actively dealt with in other countries, which have introduced legislation specifically targeting unfair contract terms. This legislation is predicated upon the concept of 'abstract control' which, in the absence of knowledge about the precise circumstances of the transaction, generally prevents businesses from imposing 'unfair' contract terms on consumers. A useful approach to defining 'unfairness' is the European Directive on Unfair Terms in Consumer Contracts.[155] Reforms in the United Kingdom,[156] implemented in accordance with the European Directive treat a contractual term as unfair if, 'contrary to the requirement of good faith, it causes a significant imbalance the parities' rights and obligations under the contract'.[157] Examples include 'terms irrevocably binding the

---

154 Mirrored in the State/Territory Fair Trading Acts.

155 European Council Directive 93/13.

156 *Unfair Terms in Consumer Contracts Regulations* 1999 (UK).

157 Ibid reg 5 (1).

consumer to terms with which he or she had no real opportunity of becoming acquainted before the conclusion of the contract'.[158] The term will not be binding and consumer regulators or other bodies can seek, by injunctive measures, to prevent their use. The UK Law Commission noted that the *Unfair Terms in Consumer Contracts Regulations 1999* (UK) effect a form of 'abstract control'.[159] There are some specific examples of 'abstract control' in Australian legislation[160] aimed at preventing entry into contracts with unfair terms. A recent example is the *Fair Trading Act 1999*(Vic) which includes provisions directly targeted at unfair consumer contracts.[161] The Victorian legislation is largely based on the *Unfair Contract Terms Act 1977* (UK) and, more specifically, on the *Unfair Terms in Consumer Contracts Regulations 1999* (UK). These UK regulations are, in turn, modelled on the relevant European Council Directive. Section 32W provides:

> A term in a consumer contract is to be regarded as unfair if, contrary to the requirements of good faith and in all the circumstances, it causes a significant imbalance in the parties' rights and obligations arising under the contract to the detriment of the consumer.

Consumer Affairs Victoria, in its recently published guidelines on the new unfair contract terms provisions, adopts the following definition of good faith:

> A principle of fair and open dealing; that is 'playing fair', especially when one party is in a position of dominance over a consumer who is vulnerable relative to that dominance or power…Good faith is intended as a broad term. For example, it could ask that a supplier take positive steps to make sure that a contract is fair, such as bringing a term in question to the consumer's attention, rather than hiding it away in the 'small print'.[162]

---

158 Ibid Sch 2 cl 1 (i).

159 '[I]t must be the case that substantive unfairness alone can make a term unfair under [the UK Regulations]. This is because the Director General of Fair Trading and the bodies listed in Schedule 1 have the power to prevent the use of unfair terms and this may be done 'in the abstract' in the sense that the precise way in which the clause is presented to the consumer is unknown'. *Unfair Terms in Contracts* (2002) The Law Commission of England and Wales (Consultation Paper 166).

160 An example, in the context of consumer guarantees, is the legibility and print size requirements under s 162(1) of the UCCC, which came into effect in 1996. A court may on the application of the relevant State Government consumer agency, prevent the lender using a non-complying provision in any guarantee.

161 The *Fair Trading (Amendment) Act 2003* introduced Part 2B, Unfair Terms in Consumer Contracts, into the FTA.

162 Consumer Affairs Victoria, 'Preventing unfair terms in consumer contracts", Guidelines on unfair terms in consumer contracts- Amendments to the *Fair Trading Act 1999* (2003) 4 (available at http://www.consumer.vic.gov.au).

Unfair credit contracts and guarantees are not covered by the Victorian fair trading legislation, presumably on the basis that they are covered by the UCCC.[163]However, as has already been noted in this paper, the UCCC provisions on unjust contracts deal with procedural deficiencies, they do not operate to set aside a guarantee that is substantively unfair. Significantly, in the guarantee context, the UCCC does not catch guarantees given to support credit contracts for business borrowing.

The issue of uniform unfair contract terms has recently come under the scrutiny of the Standing Committee of Officials of Consumer Affairs ('SCOCA') national working party, which released its Discussion Paper on February 1, 2004.[164] The paper noted that in recent times it is the standard form contract which has become the focus of allegations of unfairness[165]. The Discussion Paper referred to overseas statistics indicating that the financial services sector, along with real estate, was the industry sector in which unfair contractual terms were most commonly encountered.[166] The Discussion Paper highlighted the limitations of the common law and existing statutory provisions in providing systemic regulation of unfair contractual terms and considered five models[167] which could be implemented, providing a preliminary analysis of the costs and benefits of the various options.[168] The working party considered the UK model in detail, and is presently seeking responses to its options. Consumer bodies have responded favourably to the SCOCA's proposal for a national approach to regulating unfair contract terms.[169]

It remains to be seen whether the Commonwealth Government will be persuaded to implement unfair contract terms legislation. While the State and Territory governments have demonstrated their commitment to work towards a national regulatory response,[170] it is not clear whether the Commonwealth Government supports the case for unfair contracts terms legislation.[171] The financial services

---

163 Ibid 30.

164 Discussion Paper, *Unfair Contract Terms*, January 2004.

165 Ibid 16-17.

166 Ibid, Annexure III.

167 Ibid, 40-55.

168 Ibid.

169 See above n 32.

170 Joint Communique, Ministerial Council on Consumer Affairs, Friday 2 September 2005 (available at www.consumer.gov.au).

171 Chris Pearce, Parliamentary Secretary to the Treasurer has indicated that the federal government has no plans to initiate any legislation on unfair terms and that 'the case for intervening in the market to regulate unfair contract terms has not been made', Financial Review, 20 May 2005

regulator, ASIC points out that a new unfair contract terms regime would impose additional regulatory costs, but does not currently have a view on whether the reforms are needed.[172]

This article has demonstrated that concerns about unfair or unconscionable guarantees, to the extent that they go to the inherent or substantive unfairness of the transaction, cannot be adequately addressed by the financial services unconscionability provisions now contained in the ASIC Act. To some extent the problem is one that centres on concerns about 'ethical lending'. Corporate social responsibility is increasingly seen as a concept that should be embraced and promoted by financiers, particularly in regard to their lending policies. While banks which adhere to the Code of Banking Practice have to some extent accepted this responsibility[173] there is a case for arguing that such values need the imprimatur of legislation which covers all lending transactions, not just those in the banking sector. Relying on voluntary codes to target unfair practices in the taking of guarantees will ultimately prove ineffective. It is to be hoped that unfair contract terms legislation can be implemented, if not at a commonwealth level, then at least by an appropriate model, perhaps similar to the UCCC model, binding all states and territories to a systemic approach to eliminating substantively unfair terms in contracts.

The Ministerial Council on Consumer Affairs has agreed to progress a national regulatory response to unfair contract terms as a matter of urgency, following extensive consultation by the national working party with consumers and business and successful implementation of unfair contract terms legislation in Victoria.[174] The Council has noted that the preferred option of the working party for regulation is nationally consistent state and territory legislation in line with unfair contract terms provisions in the Victorian *Fair Trading Act*.

---

172 ASIC Submission, March 2004, to SCOCA Working Party on Unfair Contract Terms, above n 18.

173 Articulated in Clause 25.1 of the Code, discussed above.

174 Joint Communiqué, Ministerial Council on Consumer Affairs Meeting, Friday 22 April 2005 (http://www.consumer.gov.au/html/joint_communique/jointcommunique_April 2005.htm).

THE CORRECTIVE SERVICES ACT 2006 (QLD): AN EROSION OF PRISONERS' HUMAN RIGHTS

Tamara Walsh[*]

Introduction

In 2004, a range of Brisbane-based community organisations providing services to prisoners combined to fund an external research project investigating prison release practice and policy. It was intended that the project would be the first wide-scale independent investigation of corrective services law, policy and practice since that conducted by Kennedy in 1988.[1]

The project, undertaken by myself, experienced a number of obstacles early on. Most importantly, access to prisoners and corrective services staff for interviews was denied by the Department of Corrective Services, under section 100 of the *Corrective Services Act 2000* (Qld) which stated that it is an offence to interview a prisoner[2] without the chief executive's permission, and the Department of Corrective Services' *Code of Ethics* which states that any public comment in relation to the Department or their work can only be made after an 'authorised person has given official permission'.[3] As a result of these external constraints, the project was restricted to an analysis of relevant (and publicly available) statistics; legislation, policies and procedures documents; reported judicial review decisions; and oral and written evidence from a total of 20 former prisoners and 18 prisoner service providers. While the number of respondents to the research was relatively low, the scope of consultation conducted compared well with

---

[*] LLB, BSW (Hons1), PhD. Lecturer and fellow of the Centre for Public, International and Comparative Law, TC Beirne School of Law, University of Queensland.

1 J.J. Kennedy, *Commission of Review into Corrective Services in Queensland: Interim Report*, 1988 and J.J. Kennedy, *Commission of Review into Corrective Services in Queensland: Final Report*, 1988.

2 'Prisoner' was defined under s153 as a person in secure or open custody, and a person subject to any post-prison community-based release order. This expansive definition of prisoner is retained in the new *Corrective Services Act 2006* (Qld) s214.

3 Queensland Department of Corrective Services, *Code of Ethics*, 2000.

that of the Department in its own review of the *Corrective Services Act 2000* (Qld).[4] Further, the data obtained from respondents was confirmed by the statistics and the reported case law. The research culminated in the *Incorrections* Report.[5]

In *Incorrections*, I argued that while corrective services law and policy in Queensland seemed to adopt a best practice approach to prisoner rehabilitation, in practice, the reality fell far short of the rhetoric. For example, while a best practice system of gradual release was established under the Act, very few prisoners were actually being released prior to the expiration of the full-term of their sentence, so most prisoners were released into the community without being subject to any kind of supervision.[6] And while the Act provided for the education and employment of prisoners for their rehabilitation, education and employment rates amongst prisoners fell well short of national averages.[7]

Since then, the new *Corrective Services Act 2006* (Qld) has been introduced and passed. Alarmingly, this Act eliminates many of the aspects of the old Act that were worthy of praise. In particular, the new Act erodes prisoners' access to gradual release; blatantly states that prisoners with mental illness or psychological impairment are to be treated in the same way as prisoners subject to disciplinary procedures; and practically eliminates the only existing means by which prisoners may communicate confidentially with lawyers and complaint-handling bodies.

Each of these changes, and many others besides, breaches a number of United Nations treaties, including the *International Covenant on Civil and Political Rights* (1966)[8] and the *Standard Minimum Rules on the Treatment of Prisoners* (1955).[9] They also directly contravene the *Standard Guidelines for Corrections in Australia*, a set of aspirational targets based on the *Standard Minimum Rules on the Treatment of Prisoners*, agreed to by

---

4 The Department consulted only 42 non-government stakeholders in its legislative review; see Dominic Katter, *Review of the Corrective Services Act 2000*, 2005 and Tamara Walsh, *Incorrections II: Correcting Government,* 2005.

5 Tamara Walsh, *Incorrections: Investigating Prison Release Practice and Policy in Queensland and its Impact on Community Safety*, Queensland University of Technology, Brisbane, 2004. See also Tamara Walsh, 'Is corrections correcting? An examination of prisoner rehabilitation policy and practice in Queensland' (2006) 39(1) *Australian and New Zealand Journal of Criminology* 109.

6 *Incorrections* at 76-79, 102-109.

7 *Incorrections* at 82-84, 116-120. Indeed, prisoner education rates in Queensland are the lowest in Australia; see Productivity Commission, *Report on Government Services*, 2006 at 7.17.

8 UN Doc A/6316 (1966), entered into force for Australia on 10 March 1976.

9 UN Doc E/5988 (1977).

corrective services in all Australian States and Territories.[10] The changes brought about by the new Act represent another move away from best practice and evidence-based approaches to policy and law reform. This article will examine the 'reforms' in light of human rights instruments, and will then address the question of whether any recourse is available to prisoners who fall foul of the new provisions.

Human rights implications of the new *Corrective Services Act 2006* (Qld)

It is well-established that the incarceration of prisoners does not remove them from the protection of the law;[11] it has long been acknowledged that prisoners retain all civil rights that are not expressly or impliedly denied them.[12] Indeed, in recognition of this, the Victorian and Tasmanian *Corrections Acts* contain statutory charters of prisoners' 'rights'.[13] It is reasonable, therefore, to refer to international human rights instruments in determining the appropriateness of laws regulating the treatment of prisoners. Prisoners are, after all, human, regardless of what governments and the media would have us believe.[14] Indeed, the reality is that the majority of prisoners are amongst those most disadvantaged in our society: homeless, reliant on social security benefits

---

10 Conference of Correctional Administrators, *Standard Guidelines for Corrections in Australia*, Revised Edition, 2004 available at http://www.aic.gov.au/research/corrections/standards/aust-stand.html.

11 Richard Edney, 'Judicial deference to the expertise of correctional administrators: The implications for prisoners' rights' [2001] *Australian Journal of Human Rights* 5; Matthew Groves, 'Administrative segregation of prisoners: Powers, principles of review and remedies' (1995-1996) 20 *Melbourne University Law Review* 639 at 640.

12 *Raymond v Honey* [1983] 1 AC 1 at 10, per Lord Wilberforce. For examples of rights necessarily denied to Australian prisoners, see Richard Edney, 'Judicial deference to the expertise of correctional administrators: The implications for prisoners' rights' [2001] *Australian Journal of Human Rights* 5 at footnote 27.

13 See *Corrections Act 1986* (Vic) and *Corrections Act 1997* (Tas). Having said this, some of the shortcomings of these charters are recognized in Matthew Groves, 'International law and Australian prisoners' (2001) 24(1) *University of New South Wales Law Journal* 17 at 21-23.

14 I am thinking particularly of comments such as that made by the Minister for Corrective Services, The Hon Judy Spence MP, in the Second Research Speech of the Corrective Services Bill 2006:
'The bill gets tough on prisoners and makes it clear that prisoners do not have the same rights to access things the way law-abiding community members do.'
And in the related media release:
'This new Bill makes it clear that going to jail means losing the rights that other law abiding people – the majority of Queenslanders – take for granted.'

for survival, undereducated or illiterate, underemployed, suffering from mental illness, and/or indigenous.[15] They are, therefore, most in need of the law's protection.

Principles and purposes underlying the Act

The logical place to start in an examination of the new Act is to note that it contains a number of changes in terminology. The Department of Corrective Services will now be known as 'Queensland Corrective Services';[16] the community corrections boards' will be renamed 'parole boards';[17] and the term 'strip searches' has been removed and replaced by the descriptor 'searches requiring the removal of clothing'.[18]

While changes to the way something is labelled may seem to effectuate no legal change on their face, it is clear that the Department is attempting to influence perceptions regarding these things, although the 'reforms' do not seem to be accompanied by a genuine revision, let alone improvement, of ideology.

The removal of the word 'Department' from the department's name follows the current trend in Queensland to de-departmentalise government departments; others include 'Education Queensland' and 'Queensland Health'. It seems that the motive behind such name changes is to encourage a perception of independence of the public service, and achieve a distancing of their activities from the political realm. Put blatantly, it seems that Ministers may be seeking to absolve themselves from responsibility for their portfolios, and to shield themselves from the foibles and failures that occur within their departments. Obviously, this is a direct contravention of the idea that 'ministerial responsibility' forms the backbone of responsible government (something we are trying to convince our law students still exists).

Similarly, the removal of the term 'strip search' from the Act is seemingly an attempt to dissociate these kinds of searches from the negative publicity they have attracted in recent years. It has now been established that strip searches are over-used in

---

15 See for example Eileen Baldry, 'Ex-prisoners, homelessness and the State in Australia' (2006) 39(1) *Australian and New Zealand Journal of Criminology* 20; Stuart Ross, *Bridging the Gap: A Release Transition Support Program for Victorian Prisoners: Final Evaluation Report*, 2003; Maria Boryzycki and Eileen Baldry, 'Promoting integration: The provision of prisoner post-release services' (2003) 262 *Trends and Issues in Crime and Criminal Justice*; Social Exclusion Unit (UK), *Reducing Re-Offending by Ex-Prisoners*, 2002.

16 See The Hon Judy Spence MP, 'New Corrective Services Bill introduced to Parliament', Media Statement, 28 March 2006.

17 *Corrective Services Act 2006* (Qld) s216.

18 *Corrective Services Act 2006* (Qld) ss35-38.

Queensland, or used in situations in which they are not reasonably necessary (particularly in women's prisons),[19] and it has been alleged that they are frequently executed in an unlawful manner.[20] Strip searches represent a direct contravention of the UN *Standard Minimum Rules on the Treatment of Prisoners*, which states in article 65 that prisoners should be treated in such a way that will encourage their self-respect. Their excessive use also breaches the prohibition against cruel, inhuman and degrading treatment enshrined in article 7 of the *International Covenant on Civil and Political Rights* and article 31 of the *Standard Minimum Rules on the Treatment of Prisoners*.

Having said this, one substantive change has been made to the manner in which strip searches are executed. The new Act adds a requirement that prisoners be given the opportunity to remain partly clothed during a strip search, including for example, allowing the prisoner to dress their upper body before being required to remove lower clothing.[21] This is consistent with the *Standard Guidelines for Corrections in Australia* which states in article 1.51 that searches should be conducted in a manner which ensures the 'dignity and privacy of the person being searched, as far as is practicable'. It is unfortunate that corrective services officers must be directed to conduct a search in this way, and cannot be trusted to respect prisoners' dignity and privacy in this manner without legislative direction.

The reversion to 'parole boards' from 'community corrections boards', a change introduced by the *Corrective Services Act 1988* (Qld) following the Kennedy review,[22] essentially represents an admission, and a consistent formal legal and policy shift, to the effect that community corrections (in the multifarious sense of the term) no longer exists as a means of post-prison supervision; parole is now the sole order available for

---

19 Queensland Anti-Discrimination Commission, *Women in Prison Report*, 2006 at 72-76.

20 See Debbie Kilroy, *Submission of Sisters Inside to the Anti-Discrimination Commissioner for the Inquiry into the Discrimination on the Basis of Sex, Race and Disability Experienced by Women Prisoners in Queensland,* 2004 at 25-27 and *Incorrections* at 99.

21 See *Corrective Services Act 2006* (Qld) s38(5).

22 The first, and last, major external review of corrections in Queensland was undertaken by J.J. Kennedy in 1988. Kennedy described the system as completely ineffectual in its primary aim of rehabilitating prisoners, and noted that as long as this was the case, community safety would not be ensured. Kennedy proposed a range of reforms to the system, with an emphasis in 'correction', including radical new release practices, and wide-spread use of diversion; see J.J. Kennedy, *Commission of Review into Corrective Services in Queensland: Interim Report*, 1988 and J.J. Kennedy, *Commission of Review into Corrective Services in Queensland: Final Report*, 1988.

prisoners upon their release.[23] The associated erosion of gradual release, and its impacts on the lives of prisoners, are discussed below.

Of further concern is the fact that in section 3 of the new Act, the 'Purposes' section, the commitment to addressing the 'culturally specific needs of Aboriginal and Torres Strait Islander offenders' that existed in the old section, has been removed. While a recognition of the need to take offenders' 'cultural background' into account has been retained, the removal of the explicit acknowledgement of the special circumstances of indigenous prisoners represents a blatant disregard of Guiding Principle 7 of the *Standard Guidelines for Corrections in Australia* which states that the 'design and management' of corrective services should 'take account of particular needs and disadvantages that may be faced by indigenous people'.

The removal of this 'purpose' from the Act is also a significant diversion from the *Aboriginal and Torres Strait Islander Justice Agreement* which was signed by all key Queensland Ministers (many of whom still are still members of Cabinet) in 2000.[24] The Agreement represented a formal commitment by the Queensland Government to acknowledge the impact that past policies and practices have had on indigenous people; to recognise the social values and cultural practices of indigenous people in the context of the criminal law; and ultimately, to reduce the rate of indigenous people coming into contact with the criminal justice system. By formally removing the protection of indigenous prisoners as a purpose of the Act, the Queensland Government is taking a step back from this Agreement.

Safety orders

In addition to corrective services' retreat from the recognition of the special needs of indigenous prisoners, the needs of prisoners with mental illness are also compromised under the new Act.

The *Corrective Services Act 2006* (Qld) introduces a new process by which prisoners with mental illness are managed. Previously, two orders were available to prisoners in need of 'protection': a special treatment order and a crisis support order. Special treatment orders were available to protect a prisoner or to ensure the security or good order of a prison for a seven day period (unless the chief executive authorised a longer period).[25] A special treatment order usually resulted in the prisoner being placed in an

---

23 *Corrective Services Act 2006* (Qld) s179.

24 Queensland Government, *Aboriginal and Torres Strait Islander Justice Agreement*, 2000, available at http://www.datsip.qld.gov.au/pdf/justice.pdf.

25 See particularly *Corrective Services Act 2000* (Qld) s38.

observation cell, a barren 'rubber room' with 24 hour lighting.[26] Crisis support orders were available in situations where a prisoner was considered at risk of harming themselves or another. Temporary orders (for up to five days) or longer orders (for up to three months, on the advice of a doctor or psychologist) could be made, and prisoners subject to such orders were generally housed in a 'crisis support unit', a specialised unit supposedly aimed at providing prisoners with mental illness with treatment and support.[27]

In *Incorrections*, I acknowledged that the laws related to the treatment of persons with mental illness were consistent with best practice on their face. However, I noted that, in practice, observation cells and crisis support units did not provide a therapeutic environment for prisoners with mental illness. Many former prisoners who participated in the *Incorrections* research reported that being placed in an observation cell for exhibiting symptoms of psychological distress had a detrimental effect on prisoners with mental illness. One former prisoner said:

> I would wake up, all hours of the morning, thinking it was sunny, and I'd look out my window and it was still pitch black. It plays games with you, and it sends you mad. I heard some women kicking and punching the doors…[28]

Further, service providers reported that the crisis support units were often used as a 'behavioural management tool'.[29] It seemed, therefore, that these 'specialised units' were being used to deal not only with prisoners with mental illness, but also prisoners posing a disciplinary risk, despite the contrary impression given by the Act and corrective services procedures.[30]

The response of corrective services has been to replace these two orders with a single 'safety order'.[31] Under the new Act, safety orders can be issued both to remove a prisoner with mental illness from the general prison population, and to deal with a

---

26 Prisoners' and service providers' reflections on the use of observation cells are recorded in *Incorrections* at 121-122.

27 Prisoners' and service providers' reflections on crisis support units are recorded in *Incorrections* at 120.

28 See *Incorrections* at 121.

29 *Incorrections* at 120

30 *Incorrections* at 120-122; see also Queensland Department of Corrective Services, *Procedures – Offender Management – Crisis Support Order*, 6 October 2004.

31 *Corrective Services Act 2006* (Qld) ss53-59.

prisoner who poses a risk to the good order of the prison,[32] thereby removing any pretence of differential treatment for these prisoners.

As a result, there is now no explicit recognition of the special needs of prisoners with mental illness, as compared with prisoners who are subject to disciplinary action, in the Act. This is in direct contravention of article 22 of the UN *Standard Minimum Rules on the Treatment of Prisoners* which states that sick prisoners who require specialised treatment should be transferred to specialised institutions with appropriately qualified staff, or otherwise, to civil hospitals. The new arrangement also breaches the *Standard Guidelines for Corrections in Australia* which states that prisoners with mental impairment should be individually managed in specialist health care facilities so that their special needs may be addressed.[33]

Children accommodated with their mothers in prison

The new Act does little to address the special needs of mothers and children in prison. It does not alter the current regime related to children being accommodated with their mothers in prison; that is, children less than school age are permitted to reside with their mothers, provided the chief executive decides this is suitable. Contrary to best practice[34] and article 2.55 of the *Standard Guidelines for Corrections in Australia*, these mothers and children are not accommodated in special units, or in domestic settings; rather, they are housed in larger than usual cells, often in high security facilities.[35]

However, one change regarding the treatment of these children is effectuated by the new Act. In *Incorrections*, I reported that children accommodated with their mothers in prison were being subjected to strip searches, and I called on the Department to investigate this as a matter of urgency.[36] The new Act goes some way towards addressing this concern by adding a section that states that children accommodated

---

32 *Corrective Services Act 2006* (Qld) s53.

33 See arts 1.40, 2.18, 2.36 and 2.37 and the 'Guiding Principles'.

34 See particularly Madeline Loy, 'A study of the mothers and children's program in the New South Wales Department of Corrective Services, paper presented at the conference *Women in Corrections: Staff and Clients*, 2000; Cleo Lynch, 'The Parramatta Transitional Centre integrating female inmates into the community before release', paper presented at the conference *Women in Corrections: Staff and Clients*, 2000.

35 *Incorrections* at 126.

36 *Incorrections* at 126. I was not the first person to raise this as a concern. Sisters Inside had raised this many times prior to the *Incorrections Report*.

with their mothers in prison may be subjected to a general search[37] or scanning search,[38] but they cannot be required to submit to a personal search[39] or a search requiring the removal of clothing. Again, it is unfortunate that this must be legislated and corrective services officers cannot be relied upon not to treat these children as 'little prisoners'.[40]

Privileged mail

It is generally agreed that prisoners should be able to communicate privately with certain individuals and institutions, particularly their legal representatives and those with whom they may wish to register an official complaint.[41] Article 36(3) of the *Standard Minimum Rules for the Treatment of Prisoners* states that 'every prisoner shall be allowed to make a request or complaint, without censorship as to substance' to the proper authorities. Further, article 6.9 of the *Standard Guidelines for Corrections in Australia* states that offenders should be able to lodge complaints, express concerns and seek redress without fear of retribution.

To this end, the old *Corrective Services Act 2000* (Qld) stated that 'privileged' mail could not be opened, searched or read by corrective services officials.[42] Privileged mail was defined in the regulations as mail sent to certain approved persons, including a prisoner's legal representative, the Minister, the Ombudsman, an official visitor, and the Human Rights and Equal Opportunity Commission.[43]

The new Act removes this protection of prisoners' privileged mail. Section 45(2)(b) of the *Corrective Services Act 2006* (Qld) states that 'privileged' mail may be opened and

---

37 That is, where a person is required to empty their pockets, remove outer garments, and display the contents of any luggage.

38 That is, where a person is scanned, but not touched, by an electronic device or a 'sniffer dog'.

39 That is, where light pressure is applied to the person's inner garments, but without contact with the genital, anus or breast areas.

40 In the words of the Minister; see The Hon Judy Spence MP, Second Reading Speech of the *Corrective Services Bill 2006* (Qld) in Queensland Parliament, *Hansard*, 29 March 2006 at 943.

41 Indeed, the United States Supreme Court has held that all prisoner mail should remain unsearched and uncensored, but rather should remain 'an open and substantially unimpeded channel for communication'; *Procunier v Martinez* 416 US 396 (1974) at 417 and *Pell v Procunier* 417 US 817 (1974) at 824.

42 *Corrective Services Act 2000* (Qld) s35(1).

43 See *Corrective Services Regulations 2001* (Qld) s7.

searched if a corrective services officer reasonably suspects the mail is not privileged mail. The officer is directed in section 45(3) not to read the mail, but section 45(4) goes on to say that in the event that they do, they are not to disclose its contents to another.

This change will mean that prisoners have no means of communicating privately with any member of the outside world (since prisoners' telephone calls are recorded and monitored[44]) and it may ultimately mean that prisoners subject to serious mistreatment will not lodge complaints for fear of the retributive consequences that may result. This is a serious breach of prisoners' human rights, particularly their right to be recognised as a person before the law.[45]

'Obstructionist' behaviour

In addition to all of this, behaviour considered 'obstructionist' by the Department is more tightly regulated under the new Act; both that of prisoners and of prisoner service providers.

The offence of 'obstruct corrective service officer' existed under the old Act,[46] however in the *Corrective Services Act 2006* (Qld), the maximum penalty has been increased from 40 penalty units (a $3000 fine) or one year's imprisonment, to two year's imprisonment without the express option of a fine.[47] Further, the defence of reasonable excuse that was available under the old Act has been removed.

The new Act also seeks more closely to regulate the activities of service providers funded by the Department. A new Chapter has been included in the new Act on the subject of service provider agreements.[48] Under these new provisions, service providers who receive funding from corrective services must enter into a 'financial assistance agreement'.[49] Far from being an 'agreement', this document is unilateral in nature; in it, corrective services mandates the terms of funding, outlines the circumstances in which the service provider will be considered in breach of the agreement, and proscribes the action that may be taken in the event of such a breach

---

44 Except for communications authorised by the chief executive between the prisoner and his/her lawyer, a law enforcement agency, the parole board or the ombudsman; *Corrective Services Act 2006* (Qld) s52.

45 Article 16 of the *International Covenant on Civil and Political Rights*.

46 *Corrective Services Act 2000* (Qld) s95.

47 *Corrective Services Act 2006* (Qld) s124(b).

48 *Corrective Services Act 2006* (Qld) Chapter 6.

49 *Corrective Services Act 2006* (Qld) s252.

(including de-funding).[50] Considering the strong advocacy role that many prisoner service providers undertake in Queensland, there is legitimate concern amongst stakeholders that this Chapter may be used to silence these providers, under the threat of de-funding. Indeed, some have noted that this simply formalises the current situation; Sisters Inside[51] has been prevented from entering Queensland prisons to provide support and counselling to incarcerated women since its lobbying activities culminated in the launch of an Anti-Discrimination Commission Queensland investigation in 2004.[52]

Gradual release

Perhaps the most worrying change in the new Act relates to the erosion of gradual release. In *Incorrections*, I positively acknowledged the Department's formal commitment to gradual release. The *Corrective Services Act 2000* (Qld) provided a range of 'post-prison community-based release' options, including release to work,[53] home detention and parole.[54] The Act set up a scheme whereby those sentenced to imprisonment for more than two years could apply for a post-prison community-based release order which could be tailored to the reintegration needs of the particular prisoner by combining two or more of the available release options. While the empirical research reported on in *Incorrections* established that these release options were rarely used, and prisoners were invariably released without supervision upon the expiry of their full term,[55] the capacity for something better was enshrined in the legislation.

The new Act abolishes two of these gradual release options; parole is now the only post-prison community-based release order available.[56] A program that gradually

---

50 *Corrective Services Act 2006* (Qld) s235(1)(k), (l).

51 Sisters Inside Inc. is a world renowned independent community organisation which exists to advocate for the human rights of women in the criminal justice system, and to address gaps in the services available to them; see www.sistersinside.com.au.

52 Anti-Discrimination Commission Queensland, above n19.

53 Where a prisoner undertakes work in the community under community corrections surveillance, which at a minimum included one phone check per week and one physical check per week; see Queensland Department of Corrective Services, *Procedures – Offender Management – Post-Prison Community-Based Release Orders*, 1 July 2001.

54 *Corrective Services Act 2000* (Qld) s141.

55 *Incorrections* at 102.

56 *Corrective Services Act 2006* (Qld) Chapter 5, Part 1.

releases a prisoner into a less restrictive environment over time consistent with their supervision requirements (for example, by moving from release to work, to home detention, to parole) is no longer possible. Further, parole is only available to those serving sentences of more than three years, unless the sentencing court explicitly set a parole period.[57] Maintaining the integrity of the court's orders will go some way towards addressing the frustrations that many prisoners have experienced when they have reached their parole eligibility date and yet been denied parole.[58] However, since the vast majority of prisoners serve sentences of less than three years,[59] it is likely that only a very small proportion of prisoners will be eligible for any form of supervised gradual release.

This is in direct contravention of the *Standard Guidelines for Corrections in Australia* which states at Guiding Principle 3 that corrective services should 'maximis[e] opportunities for community-based rehabilitation and integration of offenders'. It is also in breach of article 60(2) of the *Standard Minimum Rules for the Treatment of Prisoners* which states that 'necessary steps' should be taken to ensure a 'gradual return to life in society' for prisoners under appropriate supervision. Article 61 goes on to state that the treatment of prisoners should emphasise not their exclusion from the community, but their continuing part in it.

The erosion of gradual release is clearly contrary to best practice principles, and represents an obvious departure from an evidence-based approach. It is well-established that prisoner rehabilitation, and community safety, are best ensured by the release of prisoners over time to less and less restrictive environments, based on a rehabilitative release scheme, rather than being released absolutely at the expiration of their full-term, so that they may be progressively prepared for community life.[60]

---

57 *Corrective Services Act 2006* (Qld) s179.

58 See *Incorrections* where one former prisoner is reported to have said (at 105): 'Prisoners arrive in prison with an expectation that they will only be in custody until their parole eligibility date, and when they understand the truth of such matters, it seriously impedes their attitude towards rehabilitation'.

59 The Australian Bureau of Statistics reported in *Prisoners in Australia* (2005) that on 30 June 2005, 56% of prisoners in Queensland were serving a sentence of less than five years, and 38% of prisoners were serving a sentence of less than two years. Taking flow-through into account, it must be concluded that very few prisoners will be eligible for non-court-ordered parole.

60 See for example Lisa Ward, *Transition from Custody to Community: Transitional Support for People Leaving Prison*, 2001; Gary Hill, 'The Correctional System of Greece – The Prison System and Aftercare' (2000) July *Corrections Compendium* 20.

Classification

Intimately related to the issue of gradual release is prisoner classification. Best practice suggests that prisoners should be accommodated in facilities that provide an appropriate level of supervision relative to their security and escape risk.[61] Under the *Corrective Services Act 2000* (Qld), prisoners were classified as maximum, high, medium, low or open security;[62] the idea being that prisoners would, during the course of their sentence, aim to progress through the various classifications by demonstrating 'good conduct and industry'. In *Incorrections*, I noted that one problem with the current system was that, due to the paucity of beds in lower security facilities, and the fact that they are never filled to capacity, many prisoners classified as medium or low were housed in high security facilities.[63] Thus, the achievement of a lower security classification[64] was not being matched by any attendant benefits.

The response in the new Act has been to reduce the number of classification categories to three – maximum, high and low – apparently to ensure a closer match between prisoners' classification and the facility in which they are housed.[65] The consequence of this is that those previously classified as medium and low will now be classified as high security, and thus will continue (legitimately) to be held in high security facilities.[66]

This arrangement is in direct contravention of article 1.37 of the *Standard Guidelines for Corrections in Australia* which states that prisoners should be placed at the 'lowest level of security appropriate for their circumstances'. In particular, it is well-established that female prisoners do not pose a significant threat to community safety or institutional good order, and therefore, it is generally accepted that they should be subject to a specialised classification system and should normally be housed in minimum security

---

61 See for example the 'Guiding Principles for the Management of Prisoners' in the *Standard Guidelines for Corrections in Australia*; Robert B Levinson, 'Developments in the classification process: Quay's AIMS approach' (1988) 15(1) *Criminal Justice and Behaviour* 24.

62 Those classified as low or open security were eligible to be placed in open security facilities (often farm complexes), rather than residing in secure facilities.

63 *Incorrections* at 102-103, 107-108.

64 And an achievement, indeed, it is. Respondents to the *Incorrections* research reported that prisoners found it extremely difficult to progress through the classification process – indeed, the vast majority of prisoners were released at the end of their full-term still classified as high security; *Incorrections* 102-107.

65 *Corrective Services Act 2006* (Qld) s12.

66 *Corrective Services Act 2006* (Qld) s363(3).

facilities.[67] Article 1.41 of the *Standard Guidelines for Corrections in Australia* explicitly states that the management of female prisoners 'should reflect their generally lower security needs'. Yet the majority of women in prison in Queensland will now be classified as high security and will continue to be housed in high security facilities.[68]

Further, for those prisoners who have worked hard to achieve their lower security classification, there is a significant risk that their re-classification as a high security risk will have a grave impact on their morale and their attitude towards their rehabilitation. A change such as this is contrary to best practice, which suggests that classification systems should be utilised in case management to create achievable goals that prisoners can aim towards. This is so that benefits will flow to prison managers in the form of prisoners' 'good conduct and industry'.[69] Further, the UN's *Standard Minimum Rules for the Treatment of Prisoners* recognise at article 67(b) that the purpose of classification should be to 'facilitate [prisoners'] treatment with a view to their social rehabilitation', and the *Standard Guidelines for Corrections in Australia* recognise the need for prisoner management systems to be predictable, structured and transparent.[70] The changes to prisoner classification under the new Act have the opposite effect.

Removal of the right to judicial review

Of further concern is the fact that the new Act removes prisoners' right to judicial review of classification and transfer decisions.[71] It does this by stating that Parts 3, 4 and 5 of the *Judicial Review Act 1991* (Qld) (the Parts that deal with statutory orders, statements of reasons and prerogative writs and injunctions) do not apply to decisions

---

67 James Austin and Patricia Hardyman, *Objective Prison Classification: A Guide for Correctional Agencies,* United States Department of Justice, Washington DC, 2004 pp49-57; Canadian Human Rights Commission, *Protecting Their Rights: A Systematic Review of Human Rights in Correctional Services for Federally Sentenced Women,* Canadian Human Rights Commission, Ottawa, 2003; Debbie Kilroy, *Submission of Sisters Inside to the Anti-Discrimination Commissioner for the Inquiry into the Discrimination on the Basis of Sex, Race and Disability Experienced by Women Prisoners in Queensland,* 2004.

68 As was reported in *Incorrections* (at 107), there are around 350 women in prison in Queensland, but only 62 beds available in low security facilities.

69 Austin and Hardyman, above n67 at 59-60

70 *Standard Guidelines for Corrections in Australia* (2004) art 1.37

71 *Corrective Services Act 2006* (Qld) ss17 (classification), 71 (transfers).

of this nature.[72] Indeed, the new Act goes so far as to state that review is not available even for a classification or transfer decision that was made in jurisdictional error.[73]

The effect of these provisions is that prisoners are now (on the face of the law) unable to obtain a remedy in relation to a classification or transfer decision: neither a statutory order, nor a remedy under a prerogative writ or injunction.[74] Nor are they able to obtain a statement of reasons in relation to the decision. According to the Act, their only means of redress is via an internal review procedure, whereby the chief executive (ie. the Director-General) reviews their case and makes a determination, without necessarily giving the prisoner a fair hearing.[75]

There are many reasons why a prisoner would want to seek judicial review of a classification or transfer decision. As noted above, prisoners' classification impacts heavily on the manner in which they are managed, and the kind of institution in which they are accommodated. Further, progressing through the classification system has proved so difficult (often for no legitimate or apparent reason)[76] that many prisoners have turned to the courts to act as an arbiter in situations where they believe they have been treated unfairly or subjected to bias.[77] With regard to transfers, being moved to another institution might mean that a prisoner is accommodated so far away from their family and community that visits are impossible, and phone calls are expensive.[78] It is well-established that maintaining contact with family members is

---

72 See ss 17(1) (classification), 71(4) (transfers).

73 For example, where there was no power to make the decision, or the decision-maker exceeded their powers (*Craig v South Australia* (1995) 184 CLR 163 at 179); see ss 17(2) (classification), 71(5) (transfers).

74 Section 41(1) *Judicial Review Act 1991* (Qld) states that prerogative writs are no longer to be issued by the Court.

75 See ss16 (classification), 71(1)-(3) (transfers). There is no requirement in the Act that the Director-General provide the prisoner with a reasonable opportunity to prepare and present their case, and of course there is no requirement that the prisoner be provided with an oral hearing, despite the fact that, without one, many prisoners would certainly be disadvantaged (mainly due to their lack of literacy skills); see *Chen Zhen Zi v Minister for Immigration* (1994) 121 ALR 83.

76 See *Incorrections* at 103-107.

77 See *Bartz v Department of Corrective Services* [2000[QSC 336, *Crowley v Chief Executive, Department of Corrective Services* [2001] QSC 219 and *Onea v Chief Executive, Department of Corrective Services* [2002] QSC 420 where classification/transfer decisions were set aside by the Queensland Supreme Court.

78 Prisoners are required to pay for their own phone calls; see *Corrective Services Act 2006* (Qld) s50(1)(b).

critical to a prisoner's rehabilitation, a fact which is recognised in article 1.39 of the *Standard Guidelines for Corrections in Australia* which states that prisoners should be enabled to 'reside as closely as possible to their family, significant others or community of interest'.[79] Yet transfers have the capacity to substantially limit the amount of contact a prisoner has with these people.

While few prisoners actually availed themselves of their 'right' to judicial review under the old Act[80] its availability was regarded as a stop-gap by prisoners, at least against blatant unlawful action on the part of government decision-makers.[81] Removal of the capacity to have these decisions reviewed by a judge is a contravention of articles 1.24 and 6.10 of the *Standard Guidelines for Corrections in Australia* which state that prisoners should be able to appeal to an external body in situations where their grievance is not resolved to their satisfaction, and it is a clear example of a move by corrective services away from the principles of fairness, accountability and transparency it purports to uphold.[82]

The enforceability of privative clauses: Legal avenues for redress

The restriction of prisoners' access to judicial review seems dubious, particularly to lawyers, who are professionally socialised to value the rule of law. The question then becomes whether there is any capacity for prisoners or service providers to seek legal redress against this. Two possible avenues will be discussed here: first, whether prisoners can assert a 'right' to judicial review, and second, whether the principles of natural justice might assist.

---

79 See also Gearoid O'Loingsigh, *Getting Out, Staying Out: The Experiences of Prisoners Upon Release*, 2004 at 22-23; Social Exclusion Unit (United Kingdom), *Reducing Re-Offending by Ex-Prisoners*, 2002 at 111-116.

80 See Queensland Legal, Constitutional and Administrative Review Committee, *The Accessibility of Administrative Justice: Discussion Paper*, 2005, Appendix B. See also Mark Plunkett, 'The Corrective Services Bill 2006 and judicial review of corrective service decisions', paper presented at the forum *The Crime of Punishment: No More Natural Justice for People Living in Prison*, 15 May 2006, Banco Court, Brisbane at 3.

81 This was expressed at the forum *The Crime of Punishment: No More Natural Justice for People Living in Prison* held at the Banco Court in Brisbane after the release of the *Corrective Services Bill 2006* (Qld).

82 See for example the 'Guiding Principles for the Management of Prisoners' and article 6.9 of the *Standing Guidelines for Corrections in Australia*; Queensland Department of Corrective Services, *Policy Framework*, 2006.

A 'right' to judicial review?

Access to judicial review of government decisions has become so entrenched that it is generally considered to be a 'right', protected either by statute or at common law.[83] Indeed, its genesis lay in the rule of law, and the idea that a decision that went beyond the scope of a decision-maker's power was of no legal effect because the decision-maker had failed to comply with the law.[84] However, it is well-established that a statutory or common law right can be abrogated at any time by a valid Act that is inconsistent with such a right.[85]

Privative clauses (laws which seek to remove the jurisdiction of the courts to review legislation) have received much attention at the federal level, both in the case law and the literature.[86] The conclusion of the courts at Commonwealth level has been that Parliament may be able to prevent the courts from reviewing non-jurisdictional errors (for example, where the decision-maker failed to take a relevant consideration into account, or took an irrelevant consideration into account), but it cannot prevent the courts' review of decisions in jurisdictional error (for example, where the decision-maker lacked the power to make the particular decision).[87]

In addition to this, the High Court's interpretation of section 75(v) of the Constitution has influenced its views with regard to the reviewability of administrative decisions at Commonwealth level. This section outlines the inherent jurisdiction of the High Court

---

83 John McMillan and Neil Williams, 'Administrative law and human rights' in *Human Rights in Australian Law*, Federation Press, Sydney, 1998 at 64; Mark Aronson and B Dyer, *Judicial Review of Administrative Action*, Law Book Company, Sydney, 2004, at 85-93.

84 See particularly *Church of Scientology v Woodward* (1982) 154 CLR 25 at 70. See also Bradley Selway QC, 'The principle behind common law judicial review of administrative action: The search continues', (2002) *Federal Law Review* 8.

85 See *Potter v Minahan* (1908) 7 CLR 277 at 304; *Sorby v Commonwealth* (1983) 26 ALR 237 at 241-242. Although, see AW Bradley, 'Administrative Justice: A Developing Human Right?' (1995) 1 *European Public Law* 347 for a contrary view in relation to judicial review.

86 Particularly in the context of refugee law; see *NAAV v Minister for Immigration and Multicultural Affairs and Others* (2002) 193 ALR 449 and *Plaintiff S157/2002 v Commonwealth* (2003) 211 CLR 476. For a broad discussion of privative clauses, see particularly Mark Seymour, 'Privative clauses in administrative law: Recent developments' (2003) 77(11) *Australian Law Journal* 757.

87 See *Craig v South Australia* (1995) 184 CLR 163 at 179. It should be noted, however, that more recent cases suggest the distinction between jurisdictional and non-jurisdictional error may be becoming increasingly blurred; see particularly *Re Refugee Review Tribunal; Ex parte Aala* (2000) 176 ALR 219; *Minister for Immigration and Multicultural Affairs v Yusuf* (2001) 180 ALR 1 at [82].

and sub-section (v) states that all matters in which a writ of mandamus[88] or prohibition,[89] or an injunction, is sought against an officer of the Commonwealth come within this inherent jurisdiction. On this basis, the High Court has declared that it cannot be prevented from reviewing and granting relief in respect of a decision that is beyond power.[90]

Of course, the situation at State level is very different.[91] The States do not have paramount Constitutions, so strictly speaking their Constitution Acts have no greater force than any other Act of Parliament.[92] A further impediment in Queensland is that section 41(1) of the *Judicial Review Act 1991* (Qld) explicitly states that prerogative writs are abolished, and cannot be issued by the Supreme Court of Queensland. As stated above, the provisions within the *Corrective Services Act 2006* (Qld) that remove access to judicial review ensure that neither statutory orders, nor prerogative writs, are capable of being utilised, so there appears to be no available remedy.

Yet the Supreme Court of Queensland has stated in a number of cases that it cannot be prevented from reviewing decisions made without jurisdiction, apparently adopting 'complete absence of power' reasoning; that is, if the decision is made outside the scope of the Act, the Act does not apply to it, and thus the Act cannot prevent the

---

88 That is, a prerogative order compelling a Crown official to do something.

89 That is, a prerogative order compelling a Crown official to stop doing something.

90 See particularly *Plaintiff S-157S157/2002 v Commonwealth of Australia* (2003) 195 ALR 24, Gaudron and Gummow JJ (with whom McHugh and Hayne JJ agreed) at 614 at [51] Gaudron, McHugh, Gummow, Kirby and Hayne JJ; *R vHickman; Ex parte Fox and Clinton* (1945) 70 CLR 598; *Darling Casino Ltd v New South Wales Casino Control Authority* (1998) 191 CLR 602 at 633. See also DF Jackson QC, 'Development of judicial review in Australia over the last 10 years: the growth of the constitutional writs' (2004) 12(1) *Australian Journal of Administrative Law* 22.

91 See Michael Sexton and Julia Quilter, 'Privative clauses and State constitutions' (2003) 5(4) *Constitutional Law and Policy Review* 69 at 73.

92 See Denise Meyerson 'State and federal privative clauses: not so different after all' (2005) 16(1) *Public Law Review* 39 at 50.

courts from reviewing it.[93] Indeed, such an approach seems to apply to both jurisdictional and non-jurisdictional errors.[94]

Further, there is a section of the *Constitution Act 2001* (Qld) that suggests some relief may still be available. Section 58 reads as follows:

> (1) The Supreme Court has all jurisdiction necessary for the administration of justice in Queensland.
>
> (2) Without limiting subsection (1), the court—
>
> (a) is the superior court of record in Queensland and the supreme court of general jurisdiction in and for the State; and
>
> (b) has, subject to the Commonwealth Constitution, unlimited jurisdiction at law, in equity and otherwise.

This section establishes the Supreme Court of Queensland as a superior court of record; as such, nothing can be considered outside its jurisdiction unless there is a clear statement to this effect.[95] Further, as the High Court said in *Ainsworth v Criminal Justice Commission*, 'it is now accepted that superior courts have an inherent power to grant declaratory relief' to determine legal controversies.[96] So, it seems that at the very least, declaratory relief would be available to a prisoner in respect of a decision that was beyond the scope of the decision-maker's power.

In addition to this, section 27B of the *Acts Interpretation Act 1954* (Qld) states that in situations where a government decision-maker is required to give written reasons (or grounds) for a decision, the instrument giving the reasons must set out the findings on material questions of fact and refer to evidence or other material on which those findings were based. Sections 15 and 26 of the *Corrective Services Act 2006* (Qld) state

---

93 Most recently, see *Batterham v QSR Ltd* [2006] HCA 23 particularly per Kirby at [54] and [66] at [26]. See also *Squires v President of the Industrial Court of Queensland & Ors* [2002] QSC 272 per Mullins J at [34]; *Carey v President of the Industrial Court of Queensland and Department of Justice and Attorney-General* [2004] QCA 62 per McPherson JA at [4] and [22]; *Emerald Developments Pty Ltd v Minister for Environment, Local Government, Planning and Women* [2006] QSC 073 at [39]. Further, see Plunkett above n80 at 9.

94 See particularly the recent Queensland case of *Emerald Developments Pty Ltd v Minister for Environment, Local Government, Planning and Women* [2006] QSC 073 at [39]. See also the New South Wales case of *Timbarra Protection Coalition Inc v Ross Mining NL* (1999) 46 NSWLR 55.

95 *Re Totalisator Adminiatration Board of Queensland* [1989] 1 Qd R 215; Plunkett, above n80 at 9

96 (1992) 175 CLR 564, per Mason CJ, Dawson, Toohey and Gaudron JJ at 581.

that where a prisoner seeks an internal review by the chief executive of a classification decision, the chief executive must provide the prisoner with an 'information notice about the chief executive's decision'.[97] On this basis, regardless of corrective services' explicit attempt to prevent prisoners from obtaining a statement of reasons for decisions made about their classification, prisoners would be entitled to information, including the factual bases for the decision, under the *Acts Interpretation Act 1954* (Qld).[98]

It must be concluded, therefore, that the Supreme Court of Queensland could probably still grant (at the very least) declaratory relief to prisoners in situations where the classification or transfer decision at issue was beyond power, and a prisoner could probably obtain 'reasons', in the form of factual bases for the decision, for a classification decision. The Queensland Government's attempt to prevent judicial review of transfer and classification decisions may, therefore, prove ineffective.

A 'right' to natural justice?

In addition to this, or perhaps as an aside, the internal review procedures established under the *Corrective Services Act 2006* (Qld) may not sufficiently comply with the rules of natural justice. In the Second Reading Speech of the *Corrective Services Bill 2006* (Qld), the Minister for Corrective Services stated, in the context of the removal of judicial review for certain decisions, that prisoners would continue to have 'unfettered access to complaint mechanisms such as official visitors' and internal review procedures.[99]

It is now generally accepted that natural justice requires that a person be granted an opportunity to be heard before a decision is made against them[100] and that the decision be made by a non-biased decision-maker, who has approached the task with an open mind and is free from preconceptions.[101] The applicable test for non-pecuniary bias in an administrative context is whether, in the circumstances, a

---

97 The situation as regards transfer decision is less clear, since s71 which deals with the reconsideration of such decisions does not require the chief executive to provide any written information to the prisoner regarding his/her decision. Rather, s71(3) says that after reconsidering the decision, the chief executive may 'confirm, amend or cancel' it.

98 See Plunkett, above n80at 2.

99 Queensland Parliament, *Hansard*, 29 March 2006 at 943.

100 See for example *Re Macquarie University; Ex parte Ong* (1989) 17 NSWLR 113; *Herscu v Queensland Corrective Services Commission* [1995] 2 Qd R 481.

101 *Webb v R* (1994) 181 CLR 41 at 53 (per Mason CJ and McHugh J) and 67-75 (per Deane J).

reasonable apprehension of bias might be entertained;[102] that is, the test is an objective one. In *Incorrections*, I reported on the perceptions of prisoners regarding the objectivity (or lack thereof) of decision-makers within the Queensland corrective services system. Respondents were overwhelmingly of the belief that internal review mechanisms were not free from bias. Official visitors, for example, were generally perceived to be 'part of the system',[103] a fact reflected in the significant decrease in the number of complaints dealt with by official visitors in recent years.[104] Indeed, even the independence of the Queensland Community Corrections Boards was questioned by respondents, with many indicating that Ministerial Guidelines dictate the decision-making process, even to the point of being slavishly followed.[105]

There is, therefore, a strong perception amongst prisoners and their advocates that internal review, and review by official visitors, of decisions will not be free from bias. If no other review mechanism is available to prisoners, it might be concluded that the rules of natural justice are breached by such an arrangement.

Conclusion

In article 9 of the Guiding Principles in the *Standard Guidelines for Corrections in Australia*, a commitment is made to ensure that corrections policies, programmes and services are 'informed by research, and reflect sound evidence-based practice'. Yet, in Queensland in the past two years, at least two independent reports have been released, both of which have made serious allegations of prisoner mistreatment and unlawful corrective services practices.[106] The Department of Corrective Services

---

102 *Webb v R* (1994) 181 CLR 41 at 53 (per Mason CJ and McHugh J) and 67-75 (per Deane J).

103 *Incorrections* at 107

104 There was a decrease of 19% between 2001/02 and 2002/03 in the number of complaints dealt with by official visitors; Queensland Department of Corrective Services, *Annual Report 2003/04*, 2004.

105 There has been at least one application for judicial review based on this allegation. In *Butler v Queensland Community Corrections Board* [2001] QCA 323, the applicant argued that Ministerial Guidelines were slavishly followed by the Board. The Supreme Court of Queensland dismissed the application. See *Incorrections* at 108.

106 They are *Incorrections* and the Anti-Discrimination Commission Queensland's report. In addition to this, many reports authored by service providers and advocates, including Sisters Inside and the State Incorrections Network, have been released during this time (available at www.sistersinside.com.au).

drafted a response to each of these reports,[107] yet neither of these responses conceded any of the points made. Rather, the Department's response in both cases was publicly to condemn the reports, claiming that the issues raised had already been investigated and dealt with by the Department; that many of the recommendations had already been implemented on the Department's own initiative; and even to deny that any adverse findings were made against the Department.[108] This is despite the fact that many of the recommendations made in these reports were subsequently acted upon by the Department.[109]

As further steps away from aspirational targets, in the form of the *Standard Guidelines* and United Nations treaties, are taken, Queensland prisoners are being increasingly dehumanised and silenced. In a policy environment characterised by lack of transparency and denial of accountability, the prospect for prisoners' human rights to be realised is grim indeed.

---

[107] See Queensland Department of Corrective Services, *Response to the Incorrections Report*, 2005; Queensland Department of Corrective Services, *Response to the Anti-Discrimination Commission Queensland Women in Prison Report*, 2006.

[108] Ibid. See also The Hon Judy Spence MP, 'Report shows no discrimination of women in Queensland jails' Media Statement, 8 March 2006.

[109] For example, as discussed above, *Incorrections* raised the issue of children accommodated with their mothers in prison being unlawfully searched, and this problem was addressed in the new Act.

DRUG USE AND THE DEFENCE OF MENTAL IMPAIRMENT: SOME CONCEPTUAL AND EXPLANATORY ISSUES

Steven T Yannoulidis[*]

Recent work has highlighted the problem of amphetamine use in Queensland. In particular, concern has been expressed in relation to what has been described as a 'disturbing trend' towards the provision of exculpatory defences based on findings of 'amphetamine-induced psychosis'. While the subject of a spate of cases before the then Queensland Mental Health Tribunal, instances of drug induced psychosis are not restricted to Queensland. Similar issues have arisen in the context of the Victorian Law Reform Commission's recent analysis of people with mentally impaired functioning who kill. What has become apparent in both case law and reform proposals dealing with this area is the interconnectedness of issues of principle and policy. The present article endeavours to clarify the thicket of conceptual and explanatory issues underlying voluntary drug use and the subsequent raising of a defence of non-responsibility.

Introduction

In a recent article Russ Scott and William Kingswell have drawn attention to certain problems associated with what they describe as 'amphetamine-induced psychosis'.[1] According to the authors such a condition is characterised by two distinct types: an acute onset psychotic state that resolves when the individual stops using the drug, and a chronic psychotic state occurring with a history of prolonged usage.[2] From a forensic point of view it is distinguishing between 'intoxication', 'psychosis' and the

---

[*] PhD Researcher, Monash University, Clayton, Victoria.

The author gratefully acknowledges the assistance received from Dr Andy Carroll, Dr Debra Wood and Professor Bernadette McSherry, all of whom were germane in the elaboration of the ideas herein expressed. Faults and omissions remain the author's. The author would also like to acknowledge the support received from the Monash University Postgraduate Publication Award Scheme in the writing of this article.

1 R Scott and W Kingswell, 'Amphetamines, Psychosis and the Insanity Defence: Disturbing Trends in Queensland' (2003) 23 *Queensland Lawyer* 151.

2 Ibid 4.

legal definition of 'mental disease' that is most problematic. Of particular concern is the question of whether, given a psychobiological point at which a mental disorder exists independently of the consumption of a drug, this should give rise to a defence of mental impairment?[3]

The growing use of amphetamines Australia-wide cannot be gainsaid. Figures released by the Australian Institute of Criminology last year indicate that an overview of arrest patterns for offenders between 1995-6 and 2003-4 reveal that 'arrests for amphetamines have almost doubled'.[4] The difficulties faced by forensic psychiatrists, lawyers and policy makers in the face of such an ever-growing problem have been the subject of a recent paper.[5] Rather than rehearse these broad concerns the present paper endeavours to explore some of the underlying conceptual issues faced by legal theorists in responding to the voluntary consumption of psychoactive drugs.

As a starting point it is important to distinguish between two distinct types of question. On the one hand is the question of whether a diagnosable mental disorder exists independently of the consumption of the drugs. Such a question is clearly a matter for expert evidence. On the other hand, is the question whether a defence of mental impairment should be available to an accused who has voluntarily chosen to consume drugs? In particular, should such an individual's mental state at the time of the offence be taken into account? These two latter questions form the focal point of the ensuing discussion. Indeed it is a principal contention of this paper that both of these last two questions raise normative issues falling outside the purview of contemporary clinical research.

In order to answer these questions this paper will critically examine three approaches that have been adopted in assigning criminal responsibility in situations of drug related states of mental impairment.[6] The first approach is one espoused by the Victorian Law Reform Commission during its examination of homicide offences committed by people with mentally impaired functioning.[7] Succinctly stated, the Commission adopted the view that the causal antecedents of an accused's behaviour are extraneous to the question of his or her culpability. According to the Commission

---

3 'Mental Impairment' is the terminology employed in the Victorian statute *Crimes (Mental Impairment and Unfitness to be Tried) Act* 1997, s 20, replacing the common law defence of insanity.

4 Australian Institute of Criminology, *Australian Crime: Facts and Figures* (2006) 42.

5 A Carroll, B McSherry, D Wood and S Yannoulidis, 'Drug-Associated Psychoses and Criminal Responsibility' in *Psychology, Public Policy and Law* [in press].

6 See ibid.

7 Victorian Law Reform Commission, *Defences to Homicide: Final Report* (2004) 203-52.

if there is a cognitive failure on the part of an accused non-responsibility should follow as a matter of principle.

The VLRC view that cause is irrelevant to the question of an accused's culpability is diametrically opposed to the second approach adopted in such cases. This is the view put forward by Edward Mitchell in the course of his discussion of the notion of 'meta-responsibility'.[8] Focusing on prior conduct Mitchell would hold an accused responsible where he or she has engaged in prior culpable behaviour that in turn constitutes the conditions of the defence relied upon. Mitchell's views will be examined in Part IV below.

The third approach I will consider is one that has found favour with some American theorists who have called for a test of 'settled insanity'.[9] This latter approach is predicated on a finding of a state of mental impairment subsequent to drug use. Such theorists argue that such a state may be ascertained either through evidence of a fixed, stable state independent of drug use, or upon a demonstrated predisposition to psychosis. A consideration of both these views will form the substance of Part V.

Prior to considering the relative merits of each of these views by way of introduction I propose to outline certain general principles of criminal responsibility. Such principles, while necessarily contentious at points, will nevertheless facilitate the ensuing discussion by providing a normative background. The application of these principles to the highly specific defence of mental impairment will in turn form the final part of the expository section. Concluding remarks will endeavour to provide an overview of the preceding analysis and in so doing place in better focus the true nature of the enquiry.

General Principles

(a) Criminal Responsibility: Criminal responsibility requires an account of the conditions that lead to the attribution of fault for particular actions by an individual.

---

8 Edward Mitchell, 'Meta-Responsibility and mental disorder: causing the conditions of one's own insanity plea' (1999) 10 *Journal of Forensic Psychiatry* 597; *Self-Made Madness: Re-thinking Illness and Criminal Responsibility* (2004); 'Culpability for inducing mental states: The insanity defence of Dr Jekyll' (2004) 32(1) *The Journal of The American Academy of Psychiatry and the Law* 63.

9 J Reid Meloy, 'Voluntary Intoxication and the insanity defence' (1992) Spring *The Journal of Psychiatry and Law* 439.

Two such accounts are termed 'subjective' and 'objective'.[10] These accounts are distinguished on the basis of the criteria each employs in ascribing proscribed acts to a particular actor. 'Subjectivist' accounts seek to utilise subjective terms in such ascriptions while 'objectivist' accounts employ terms of an objective character.[11]

Further, subjectivist accounts may be divided into analyses involving a focus on either 'choice' or 'character'. A choice analysis highlights actions which an agent has chosen as being truly his or her own. In contrast, a character analysis seeks to underpin ascriptions of action to an individual on the basis of his or her character traits. Such traits, according to a character theorist, are revealed in his or her dispositions, attitudes or motivations.[12] Yet, it has been maintained that such a focus on character misses the true focus of liability, namely, action. This is readily apparent from the fact that proscribed action extends to behaviour that may be viewed as 'out of character'. As Antony Duff has noted, criminal action is not merely evidence from which guilt is to be inferred but in fact constitutes criminal guilt.[13] It may be noted that the force of this point applies equally to a choice theorist's ascription of criminal action on the basis of an individual's choice.

Yet, the apparent conflict between these two subjectivist accounts, choice and character, stems from a false dichotomy between what a person 'is' and what he or she 'does'. As Silber has argued, 'what a man [or woman] does is a function of what he [or she], in the context of his [or her] situation, is, and what he [or she] is within this context is revealed by what he [or she] does'.[14] An individual's criminal responsibility is not determined by either his or her choice or character but action. Such action reveals, as the action of an independent agent, some inappropriate attitude towards the values the law protects.[15]

The following analysis will take, for ease of explication, the choice theory account of fault. As noted above, the choice theory ascribes actions to an individual on the basis

---

10 See Allan Norrie, 'Subjectivism, Objectivism, and the Limits of Criminal Recklessness' (1992) 12 *Oxford Journal of Legal Studies* 45.

11 R Antony Duff, 'Subjectivism, Objectivism and Attempts' in AP Simester, ATH Smith (eds), *Harm and Culpability* (1996) 19, 20.

12 Peter Arenella, 'Character, Choice and Moral Agency' in EF Paul (ed) *Crime, Culpability, and Remedy* (1990) 78.

13 R Antony Duff 'Choice, Character and Criminal Liability' (1993) 12 *Law and Philosophy* 345, 379.

14 J R Silber, 'Being and Doing: A Study of Status Responsibility and Voluntary Responsibility' (1967) 35 *University of Chicago Law Review* 47, 65 (emphasis in original).

15 Duff, above n 13, 380.

of his or her choices.[16] Such choices belong to an individual to the extent that they manifest his or her *intentions* and *beliefs*. As expressed by Ashworth agents should be held 'criminally liable for what they chose to do, and not according to what actually did or did not occur' and must be 'judged on the basis of what they believed they were doing, not on the basis of actual facts and circumstances which were not known to them at the time'.[17] What is required as a necessary precondition of criminal liability is subjective advertence. In the absence of such advertence there can be no liability.[18]

In contrast, an objectivist ascribes actions to an individual on the basis of one of two aspects of action, namely,[19] a result which occurs in the real world due to the action being performed, or what a 'reasonable' person would believe or realise would be the result of performing such an action. Hence, an objectivist will focus on what an individual actually does as distinct from whether he or she has adverted to what is done.[20]

The distinctions alluded to above between the different bases of ascription are made most apparent when we come to consider the various fault elements of an offence. So, for example, it may be noted that recklessness is often viewed as the minimum fault requirement for a criminal offence.[21] This is exemplified by the Commonwealth Criminal Code which proceeds on the basis that conscious advertence is a precondition of criminal liability.[22] Hence, in the context of a result-crime what is required is advertence on the part of the accused individual as to the consequence ensuing in the completed offence.[23]

Such recklessness has two aspects, namely, a mental element and a failure to comply with a standard of conduct.[24] Moreover, the former mental element is comprised of both deliberation and knowledge. That is, reckless conduct in addition to being

---

16 Ibid.

17 Andrew Ashworth, 'Criminal Attempts and the Role of Resulting Harm' (1988) 19 *Rutgers Law Journal* 725, 736.

18 Simester, Smith, above n 11, 8.

19 Duff, above n 11, 21.

20 Simester, Smith, above n 11, 8.

21 Cp *Criminal Code* 1995 s 5.4 (4).

22 Ian Leader-Elliott, The Commonwealth Criminal Code: A Guide for Practitioners (2002) 73.

23 *Criminal Code* 1995 (Cth) s 5.4(2); Cp s 5.4(4): 'If recklessness is a fault element for a physical element of an offence, proof of intention, knowledge or recklessness will satisfy that fault element'.

24 Peter Cane, *Responsibility in Law and Morality* (2002) 80 for the following analysis of respective fault elements.

deliberate also requires an awareness of a risk that is unreasonable in the particular circumstances. The requisite degree of such risk, while a matter of some contention, is considered by many analysts to require the possibility of a given result ensuing. [25]

In contrast to the subjective advertence required for the fault element of recklessness, negligence consists of a failure to reach a standard of conduct. The requisite standard requires that reasonable care be taken in the course of engaging in a given activity. Hence, negligence is the failure to take reasonable precautions against a foreseeable risk of injury.[26] However, such failure to meet the requisite standard will warrant censure only where it may be characterised as 'gross'. It has been remarked that such a requirement is meant to 'blunt the argument that justice requires advertent wrongdoing before criminal liability is imposed'.[27] It is only in taking such grossly obvious risks that liability ensues, for such risks would be readily apparent to a reasonable person in like circumstances. [28]

Clearly, such a basis for ascription runs counter to the subjectivist dictum calling for advertence.[29] However, equally clear from the objectivist viewpoint is the argument that the very unreasonableness of the behaviour engaged in warrants condemnation notwithstanding any failure of advertence on the part of the accused.[30]

(b) Mental Impairment and Criminal Responsibility: The M'Naghten Rules,[31] establish that cognitive incapacity, 'defect of reason', must be caused by a 'disease of the mind'. That is, a causal link between an accused's impaired reasoning and an underlying 'disease' is required.

What will constitute a 'disease of the mind' is a matter to be determined by law not psychiatry.[32] The relevant question is not whether the accused suffers from a diagnosable mental illness, but rather whether there is an adequate degree of impairment.[33] As expressed by Glanville Williams, insanity is a social judgment

---

25 Simon Bronnitt and Bernadette McSherry, *Principles of Criminal Law* (2001) 183, for all offences other than murder.

26 Cane, above n 24, 80.

27 Leader-Elliott, above n 13, 85 citing *The Queen v Nydam* [1977] VR 430.

28 Cane, above n 24, 80.

29 Jerome Hall, 'Negligent Behaviour should be Excluded from Penal Liability' (1963) 63 *Columbia LR* 632.

30 Michael Moore, 'Choice, Character and Excuse' in EF Paul et al above n 12, 29 at 58.

31 *R v M'Naghten* (1843) 4 St Tr (ns) 847.

32 *R v Kemp* [1957] QB 399,406.

33 *R v Sullivan* [1984] AC 156.

founded upon but not precisely representing, a medical diagnosis.[34] The issue of whether such impairment satisfies the legal definition of 'disease of the mind' is at all times a question of law not fact. [35] What remains a question of fact for the purposes of analysis is whether the accused suffered from a 'disease of the mind' at the relevant time in question.[36] This latter factual question will be the subject of expert evidence.[37]

No organic basis, in the sense of physical pathology, need be established for the condition under which the accused may be said to be suffering. Yet, what is called for is that the 'functions of the understanding are through some cause, whether understandable or not, thrown into derangement or disorder'.[38] 'Disease of the mind' will cover all conditions recognised by psychiatry,[39] including those with a physical basis such as arteriosclerosis,[40] epilepsy,[41] and diabetes. [42] Nevertheless, 'the condition of the brain is [ultimately] irrelevant as is the question of whether the condition of the mind is curable or incurable, transitory or permanent'.[43]

Hence, it has been said that the limits to the notion of 'disease of the mind' are determined by the concept of pathology.[44] As a result 'conditions of intense passion and other transient states attributable to the fault or to the nature of man' do not constitute a 'disease of the mind'. [45] Such latter conditions may be characterised as 'the reaction of a healthy mind to extraordinary external stimuli'.[46]

McAuley and McCutcheon have noted several reasons why the law of insanity is not coextensive with the psychiatric conception of mental disorder.[47] The authors note the following. First, legal insanity is an excuse for wrongdoing, not a diagnosis of the

---

34 Glanville Williams, *Textbook of Criminal Law* (2nd ed, 1983) 644.

35 *R v Kemp* [1957] 1QB 399, 406; *R v Falconer* (1990) 171 CLR 30, 48-9.

36 *R v Porter* (1933) 55 CLR 182, 188.

37 *R v Bromage* [1991] 1Qd R 1, 6.

38 *R v Porter* (1933) 55 CLR 182,189.

39 Owen Dixon, 'A Legacy of Hadfield, McNaghten and Maclean' (1957) 31 *Australian Law Journal* 255, 260.

40 *R v Kemp* [1957] QB 399.

41 Bratty v A-G for Northern Ireland [1963] AC 386.

42 *R v Quick* [1973] QB 910.

43 *R v Kemp* [1957] 1 QB 399, 407.

44 Finbarr Mc Auley and John Paul McCutchen, *Ciriminal Liability* (2000) 683.

45 Dixon, above n 39, 260.

46 *R v Radford* (1985) 42 SASR 266, 274, (approved *R v Falconer* (1990) 171 CLR 30).

47 Mc Auley, McCutchen, above n 44, 638-44 from where the following analysis is taken.

accused's mental condition. As such, the scope of the defence is not amenable to expert evidence alone.[48]

Secondly, something more than an organic basis is required, it being the degree of impairment not its cause that is determinative. The legal construct, 'disease of the mind', operates as a criterion of non-responsibility. It provides a means by which to distinguish exculpatory from non-exculpatory conditions. [49] The legal issue remains whether the accused's ' illness' is sufficient to exculpate, not whether it satisfies a recognised diagnostic category.[50]

Thirdly, it must be shown that such a 'disease of the mind' was the cause of the wrongdoing.[51] That is, the defence of insanity must be connected with the proscribed conduct if it is to have any exculpatory force.

Fourthly, the insanity defence at common law is not to be confused with the 'product test'. This latter test seeks to provide an excuse to an accused 'if his [or her] unlawful act was the product of mental disease or mental defect'.[52] Such an argument is misconceived in that it fails to account for the way in which disease affects action. According to the 'product test' any organically determinable condition causing a particular reaction will exonerate. This will be the case whether or not the particular individual concerned could have refrained from reacting in that particular way. [53] Yet, as Morse has noted, such an analysis confuses causation with excuse. For, causation is not necessarily compulsion. Succinctly put, 'if causation were an excuse, no one would be held responsible for any behaviour, criminal or not'.[54]

The manner in which physical disease operates in excusing an individual differs from that in the case of mental disorder. In the former instance excuse stems from the fact that an individual is acted upon rather than acting. That is, there is nothing which an individual suffering from a physical disease actually does. While there may be, for example, in the instance of Parkinson's disease a movement by the agent, such

---

48 Jennifer Radden, *Madness and Reason* (1985) 28.

49 Ian G Campbell, Mental Disorder and Criminal Law in Australia and New Zealand (1988) 128.

50 McAuley and McCutcheon, above n 44, 640.

51 *R v Kemp* [1957] QB 399, 407 per Devlin J.

52 *Durham v United States* (1954) 214 F.2d 862, 874.

53 McAuley and McCutcheon, above n 44, 641-4.

54 Stephen J Morse, ' Excusing the Crazy: The Insanity Defense Reconsidered' [1985] 58 Southern California Law Review 777, 789; Cp Michael Moore *Law and Psychiatry: Rethinking the Relationship* (1984) 362.

movement is not an action of the individual. However, this is not the manner in which mental 'disease' operates.

In the instance of mental disorders the 'disease' manifests itself through the individual's psychological state. Such states consist of that individual's desires, beliefs and intentions. Responsibility in relation to these states is not automatically negated because of the existence of an underlying 'disease'. Such reasoning would follow only where the claim 'Mental disorder is not blameworthy' is held to entail the belief that 'Mental disorder mitigates blame for wrongdoing'.[55] Yet, this need not necessarily follow as an individual's behaviour may be condemned even where such behaviour is recognised as part of his or her character. It may be such an individual's very failure to keep such characteristics in check, assuming that others would be able to do so, which grounds our condemnation.[56]

The legal construct, 'disease of the mind', depends on an overly broad conception of disorder as a criterion of responsibility: where such disorder is deemed to deprive the accused of certain capacities he or she is held non-responsible. Such breadth of application has been restricted, however, by explicit recourse to policy-oriented questions pertaining to an individual's future dangerousness and his or her disposition. McSherry has noted three tests that have been explicitly referred to in case law relating to the 'threshold question of did the accused at the time of the commission of the offence suffer from a disease of the mind?'[57]

The *recurrence test*[58] which focuses attention on the likelihood of the operative mental state's recurrence; in cases where such a condition is likely to recur it may be characterised as a 'disease of the mind';

the *internal/external test*[59] seeking to delimit the characterisation of 'disease of the mind' to mental states which have an internal rather than an external aetiology; and

the *sound/ unsound mind test*[60] where the characterisation 'disease of the mind' describes the responses of an unsound mind to both its own internal and any external stimuli.

---

55 See,Radden, above n 48, 30-3.

56 McAuley and McCutcheon, above n 44, 643.

57 Bernadette McSherry, 'Defining What is a "Disease of the Mind": The Untenability of Current Legal Interpretations' (1993) 1 *Journal of Law and Medicine* 76, 82-89.

58 *Bratty v Attorney General for Northern Ireland* [1963] AC 386.

59 *R v Rabey* (1977) 37 CCC (2d) 461; *R v Hennessy* [1989] 1 WLR 287.

60 *R v Falconer* (1990) 171 CLR 30.

As previously noted, the legal construct 'disease of the mind' is restricted by application of a policy-oriented question. Succinctly stated, this involves an enquiry as to the 'dangerousness' of a given individual. Where it is held that such an individual is likely to harm others his or her mental condition will be characterised as a 'disease of the mind'. While no longer held responsible, such an individual will be the object of a therapeutic regime.

The following three sections will analyse several approaches that have been adopted in assigning criminal responsibility in cases of drug-induced states of mental impairment.

Causal Irrelevance

In the context of their discussion of people with mentally impaired functioning who kill the Victorian Law Reform Commission (VLRC) turned its attention to clarifying the current scope of mental impairment.[61] The VLRC objected in particular to the Supreme Court of Victoria's interpretation of 'mental impairment' as being 'merely a legislative restatement of the common law'.[62] The Commission argued that adopting a common law definition of mental impairment will not provide a sufficient level of flexibility in the application of the defence.[63] By way of example the Commission cited the case of *R v Sebalj* [64] where the accused was found guilty of murder while in a psychotic state directly related to drug use. During the course of treating his addiction the accused killed his girlfriend. As the accused's psychosis had arisen due to drug use, the court held that it did not fit the definition of 'mental impairment'[65] which, following the common law, required the illness to be a 'disease of the mind'.[66] Not able to use the defence of 'mental impairment' the accused was sentenced to 15 years imprisonment for murder.[67]

---

61 VLRC op cit n 7, 203-52, 213 para 5.34.

62 Ibid para 5.34 citing *The Queen v R* [2003] VSC 187 (Unreported, Supreme Court of Victoria, Teague J, 5 March 2003), *The Queen v Sebalj* [2003] VSC 181 (Unreported, Supreme Court of Victoria, Smith J, 5 June 2003).

63 Ibid para 5.35.

64 [2003] VSC 181.

65 The term 'mental illness' while not defined appears in *Crimes (Mental Impairment and Unfitness to be Tried) Act 1997*, s 20.

66 *The Queen v Sebalj* [2003] VSC 181 para 14 per Smith J.

67 *The Queen v Sebalj* [2004] VSC 212 (11 June 2004) per Williams J.

The Commission noted that there are several objections to allowing a person whose psychosis is induced by alcohol or drugs to rely on mental impairment. In particular 'there is a moral argument that a person ought not to be able to raise the mental impairment defence if they were responsible for causing their condition'.[68] The Commission stated that it did not think this should be a 'consideration in applying the defence of mental impairment'.[69] By way of argument they drew an analogy with current Victorian law in regards to gross intoxication. They noted that *O'Connor*,[70] as the leading case in Victoria, was authority for the view that where an accused's state of intoxication is such as to preclude the capacity to form an intention, he or she must be acquitted. By parity of reasoning the Commission argued an accused's drug-induced state of mental impairment should, where the requisite incapacity is established, lead to a like acquittal.[71]

As a result, the Commission noted, on the current approach in Victoria accused who are so intoxicated that they cannot form an intention to commit a crime cannot be held criminally liable. Yet, the Commission went on to note, accused who are so intoxicated or affected by drugs that they experience a psychotic episode, such that they are either unable to understand what they are doing or that it is wrong, may be held criminally responsible.[72]

The Commission remarked on two principal issues which render drug related states of mental impairment particularly problematic. As the Commission put it, there is a great deal of difficulty in delimiting the notion of responsibility.[73] For example, in the case of *Sebalj* the accused was responsible for the psychosis as he had chosen to engage in the use of drugs the effects of which ultimately had brought on a psychotic episode.

Secondly, the Commission remarked on the fact that the court's ruling on mental impairment in the case of *Sebalj* failed to acknowledge the defence's true purpose. Such purpose, according to the Commission, would preclude taking into account the causal antecedent of the state of impairment. What is critical for the operation of the defence is the cognitive deficiency that gives rise to the requisite incapacities, not how such deficiency came about.[74] An examination of these two issues, that of

---

68 VLRC op cit n 7, 214 para 5.37.

69 Ibid 215 para 5.40.

70 *The Queen v O'Connor* (1980) 146 CLR 64.

71 VLRC, above n 7, para 5.42.

72 Ibid.

73 VLRC, above n 7, para 5.43.

74 Ibid.

responsibility and the purpose of the mental impairment defence in the context of drug induced psychoses will be the subject of this section.

As stated above, the Commission remark on the fact that the notion of 'responsibility' in the context of *Sebalj* is difficult to limit. [75] In that case the accused was responsible for the psychoses as he had chosen to withdraw from drugs and it was that withdrawal which had brought on a psychotic episode. Whether an accused should be permitted to benefit from the excusing condition he or she brings about is a moral (and by implication legal) question.

Clearly, the Commission is correct in noting that there is a moral concern attendant upon allowing a person to benefit as a result of a condition that he or she has been responsible in producing.[76] Both subjectivist and objectivist accounts of criminal responsibility endeavour to answer such a concern. As noted above, a subjectivist requires as a necessary minimum for the ascription of actions to an accused, a degree of subjective advertence. Such advertence may be readily established in the context of drug taking given the widespread acknowledgment of the risks attendant on the use of drugs. Similarly, an objectivist account may view drug use that leads to criminal behaviour, in and of itself, commensurate with a degree of gross negligence. However, in both instances, subjectivist and objectivist alike, there is an assumption that the creation of the conditions under which an offence is committed is tantamount to an intention to commit or risk committing such an offence. This issue will be further discussed in the context of my analysis of the concept of 'prior fault', or 'meta-responsibility', in section IV below.

Implicit in the Court's judgment in the *Sebalj* case above is the choice of test used in response to the threshold question of did the accused at the time of the commission of the offence suffer from a 'disease of the mind'?[77] The Court seemed to favour the 'internal/external' test that provides that if the mental state operative at the time of the offence is 'internal' to the accused, as opposed to arising from an external cause, it should be defined as a disease of the mind.[78] The threshold question is policy-oriented asking in effect whether the accused is likely to cause harm to others if not confined? If so, then his or her mental condition will be considered to be a disease of the mind for the purposes of the mental impairment defence. It is implicit in the argument that

---

75 Ibid.

76 Ibid, para 5.40.

77 McSherry, above n 57, 82-9.

78 *R v Rabey* (1977) 37 CCC (2d) 461; *R v Hennessy* [1989] 1 WLR 287.

only those with mental conditions arising from some internal source should be detained.[79]

Many of the current legal 'tests' utilise aspects of the conventional 'disease' model of mental impairment. The M'Naghten Rules[80] are limited by the fact that the exculpatory mental state of the accused must proceed from a disease of the mind. For example, in *R v Quick*[81] the Court stated that insane automatism, arises from a 'disease of the mind' which is a 'malfunctioning of [the] mind…caused by a bodily disorder in the nature of a disease'.[82] On the other hand, the Court held that 'non-insane' or 'sane' automatism is not due to a disease of the mind but is the result of 'the application to the body of some external factor such as violence, drugs, including anaesthetics, alcohol and hypnotic influences…'.[83]

However, as noted above, causal questions should be distinguished from questions of responsibility. It is not necessarily exculpatory to be informed in the context of mental states what the cause of the disorder is. The issue is not what was the cause but rather whether there has there been the requisite loss of capacity. For example, Mackay has remarked that in the case of somnambulism it is sleep itself that is the cause of the automatism.[84] As sleep is a 'normal' condition the accused's 'disorder' is internal.[85] However, as Mackay remarks, even if one accepts that sleepwalking can properly be described as pathological, there is still the problem that the cause of the impairment was sleep rather than sleepwalking itself. Mackay concludes by noting that unless the requirement of a causal link between disease of the mind and defect of reason within the *M'Naghten Rules* is to be disregarded in the case of sleepwalkers, it seems difficult to accept that a sleepwalker who is otherwise mentally normal should be classed as legally insane.[86]

The issue of an individual's rationality may be distinguished from the cause of his or her rationality. While the former is of concern to the question of responsibility the latter is not. Similarly, the cause of an individual's conduct is not of immediate concern in relation to his or her conduct. It has been remarked that where control is

---

79 McSherry, above n 57, 89.

80 *R v M'Naghten* (1843) 4 St Tr (ns) 847.

81 [1973] 3 All ER 347; Cp *Bratty v Attorney-General for Northern Ireland* [1961] 3 All ER 523.

82 Ibid 356.

83 Ibid.

84 Ronald D Mackay, Mental Condition Defences in the Criminal Law (1995) 49.

85 *R v Burgess* 2 All ER 769, 775 per Lord Lane; *R v Parks* (1990) 78 CR (3d) 1, 19 per Galligan JA.

86 Mackay, above n 55, 50.

extant an individual may be responsible for a spontaneous reflex movement, even where such a movement is caused by disease.[87] As has been noted, so long as such control is in place an individual remains a fit subject of responsibility: control should not be confused with cause as an individual may be in control of his behaviour even where he or she has not caused it. Further, one need not be the cause of something to be in control of that thing. [88] This is made explicit in Frankfurt's account of action as responsible movement.

Frankfurt differentiates action from movement on the basis of an agent's being 'in touch with the movements of his [or her] body in a certain way'.[89] According to Frankfurt what is required for responsibility is 'guidance control'. Such guidance guides behaviour by subjecting it to adjustments compensating for the effects of interfering causes.[90] However, the operation of the guidance system is not dependant on the nature of the interfering causes. It has been remarked that as responsible agents individuals may be viewed as guidance systems responding to the demands of both law and morality. [91] A corollary of this is that as a guidance system the cause of any given individual's mental state and attendant action cannot be determinative of his or her responsibility.

Notwithstanding the above I would contend that the VLRC is mistaken in thinking that the underlying purpose of the insanity defence is satisfied by focusing exclusively on an individual's rational capacity.

It was noted above that the law's construct of disease does not reflect concepts of responsibility. 'Disease of the mind' as a disorder of the reasoning process leading to a specified lack of capacity is delimited in its application by means of the question of dangerousness and clinical intervention. However, as Campbell has noted, as a criterion of mental illness a concern with the disposition of those considered dangerous conflates two issues: the 'utilitarian' issue of what to do with the non-responsible dangerous individual and the moral issue of when to hold an individual criminally responsible.[92] The VLRC approach seeks to establish equivalence between

---

87 *Ryan v The Queen* (1967) 40 ALJR 488, 505 per Windeyer J; ID Elliott 'Responsibility for Involuntary Acts: Ryan v The Queen' (1968) 41 *Australian Law Journal* 497,500.

88 S Shute, J Gardner and J Horder, 'Introduction: The Logic of Criminal Law' in Stephen Shute, John Gardner and Jeremy Horder (ed), *Action and Value in Criminal Law* (1993) 1, 19.

89 Harry Frankfurt, The Importance of What we Care About (1990) 71.

90 Michael Smith ,'Responsibility and Self-Control' in Peter Cane and J Gardner (ed), *Relating to Responsibility* (2001) 1, 12 noting 'back-up capacities'.

91 Shute, Gardner and Horder, above n 88, 19; See, Joseph Raz, *Practical Reason and Norms* (1990) 140-6.

92 Campbell, above n 49,129.

the notion of 'disease' and 'non-responsibility' absent consideration of any operative policy issues. The force of such policy issues and the way in which they impinge on questions of responsibility is transparent in the approach courts adopt to the issue of mens rea and insanity.

An accused raising a defence of mental impairment may argue a lack of mens rea where such a mental element is called for by the relevant offence. However, there has been reluctance on the part of courts to characterise such denials of fault as amounting to the prosecution case not being proved. Rather, courts have consistently 'refused to permit evidence of insanity to be used to support an argument of lack of mens rea independently of the insanity defence.'[93] The reason for this reluctance on the part of the courts stems from the fact that such an approach fails to pay heed to the policy questions at the heart of the insanity defence.

Similarly, the VLRC approach fails to acknowledge that the threshold question of what is a 'disease of the mind' is a policy question relating to the likelihood of social danger in the absence of clinical intervention. Recourse to an internal/external and recurrence test, while, as we shall see below, less than a complete safeguard, does nevertheless provide a means by which to ensure such societal protection.

In the following section I turn to a consideration of an approach that stands in marked contrast to the VLRC position, namely, Edward Mitchell's account of 'meta-responsibility'.

'Meta-Responsibility'

In his discussion of 'meta-responsibility', or a person's responsibility for non-responsibility, Mitchell begins by noting an historical precedent, namely Stephen's denial of an insanity defence to any person whose 'absence of the power of control has been produced by his [or her] own default'.[94] The idea of prior fault, according to Mitchell, finds further support in modern legal research, principally in the work of the American scholar Paul Robinson.[95]

In developing the concept Mitchell argues, first, that a defendant's culpability for his or her own meta-responsibility should be reflected in a denial of a mental condition

---

93 Paul A Fairall, Stanley Yeo, *Criminal Law Defences in Australia* (4th ed, 2005) 267.

94 Mitchell (2004), above n 8, 66, citing Fitzjames Stephen, *A History of the Criminal Law of England* (1883).

95 Ibid 66, citing Paul H Robinson, 'Causing the Conditions of One's Own Defense' (1985) 71 *Virginia Law Review* 1.

defence.[96] Secondly, in responding to the question of how broadly to frame the time period during which one may look back to find such culpability Mitchell acknowledges 'that some practical limit will have to be set (on a case-by-case basis)'.[97] Such a search will, he claims, be necessarily delimited by an evidentiary test of relevance. Similar issues have been addressed by Paul Robinson in defending his own doctrine of 'creating the condition of one's own defence'.

According to Robinson, of primary interest is the question of whether an accused responsible for bringing about the conditions of his or her defence should be allowed to benefit from the defence? Robinson proposes a general principle to govern all instances of an accused's culpability in bringing about the conditions of his or her defence. Such a principle takes as the most critical factor an accused's culpability at the time he or she causes the conditions of his or her defence, rather than at the time of the offence.[98] On Robinson's analysis an accused will be provided with a defence for the immediate conduct constituted by the offence, but will be liable for the earlier conduct which culpably caused the conditions of the defence. Such a '*conduct-in-causing*' analysis ensures that an accused has an excuse or justification where he or she satisfies the conditions of such a defence.

Nevertheless, on Robinson's analysis, such an accused may be liable for the ultimate offence given that he or she has caused the excused conduct and has an accompanying culpable state of mind with respect to the commission of the offence. For example, a hypothetical cited by Robinson involves an accused that knows of his propensity to violent behaviour when intoxicated. This particular individual decides to kill his wife, and to this end begins to drink excessively with the express intention of bringing about her death. It may indeed be the case that at the time of the killing, given the accused's state of intoxication, he in fact lacks the requisite mental element of the offence, namely, intent or recklessness. Equally, he may have lacked any awareness of his conduct.

On similar facts the accused in *Attorney-General for Northern Ireland v Gallagher*[99] argued that his consumption of alcohol was liable to bring into operation a condition amounting to insanity. He argued that at the time of the offence he was in fact suffering from such a state and as such could not be held responsible for the resulting death. On appeal to the House of Lords it was held that a preliminary question that needed to be resolved concerned the circumstances under which the state of insanity

---

96 Ibid.

97 Ibid.

98 Ibid 27.

99 [1963] AC 349; Cp *The Queen v O'Connor* (1980) 146 CLR 64, 73 per Barwick CJ.

had been brought about. Their Lordships held that as the alcohol was consumed in order to bring about the state of insanity the appellant could not have recourse to the defence. Any other decision would be tantamount to condoning the means chosen consciously to bring about the state of non-responsibility.[100]

Fisse has noted that cases such as *Gallagher* highlight the purpose of the distinction drawn between voluntary and involuntary intoxication.[101] Fisse further notes that what such a case highlights is that where drunkenness is put in issue by an accused 'the inquiry, which is to say the relevant time, must be carried back far enough to ascertain whether the intoxication was entered into as a means to an end'.[102] For, if it had been the case that the state of intoxication was entered into for such a purpose it would be manifestly wrong to acquit.[103]

It has been suggested that this type of argument reflects the elasticity of time and blameworthiness in the criminal law.[104] Yet, such an analysis of the backward extension of the relevant time fails to pay sufficient heed to the relevant time frame for viewing an offence in the criminal law. As Fisse has observed, 'the time finally decided on as marking the first significant event in no way affects the time and occurrence of the last significant event, which remains relevant throughout'.[105] It is a series of events rather than a 'time-slice' view of events that is determinative of criminal responsibility. The evidentiary import of such a series of events will, of course, be delimited by the causal connection existing between the accused and the proscribed outcome. [106]

The causing-one's-defence doctrine is ultimately predicated on the view that where an accused causes the conditions of his or her own defence, and attendant upon the creation of these conditions he or she has the required fault element of the offence, he or she will be guilty of the completed offence. Such an account is analogous to the doctrine of 'innocent agency' whereby the excused accused is the 'innocent agent' who was caused to engage in the criminal conduct by the accused's prior, culpable

---

100 [1963] AC 349, 382 per Lord Denning.

101 Brent Fisse, *Howard's Criminal Law* (5th ed, 1990,) 15-16.

102 Ibid.

103 Ibid; AG For Northern Ireland v Gallagher [1963] AC 349, 379.

104 Mark Kelman, 'Interpretive Construction in the Substantive Criminal Law' (1980) 33 *Stanford Law Review* 591; Contra: L Shwarz, 'With Gun and Camera through Darkest CLS-Land' (1984) 36 *Stanford Law Review* 413 both cited by Fisse, above n 101, 13n 65.

105 Ibid 15.

106 Ibid 16.

actions.[107] As Robinson's account of the operation of the doctrine indicates, after the jury is instructed as to the state of mind requirements of the underlying offence, they may be instructed that 'the accused may satisfy these elements either at the time of the offence or at the time he or she causes the conditions of a defence'.[108]

Nigel Walker has contended that '"meta-responsibility"… raises genuine issues for retributively-minded moralists'.[109] Such issues include both the legal and moral problems arising from the view that an individual is 'punishably culpable for doing something which merely risked an unintended and far from certain result'.[110]

Lawton LJ in *R v Quick* noted that '[a] self-induced incapacity will not excuse (see *Lipman* [1970] QB 152) nor will one which could have been reasonably foreseen as a result of doing, or omitting to do something…'.[111] As noted above, in a case such as *Gallagher* an accused may, intending to kill another, set about becoming intoxicated in order to undermine his or her requisite culpable mental state. In such a case he or she will be held to have the pre-existing intention. Clearly, in such an instance it is not the mental state of the defendant at the time of the proscribed act that is of utmost importance but rather the subjective state preceding such an act. Having placed him or her self in such a position he or she is responsible for any resulting harm, regardless as to whether he or she has the requisite mental state for the harm so produced. Such a requisite state may involve either intent or a subjective advertence grounding recklessness.

However, 'the transfer of a pre-existing intention and/or recklessness from one time to another is one thing; the deeming of that transfer is another'.[112] This is the kernel of truth in the case of *The Queen v O'Connor* where an accused becomes so intoxicated that he or she cannot form an intention to commit a crime cannot be held criminally liable.[113] At common law it is not correct to say that there may be an automatic transfer of deemed fault to a proscribed result from previous behaviour.[114] The reason why this is so is that a distinction is to be drawn between 'on the one hand, the result of an act

---

107 Robinson, above n 98, 54; on 'innocent agency' see Bronitt and McSherry, above n 25, 402-4.

108 Ibid.

109 N Walker, 'Book Review: Mitchell, EW (2003) Self-Made Madness. Rethinking illness and criminal responsibility' (2005) 16 (3) *Journal of Forensic Psychiatry and Psychology* 612, 614.

110 Ibid 614.

111 (1973) 57 Cr. App. R. 722, 735.

112 Mathew Goode, 'On Subjectivity and Objectivity in Denial of Criminal Responsibility: Reflections on Reading Radford' (1987) 11 *Crim LJ* 131, 146.

113 *The Queen v O'Connor* (1980) 146 CLR 64.

114 Ibid 80 per Barwick CJ.

of the conscious will and, on the other, an act of the conscious will with intent to do that act'.[115] As Goode observes, recklessness as to resulting harm is not relevant to the issue of whether an accused acted at all as these issues are distinct from one another.[116]

Nevertheless, our intuition suggests that any defendant who has voluntarily consumed drugs and then caused harm is responsible. This intuition is operative even where at the time of becoming intoxicated the defendant has not formed the requisite fault element for the offence.[117] Given that the effects of drug taking are widely known it may be argued that an accused that places him or herself in such a situation has in fact satisfied the subjective fault requirement of advertence. Yet, as Goode observes, such a claim of common knowledge grounding advertence amounts to '*constructive* or *deemed* fault or culpability. The accused *must be taken to know* that a high degree of intoxication may lead to some kind of violence'.[118] However, such a claim to constructive knowledge is, at best, arguable.

Indeed, the contention that a self-intoxicated accused should be held liable notwithstanding the absence of a requisite mental state on his or her part amounts to an argument for a general exception to the principle of subjective fault.[119] To that degree it is open to question whether 'retributively-minded moralists' such as Mitchell and Robinson would be able to defend their view of 'meta-responsibility' without showing more. Mitchell and Robinson would need to show that such a shift away from subjective fault is justified under the circumstances.

The next section will consider the merit of the views expressed by those theorists who argue for a 'settled insanity' approach to the question of drug-induced psychoses.

'Settled Insanity'

At common law a distinction is drawn between 'an underlying mental infirmity' required for a 'disease of the mind' and a transient, externally caused, non-recurrent mental malfunction.[120] Where such mental malfunction is caused by something internal to an accused and is prone to recur, a 'disease of the mind' is established. Lord Denning has held that while not criminally responsible such an individual

---

115 Goode, above n 112, 147.

116 Ibid.

117 *The Queen v O'Connor* (1980) 146 CLR 64, 97 per Mason CJ (diss).

118 Goode, above n 112, 150 (emphasis added).

119 Ibid.

120 *R v Radford* (1985) 42 SASR 266 per King CJ.

should be detained on the grounds of societal protection.[121] Alternatively, where the mental malfunctioning results from a cause lying external to the accused, is transient and unlikely to recur, civil detention is not required.[122] The *M'Naghten Rules* call for a defect of reason to be present at the time of the act. Hence, legal impairment may be established even in situations where a mental malfunction is not likely to continue.[123]

The Victorian case of *R v Meddings*[124] is authority for the view that any condition which is likely to recur, and to that degree presents an element of risk to public safety, should be treated as a disease of the mind. However, the task of predicting future dangerousness with any degree of certainty remains problematic.[125] Additionally, any wholesale adoption of recurrence as a criterion delimiting the scope of the 'disease of mind' enquiry fails to acknowledge any serious mental disease not likely to recur.[126]

The 'internal/external' test has sought to ground the 'disease of the mind' enquiry on a distinction between 'the reaction of an unsound mind to its own delusions or to external stimuli on the one hand' and, on the other hand, 'the reaction of a sound mind to external stimuli, including stress producing factors.'[127] Similarly, in *Meddings* it was held that a 'predisposition, whether resulting from injury or from some other idiopathic cause…[is] a disease of the mind'.[128] Where there is such a predisposition it matters not whether the trigger is 'alcohol…a set of surrounding circumstances…a provocative word or some object which arouses recollection or emotion'.[129]

However, the underlying condition/external cause distinction has not met with universal approval. Most problematic is the over-exclusiveness of the test and its potential to lead to inconsistent results.[130] To take some examples, hyperglycemia (an excess of blood sugar) has been treated as an internal cause, giving rise to a mental impairment defence, [131] while hypoglycemia (a blood sugar deficiency) has been

---

121 *Bratty v Attorney-General (Northern Ireland)* [1963] AC 386, 412.

122 *Cottle* [1958]NZLR 999 at 1015 per Gresson P.

123 *Kemp* [1957] 1 QB 399

124 [1959] VR 105.

125 Steven Yannoulidis, 'Negotiating Dangerousness: Charting a Course between Psychiatry and Law' (2002) 9 (2) *Psychiatry, Psychology and Law* 151.

126 McSherry above n 57, 83.

127 *R v Radford* (1985) 42 SASR 266, 276 per King CJ; applied *R v Falconer* (1990).

128 [1966] VR 306, 310.

129 Ibid.

130 *The Queen v Falconer* (1990) 171 CLR 30, 75 per Toohey J.

131 *R v Hennessy* (1989) 1 WLR 287.

viewed as an external source precluding such a defence.[132] Equally, as noted above, states of somnambulism have proven to be problematic it not being apparent that such may be convincingly described as 'externally caused'.[133]

A test that seeks to distinguish internal from external causes is too vague to do the work required of it. As Mathew Goode has asked, why is a hypoglycaemic episode precipitated by a combination of insulin and a lack of food externally caused rather than being viewed as a concomitant of the diabetic condition?[134] The only cogent response to this type of difficulty is to acknowledge that 'such conditions are brought about by a variety of factors, some externally based and some not'.[135]

Indeed, such criticisms have caused Toohey J to view the internal/external test as 'artificial…paying [as it does] insufficient regard to the subtleties surrounding the notion of external disease'.[136] Similarly, the Supreme Court of Canada has held that both the internal/external test and the recurrence test while of evidentiary import are not determinative of the 'disease of the mind' enquiry.[137] According to the Court such tests, along side other 'valid policy concerns', are to be viewed as analytic tools. Tools that will assist trial judges in 'answering the fundamental question of mixed law and fact which is at the centre of the disease of the mind inquiry: whether society requires protection from the accused'.[138]

An alternative approach to the internal/external test is to afford a defence in circumstances of 'settled insanity'. The 'settled insanity' approach focuses on the mental state of the accused *after* the criminal conduct. As originally outlined in the case of *People v Skinner* the hallmarks of such a state include a mental condition which is 'fixed and stable', lasts for a protracted period of time and, yet is not predicated on the consumption or operation of a drug.[139]

For Meloy '"settled insanity" becomes the legal point at which voluntary behaviour creates a state of mind that negates culpability for criminal behaviour'.[140] According to Meloy both temporal and causal problems remain with the formulation in *Skinner*. In

---

132 *R v Quick* (1973) 3 All ER 347.

133 *R v Parks* (1992) 75 CCC (3d) 287; *R v Stone* (1999) 134 CCC (3d) 353.

134 Goode, above n 112, 142.

135 Ibid.

136 *R v Falconer* (1990) 171 CLR 30, 75-6.

137 *R v Stone* (1999) 134 CCC (3d) 353, 439 per Bastarache J.

138 Ibid 441.

139 *People v Skinner* (1986) 185 Cal. App. 3d 1050 cited Meloy, above n 9, 438.

140 Meloy, above n 9, 439.

seeking to address the ambiguity in the term 'settled' the *Skinner* court employed 'two further semantically ambiguous and clinically meaningless terms, "fixed" and "stable"'.[141] In so doing the court had, Meloy argues, merely substituted semantically equivalent words and had failed to create a new distinction aiding clarification.[142] Additionally, the causal question raised by such cases is how to clinically distinguish the mental state from the drug's effects. Yet, as Meloy notes, the court while requiring that the mental state not be 'solely dependant' on the effects of the drugs taken failed to provide any means by which to ascertain whether this was indeed the case.[143]

In order to alleviate such concerns Meloy suggests 'that courts turn to the scientific literature for direction to formulate a judicial principle that can be applied to the facts in any one particular case'.[144] And in keeping with 'current scientific research', he proposes that 'the concept of "settled insanity" be limited to cases in which a predisposition to psychosis can be substantially demonstrated'.[145] While evidentiary issues would arise these need not be insuperable it being understood that '*but for* the presence of a vulnerability to psychosis, "settled insanity" would not apply'.[146] Several advantages are said to follow from such an approach. In particular, it would eliminate any problem associated with defining 'settled', avoid the arbitrariness associated with the notion of what constitutes a 'reasonable' duration of time, and 'flush out other factors that contributed to the mental state beyond the ingestion and duration of the drug'.[147]

Moreover, the above formulation, it is claimed, would meet squarely the issue of prior culpability in the context of drug-induced psychosis. Meloy argues that the absence of any history in the nature of psychiatric or psychological data would preclude an insanity defence in a case of voluntary intoxication, 'even if the individual had a diagnosable psychotic disorder at the time of the crime that fit the jurisdiction's definition of "insanity"'.[148] Additionally, according to Meloy, such an approach argues for the view that 'despite the individual's voluntary ingestion of drugs, his [or her]

---

141 Ibid 449, citing *People v Skinner* (1986) 185 Cal. App. 3d 1050, 1063.

142 Ibid.

143 Ibid 449 citing *People v Skinner* (1986) 185 Cal. App. 3d 1050, 1063.

144 Ibid 450.

145 Ibid 451.

146 Ibid (italics in original).

147 Ibid.

148 Ibid.

biological proneness to psychosis is beyond his volitional control and therefore becomes the most salient factor in his exculpation'.[149]

Russ Scott and William Kingswell have argued that even if it is accepted that the influence of an external factor causes an acute psychosis this need not to be understood as satisfying the 'disease of the mind' enquiry.[150] According to the authors in keeping with what was said by King CJ in *R v Radford* what is called for in a finding of a 'disease of the mind' is something more than 'a temporary disorder or disturbance of an otherwise healthy mind caused by external factors'.[151] Analogously to Meloy, Scott and Kingswell contend that drug-associated psychosis will amount to a 'disease of the mind' only when the symptoms of psychosis have persisted longer than 28 days and in the absence of positive urine drug testing.

Several criticisms may be made in regards to the reasoning of those who support a 'settled insanity' approach. For ease of analysis I will restrict my immediate comments to the work of Meloy. First, Meloy contends that the 'settled insanity' approach be limited to those with a predisposition or vulnerability to psychosis. In order to bolster his argument he cites research suggesting that subjects who become psychotic while using psychostimulants have 'symptoms associated with psychotic disorders [that are] present prior to the drug use but [which are] subclinical'.[152]

Genetically predicated brain organisation variations have been referred to as *dispositions* to forms of mental disorder.[153] The expression 'disease' as a descriptive term is best restricted to those dispositions that are most readily classified as abnormal on the continuum along which such variations are located.[154] Where the term 'disease' is understood to include such dispositions we find a corresponding broadening of the concept of mental illness to incorporate both organic disorders and any 'operative tendency-to-disorder'.[155] The significance of this for the purposes of 'settled insanity' is that any account of mental disorder-as-predisposition while descriptively sound fails to provide the means by which to ascertain the point at which disposition becomes disease. Moreover, the likelihood of discovering any such means remains in

---

149 Ibid 452.

150 Scott and Kingswell, above n 1, 175.

151 (1985) 42 SASR 226, 274 applied *R v Falconer* (1990) 171 CLR 30, 49 per Mason CJ, Brennan and McHugh JJ.

152 Meloy, above n 9,451.

153 Michael Roth, 'Schizophrenia and the Theories of Thomas Szasz' (1976) 129 *British Journal of Psychology* 317, 321-2.

154 Finbarr McAuley, Insanity, Psychiatry and Criminal Responsibility (1993) 67.

155 Ibid 67-8.

the realm of conjecture. This point is readily acknowledged by Meloy who concedes that there is no scientific basis in support of 'a relationship between measurable psychosis proneness and subsequent psychotic symptoms following the use of psychostimulants'.[156]

Further, it is suggested by Meloy that proneness 'attenuates' the volitional problem of 'settled insanity' as a 'biological proneness to psychosis is beyond [the defendant's] volitional control and therefore becomes the most salient factor in his [or her] exculpation'.[157] Several observations concerning this claim deserve mention. If, on the one hand, it is taken to mean that notwithstanding the voluntary ingestion of drugs the defendant could not have foreseen the resulting psychosis and harm caused, this brings into sharp relief the principle of subjective fault. This issue was canvassed above in relation to the concept of 'meta-responsibility' and as such will not be rehearsed here. If, on the other hand, what is meant by Meloy's claim is that a lack of 'volitional control' is exculpatory, in and of itself, this does not necessarily follow.

Stephen Morse has argued that so called 'disorders of desire' inevitably beg social, moral and empirical questions.[158] Morse does however acknowledge that some disorders may produce a degree of physical and psychological stress such that an individual's general capacity for rationality is impaired.[159] This suggestion raises the question of the direction of causation in the instance of drug-induced psychoses and biological proneness: the protracted period of drug taking may in fact be the product of a proneness to psychosis rather than the other way around. Nevertheless, Morse is prepared to concede that where an individual's disorder affects his or her capacity for rationality a defence of 'settled insanity' should be available.[160] Yet, and this is the point, whether an individual is 'excused under such a regime would depend on a normative judgment about how much the general capacity…is undermined'.[161]

Both the Scott, Kingswell, and Meloy approaches to the question of settled insanity endeavour to provide formulations which accord with contemporary clinical research. Against such an endeavour to ground criminal responsibility on the basis of scientific literature lies the argument that it is 'public policy considerations concerning criminal responsibility, and not clinical science, [which] determine the issue of innocence or

---

156 Meloy, above n 9, 452.

157 Ibid.

158 Stephen J Morse, 'Rationality and Responsibility' (2000) 74 *Sth California Law Review* 251, 263.

159 Ibid.

160 Ibid 264.

161 Ibid.

guilt'.[162] Such policy concerns, it has been held, assist trial judges in 'answering the fundamental question of mixed law and fact which is at the centre of the disease of the mind inquiry: whether society requires protection from the accused'.[163] Indeed, the concept of mental disease has been described as a 'crude but effective policy tool for maintaining a balance between social protection and humanitarian concern for those lacking mental capacity'.[164] An accused will not be excused from criminal responsibility merely on the basis of clinical research. As previously noted, the meaning of the term 'disease of the mind' is a legal rather than a psychiatric question.[165] The issue is whether the accused's mental faculties were impaired by illness, not whether he or she was suffering from a recognised mental illness.[166]

Conclusion

I began my analysis by distinguishing between two types of question. The first is a question to be answered on the basis of clinical evidence: whether a diagnosable mental disorder exists independently of the consumption of drugs. The fact that an accused suffers from cognitive incapacity is not, in and of itself, conclusive. What needs to be established is a causal nexus between such cognitive incapacity and an underlying 'disease of the mind'. The precise nature of what constitutes a 'disease of the mind' is a question informed by a social judgment. Such a social judgment, while founded upon a medical diagnosis, remains a question of law. As previously noted, mental impairment is an excuse for wrongdoing rather than a diagnosis of an accused's mental condition.

The second question is whether an individual who has voluntarily consumed drugs that have led to a state of impaired reasoning powers should be viewed as suffering from a 'disease of the mind'. The construct of 'disease' employed by the law in determining these issues is however apt to mislead. It does not follow that because an accused has impaired cognitive functioning he or she is to be excused. As noted previously, recourse to the notion of 'disease' erroneously facilitates the conflation of compulsion and causation. Where an agent is compelled to act he or she lacks agent causation and is for this reason excused. However, the same may not be said of all cases involving accused who have been caused to act. Whether these latter accused

---

162 Fairall and Yeo, above n 94, 255.

163 *R v Stone* (1999) 134 CCC (3d) 353, 441 per Bastarache J.

164 Fairall and Yeo, above n 94, 261.

165 *R v Kemp* [1957] QB 399, 406.

166 *R v Sullivan* [1984] AC 156.

will be excused is a question that will be answered on the basis of a normative judgment. Such a judgment will take into account the accused's ability to respond actively to reasons both for and against behaving in a particular way. That is, at the heart of the normative judgement is a consideration of an accused's capacity to control his or her behaviour.

As Raz has explained 'we are active when our mental life displays sensitivity to reasons, and we are passive when such mental events occur in a way which is not sensitive to reasons.'[167] The underlying force of the distinction being drawn between moral activity and passivity also informs discussions of dissociation in the criminal law. As case law indicates automatic states arising from 'psychological blows' will not, normally, ground a defence. [168] Conversely, dissociation stemming from physical trauma has been viewed as arising from 'external' sources and as such sufficient for the purposes of the defence.[169] At least part of the reason for the distinction is the intuition that an individual is more likely to be mentally and morally 'active' in the former case than in the latter.[170] Underlying such a 'control'-centred analysis, however, is the idea that the scope of the disease of the mind enquiry is delimited by questions of an accused's future dangerousness.

To turn to the above analysis, the VLRC approach outlined above focuses exclusively on an accused's state of mind at the time of the offence. The law clearly requires that an accused's mental disorder cause the wrong in question.[171] However, such an approach fails to acknowledge that an equally pertinent consideration is that such cognitive incapacity be caused by a 'disease of the mind'. That is, the requisite causal link is not merely from incapacity to behaviour but also, and primarily for the purpose of the defence, from 'disease of the mind' to 'defect of reason'. The preliminary question of what constitutes 'disease of the mind' remains the threshold question.

Mitchell and Robinson approach the issue from the perspective of an accused's prior-fault in becoming intoxicated. Yet, such an account is problematic to the extent to which it suggests derogation from general principles of criminal responsibility. In particular the authors fail to show how an accused's action may be justly ascribed to him or her. Neither on subjectivist nor objectivist criteria is such an ascription permissible. Recklessness as the minimum subjective fault element and negligence as the corresponding objective fault element, require conscious advertence, and the

---

167 Joseph Raz, *Engaging Reason* (1999) 82.

168 *Rabey v R* (1977) 37 CCC (2d) 461; But see, *The Queen v Falconer* (1990) 171 CLR 30.

169 *R v T* (1990) Crim LR 256.

170 See Meir Dan-Cohen, 'Responsibility and the Boundaries of Self' (1992) 105 (5) *Harvard Law Review* 959, 975.

171 *R v Kemp* [1957] QB 399, 407 per Devlin J.

taking of reasonable precautions against a foreseeable risk, respectively. Foreseeability is the minimum determinative factor in both accounts. A subjectivist would insist that an accused foresee the risk undertaken. An objectivist minded theorist, on the other hand, would hold that a reasonable person must have foreseen the unreasonable nature of the risk undertaken by the accused, before liability attaches. Such a degree of foreseeability is not one that may be readily acknowledged in cases of drug-induced psychoses absent recourse to a degree of constructive fault. The contention that an accused is to be held responsible on the basis of drug-use *simpliciter* is not one which stands up to scrutiny.

Finally, both the Meloy, and Scott and Kingswell, approaches have endeavoured to delimit the enquiry by postulating a state of 'settled insanity' as a precondition of an accused raising a defence of mental impairment. I have argued that these authors in suggesting reliance on 'current scientific research' as a means of formulating judicial principles fail to appreciate the true nature of the defence. The defence of mental impairment is not an attempt to ascertain whether an accused is suffering from a mental illness. Rather, the various formulations of the mental impairment defence, both statutory and at common law, represent 'means whereby juries work rough justice in a difficult area of law and morality'.[172]

Fundamentally, such formulations highlight that it is a public policy consideration, and not scientific research, which is determinative of an accused's criminal responsibility. The policy orientation of the enquiry finds expression in, amongst other things, recourse to the distinction between an internal and external state. While less than conceptually sound the 'internal/external' test provides, in the language of the Supreme Court of Canada, an 'analytic tool' in determining the question of whether clinical intervention is required for societal protection.[173] In so doing the focus does not remain on establishing whether an accused's state of cognitive impairment was 'settled' but whether such a state warrants clinical intervention.

These are difficult questions, not because they raise complex, indeterminate clinical issues, but because by their nature they bring into sharp relief the nature of our responsibility practices. The temptation to bypass such difficulties by focusing on overly narrow time frames, or by abandoning fundamental legal principles, or by suggesting that these are empirically verifiable propositions, should be resisted. It is the difficulty of the enterprise, and the goal to be achieved, which makes it worth our while to consider these questions in as clear a light as possible.

---

172 Norval Morris, ' "Wrong" in the M'Naghten Rules' (1953) 16 *Modern Law Review* 435, 437.

173 *R v Stone* (1999) 134 CCC (3d) 353, 439 per Bastarache J.