the sydney law review

contents

articles
The Efficacy of Contractual Provisions Prohibiting Assignment
GJ Tolhurst
161

Trade Marks and Language
Megan Richardson
193

What Explains Variations in Dissent Rates? Time Series Evidence from the High Court
Russell Smyth
221

before the high court
Redundancy and Interpretation in Industrial Agreements: The Amcor Case
Andrew Frazer
241

cases and comments
ACCC v Berbatis Holdings
(2003) 197 ALR 153
Nicole Dean
255

Renvoi Rejected? The Meaning of ‘the lex loci delicti’ after Zhang
Robert Yezerski
273

Shock Rebounds: Tort Reform and Negligently Inflicted Psychiatric Injury (Dyanah)
Nicole Seeto
293
EDITORIAL BOARD

Patrick Parkinson (Editor)  Fleur Johns
Emma Armson  Mary Crock  Les McCrimmon
Saul Fridman  Jenni Millbank

STUDENT EDITORIAL COMMITTEE

Fleur Beaupert  John Paul Hinojosa
Matthew Burston  Stephen Hunyor
Jeremy Liam Burgess  Rowan Mawa
Sevda Clark  Alison Mutton
Samantha Edwards  Brindan Suresh
Ben Everitt  Rebecca Ward
Leah Friedman

Coordinator of the Review: Joanna Howse

Contributions and correspondence should be addressed to:

The Editor
Sydney Law Review
Sydney University Law School
173–175 Phillip Street
SYDNEY NSW 2000
AUSTRALIA
Fax +61 2 9351 0200
Email slr@law.usyd.edu.au or, joanna@law.usyd.edu.au

The Sydney Law Review is a refereed journal.

The Sydney Law Review Online
http://www.law.usyd.edu.au/~slr

Text of Before the High Court Special Issue Feature page
Article and Case Note synopses Indices for Volume 14 onwards
Text of Style Guide Back Issue Order Form
Faculty of Law

Dean
Ronald McCallum, Bjur LLB Monash, LLM Qu, Blake Dawson Waldron Professor of Industrial Law

Pro Dean (Staff Development)
Barbara McDonald, LLM Lond, BA LLB, Associate Professor (Semester 1)
Julie Stubbs, BA W'gong MA Tor, Associate Professor (Semester 2)

Pro Dean (Teaching Programs)
Patricia Loughlan, BA LLM Tor, PhD, Associate Professor (Semester 1)
Donald Rothwell, BA LLB Qld, LLM Alta, MA Calg, PhD, Associate Professor (Semester 2)

Associate Deans
(Undergraduate)
Graeme Coss, LLB, LLM (Hons) Syd, Grad Dip Inf&LibStud Curtin, Senior Lecturer

(Postgraduate)
Jennifer G Hill, BCL Oxf, BA LLB, Professor (Research)
Roger Magnusson, BA, LLB ANU, PhD Melb, Associate Professor (Coursework)

International Students
Bernhard W Boer, BA LLB Melb, Professor of Environmental Law (Personal Chair)
Lee Aitken, BA, LLB ANU, BCL Oxf, Associate Professor
Margaret Allars, DPhil Oxf, BA LLB, Professor
Ross Anderson, LLM Lond, LLB, Senior Lecturer
Patricia Apps, BArch UNSW, MED Yale, PhD Camb, ARAIA, Professor in Public Economics in Law (Personal Chair)
Emma Armson, BEc, LLB Macq, LLM UNSW, Lecturer
Hilary Astor, BTech(Law) PhD Brunel, Abbott Tout Professor of Litigation and Dispute Resolution
Belinda Bennett, BEc LLB Macq, LLM SJD Wisc, Associate Professor
Celeste Black, BA Harv, JD Penn, LLB, Lecturer
Fiona Burns, BA LLB LLM Syd, LLM Camb, PhD ANU, Senior Lecturer
Lee Burns, BComm LLM UNSW, LLM, Sesqui Associate Professor in Taxation Law
Peter Butt, BA LLM Syd, Professor
Terry R Carney, LLB DipCrim Melb, PhD, Monash, Professor
John W Carter, PhD Camb, BA LLB, Professor in Commercial Law (Personal Chair) (fractional)
Judith Cashmore, BA, DipEd, Adel, MED Newcastle, PhD Macquarie, Research, Level B Academic (fractional)
Graeme Cooper, BA, LLB, LLM Syd, LLM Illinois, LLM, JSD Col, Professor of Taxation Law
Mary Crock, BA LLB PhD Melb, Senior Lecturer
Christopher Cunneen, BA DipEd UNSW, MA, Associate Professor
Kristen Daglish, BA, Macq, LLB(Hons) Macq, MIL Syd, Associate Lecturer (fractional)
Bernard Dunne, BA LLB Macq, Lecturer
Mark J Findlay, BA LLB ANU, DipCrim MSc Edin, LLM, Professor
Nicola E Franklin, BA Natal, LLB Natal & Camb, DipComparLegalStud Camb, Senior Lecturer
Saul Fridman, LLB W Ontario, BCL Oxf, Senior Lecturer
Jennifer Gage, BEc, LLB Syd, Lecturer (fractional)
Peter Gerangelos, BA, LLB Syd, LLM UNSW, Lecturer
Reg Graycar, LLB Adel, LLC Harv, Professor of Law
Helen Irving, BA Melb, Etudes Libres (License) Lausanne, MPhil Camb, PhD, Associate Professor
Fleur Johns, B A, LLB Melb, LLM Harv, Lecturer
Isabel Karpin, LLM Harv, SJD Col, BA LLB, Senior Lecturer
Abigail Kohn, BA Harv, MPhil (Criminology) Camb, PhD (Medical Anthropology) Calif, U2000 Postdoctoral Fellow, Leval A Academic
Ghena Krayem, BA, LLB Syd, Associate Lecturer (fractional)
Rosemary Lyster, BA LLB LLM Natal, Senior Lecturer
Ghena Krayem, BA, LLB Syd, Associate Lecturer (fractional)
Les McRittenon, BA LLB Alta, LLM Qld, Associate Professor
Rebecca Millar, BSc Auck, LLB UNSW, Senior Lecturer in Taxation
Jenni Millbank, LLM Br Com, BA LLB, Senior Lecturer
Luke Nottage, BCA, LLB VUW, LLM Kyoto, Senior Lecturer
Patrick N Parkinson, MA Oxf, LLM Ill, Professor
Les McRittenon, BA LLB Alta, LLM Qld, Associate Professor
Rebecca Millar, BSc Auck, LLB UNSW, Senior Lecturer in Taxation
Jenni Millbank, LLM Br Com, BA LLB, Senior Lecturer
Luke Nottage, BCA, LLB VUW, LLM Kyoto, Senior Lecturer
Patrick N Parkinson, MA Oxf, LLM Ill, Professor
Elizabeth Raper, BA, LLB Syd, Lecturer to 30 June 2004
Joellen Riley, MA LLB Syd, BCL Oxf, Senior Lecturer
Alpana Roy, BA UTS, LLM(Hons) Syd, MA UTS, Associate Lecturer (fractional)
Kristen Rundle, BA, LLB Syd, LLM McGill, Lecturer
Wojciech Sadurski, DrJur Warsaw, Professor of Legal Philosophy (Personal Chair)
Kristin Savell, BSc, LLM Macq, LLM McGill, PhD Camb, Lecturer
Diane Skapinker, BComm LLM Wits, LLM Associate Professor (fractional)
Belinda Smith, BEc, LLB Syd, Lecturer
Tim Stephens, BA(Hons) LLB(Hons) Syd, Associate Lecturer (fractional)
Greg Tolhurst, DipLaw SAB, LLM, Lecturer
George Tomossy, BSc, Toronto, LLM McGill, Associate Lecturer (fractional)
Richard J Vann, BA LLB Qld, BCL Oxf, Professor
Andrew Tuch, BCom, LLB Qld, LLM Harv, Lecturer
Kevin Walton, MA Lond, LLB Edin, Lecturer
Jane Wangmann, BA, LLB UNSW, Research, Level A Academic (fractional)
David Weisbrot, BA CUNY, JD Calif, Professor
Brett Williams, LLB BEc Adel, Lecturer
George Winterton, LLM, WA, JSD Columbia, Professor of Constitutional Law
K Alex Ziegert, DPhil Min, Associate Professor

Challis Lecturers
The Hon Justice Arthur R Emmett, BA LLM (Challis Lecturer in Roman Law)
The Hon Justice D Graham Hill, LLM Harv, BA LLB (Challis Lecturer in Taxation) (Stamp Duties)
The Hon Justice Roderick P Meagher, BA LLB LLD (Challis Lecturer in Equity)

Lecturers (Part-time)
The Hon Justice Robert Austin, BA LLB LLM, DPhil Oxf (Takeovers and Reconstruction)
Gerry Bates, LLB, PhD Birm, Senior Lecturer (Environmental law)
Christopher Birch, BA LLB PhD (Aspects of Legal Reasoning) (S1 & S2)
David J Harland, BCL Oxf BA LLB (Consumer Protection)
Bron A McKillop, LLM Harv; BA LLB BEc (Criminal Procedures – intensive)
Anne McNaughton, BA N’cle, LLM UNSW, LLM Tübingen, LLM UTS (Legal System of the European Law (S1)
Peter B Shea, BHA GradDip(HealthAdmin) UNSW, DipEnvironStud Macq, MB BS MPH DPM DipCrim Syd, FRANZCP, FRACMA LHA AFAIM (Challis Lecturer in Forensic Psychiatry) (Forensic Psychiatry)
Brian Preston, BA LLB Macq (Biodiversity Law)
Robin H Woellner, BA LLB (Taxation Administration)
The Efficacy of Contractual Provisions Prohibiting Assignment

GJ TOLHURST*

Abstract

In recent years, a line of decisions has recognised the legitimacy of prohibitions on assignment. Moreover, and despite holding that the effect of such prohibitions depends on construction, generally, it would appear that whatever words are used, the effect of a prohibition will be to render the attempted assignment ineffective. Nevertheless, it has never been fully explained how such provisions operate at a doctrinal level, that is, how a contractual provision inhibits a person dealing with their right of property. This paper suggests that the explanation lies with the ability of contracting parties to fashion the characteristics of the contractual rights they bring into existence not just in their character as contractual rights but also as choses in action.

* Faculty of Law, University of Sydney, Associate Editor, Journal of Contract Law, Editor, Commercial Law Quarterly.
1. Introduction

The assignment of contractual rights is an important but difficult area of the law. It forms a sub-group of the assignment of choses in action. The idea behind an assignment of a chose in action is simple: namely, a chose in action forms part of the property of the person in whom it is vested and that person may transfer it to another. The word ‘assign’ in this context simply means to transfer or convey. The person transferring the chose is the assignor and the person to whom it is transferred is the assignee. In the case of a contractual right, the other party to the contract with the assignor, who must now account to the assignee, is either the debtor in the case of an assigned contractual debt, or the obligor in the case of an assignment of some other right to performance.

Unfortunately, in the case of contractual rights, this is where simplicity ends. The efficacy of an assignment of a contractual right is governed by numerous rules which have never been the subject of any detailed analysis. These rules include, the assignee is to be no better off than the assignor, it is only possible to assign rights and not duties and personal contractual rights may not be assigned. It is not the purpose of this paper to set out, explain or resolve all the problems that exist within the rules governing the assignment of contractual rights. The paper has a narrow focus: namely, given that the current Anglo-Australian model of assignment is property-based, that is, it is concerned with transferring rights of property, to what extent can the parties to the contract, that is, the assignor and obligor, control the transferability of a right? It has been held that an express contractual provision may allow for the assignment of a right which otherwise may have been construed as being personal and not assignable.\(^1\) This is an application of the rule that the personal nature of a contractual right is a question of construction. Usually, however, contractual provisions attempt to limit or prohibit assignability.\(^2\) What is the effect of an attempted assignment in the face of such a prohibition? There are various possible answers: for example, such prohibitions may be void for being against public policy, they may be valid and the attempted assignment is ineffective or the assignment may be effective but place the assignor in breach of contract with the obligor. Surprisingly, it is only in recent years that courts have been given the opportunity to address this issue. The current weight of Anglo-Australian authority has recognised that such prohibitions are not against public policy and their effect is to render any assignment ineffective. Despite this authority, it is not clear how such prohibitions operate from a doctrinal perspective.

In this paper, it will be suggested that the answer to this issue of doctrinal efficacy depends on whether such provisions merely operate as a matter of contract or whether they impact on the nature of the relevant contractual right as a chose in action. Once an answer to this question is formulated, it is then possible to

---

1 See Devefi Pty Ltd v Mateffy Pearl Nagy Pty Ltd (1993) 113 ALR 225 at 235.
2 For a discussion of the reasons why parties may want to prohibit or promote assignment; see Andrew McKnight, ‘Contractual Restrictions on a Creditor’s Right to Alienate Debts: Part 1’ (2003) 18 JIBL 1 at 3; Richard Epstein “Why Restrain Alienation?” (1985) 85 Columbia L Rev 970 at 982.
investigate two further problematic areas, namely: what the rights of the various parties in the case of a breach of a prohibition are and what the relationship between such provisions and the rule preventing the assignment of personal contractual rights is. In addition, a related issue, namely, who can claim the benefit of such a provision, is also discussed. The paper begins by looking at some of the more fundamental issues relating to the efficacy of such provisions.

2. Fundamental Efficacy Issues

Because this is an area of law that has only recently come under the spotlight in Anglo-Australian law, it is necessary to make a number of preliminary points for the purpose of setting the scene before moving on to the discussion of the doctrinal efficacy of prohibitions on assignment.

A. The Position in England

In England, pursuant to the decision of the House of Lords in *Linden Gardens Trust Ltd v Lenesta Sludge Disposals Ltd*, full effect will be given to contractual provisions prohibiting assignment. That is, any purported assignment will be ineffective.

B. The Position in Australia

In Australia, a majority of the High Court in *Associated Alloys Pty Ltd v ACN 001 452 106 Pty Ltd*, approved the analysis of the effect of a breach of such a prohibition contained in *Don King Productions Inc v Warren* and *Devefi Pty Ltd v Mateffy Pearl Nagy Pty Ltd*. *Don King Productions* concerned the application of certain aspects of the *Linden Gardens* case and the reference to it may suggest a broader acceptance of the approach adopted in *Linden Gardens*. In *Devefi* too, the analysis adopted by the Full Federal Court, although handed down prior to *Linden Gardens*, is in line with *Linden Gardens* at least to the extent of recognising the validity of such clauses. In addition, there are now a number of decisions of Australian courts that have expressly adopted the reasoning in *Linden Gardens* and it is reasonably safe to conclude that until the High Court has an opportunity to

---

3 Although not discussed in this paper, generally, a contractual prohibition on assignment cannot stop the relevant property of a bankrupt vesting in the trustee in bankruptcy; see *Krasner v Dennison* [2001] Ch 76 at 99 (Chadwick LJ). See further Grover Grismore, ‘The Effect of a Restriction on Assignment in a Contract’ (1933) 31 Michigan L Rev 299 at 311 and see *Norglen Ltd v Reeds Rains Prudential Ltd* [1999] 2 AC 1 and *Gathercole v Smith* (1881) 17 Ch D 1.

4 *Linden Gardens Trust Ltd v Lenesta Sludge Disposals Ltd* [1994] 1 AC 85.

5 *Associated Alloys Pty Ltd v ACN 001 452 106 Pty Ltd* (2000) 202 CLR 588 at 601. See also *Broadcast Australia v Australia Pty Ltd v Minister Assisting the Minister for Natural Resources (Lands)* (2004) 204 ALR 46 at 51 (referring to the decision in *Linden Gardens* with apparent approval). See further *Bruce v Tyley* (1916) 21 CLR 277 at 292 (Isaacs J) (approval of prohibition on sub-letting a contract).


7 *Devefi*, above n1 at 234–7.

8 Compare page 183.
look at the issue in detail, the reasoning of the House of Lords will be applied by
Australian courts.  

There is, however, one decision of the High Court of Australia which appears
to stand in the way of adopting the approach taken by the House of Lords in *Linden
Gardens* and which requires further analysis before it can be said without question
that such provisions are efficacious in Australia as a matter of law. That case is
*Hall v Busst*, 9 and, in particular, the statement made by Dixon CJ to the effect that
a contractual restraint on alienation can only give rise to a right to damages. In
*Caboche v Ramsey*, 10 Gummow J said that Dixon CJ’s statement was restricted to
restraints involving interests in land and where ‘the restraint is imposed upon
dealings with debts or other choses in action, or with the benefit of performance of
contracts involving personal skill and confidence, further, and more detailed,
principles apply both as to the interpretation and efficacy of such provisions.’ For
this proposition he referred to the decisions in *Devefi* and *Linden Gardens*. The
court in *Devefi* suggested that *Hall v Busst* represented a ‘distinct body of principle
[dealing] with restraints upon alienation of land’. 12 However, earlier Needham J,
in *Reuthlinger v MacDonald*, 13 said that the doctrine of restraints on alienation
applied to both reality and personality and, moreover, the doctrine applied to mere
contractual restraints because the majority of the High Court in *Hall v Busst* did
not limit it to restraints that are conditions of a grant. 15

*Hall v Busst* concerned the sale of an island. On the day the sale contract was
entered into (but executed after the sale contract), the parties entered into a
separate indenture in which the grantor/purchaser agreed not to transfer, assign or
lease any part of the island without the consent of the grantee/vendor. In due
course, the purchaser agreed to sell the land to a third party and that third party


10 *Hall v Busst* (1960) 104 CLR 206 at 217. Earlier in *The Commissioner of Stamp Duties New South Wales v Yeend* (1929) 43 CLR 235, a case concerning a stamp duty issue, the fact the contract contained a prohibition on assignment was not even mentioned by the Court.

11 *Caboche*, above n9 at 232.

12 *Devefi*, above n1 at 236.

13 *Reuthlinger v MacDonald* [1976] 1 NSWLR 88 (affirmed in *Reuthlinger v MacDonald* (NSW Court of Appeal, Needham J, 20 October 1976)).

14 Id (NSWSC) at 97. Compare *Caboche*, above n9 at 226–28, 231–32 where Gummow J accepts that the restraints on alienation doctrine applies to both reality and personality but distinguishes it from contractual restraints relating to debts and choses in action.

15 Id (NSWSC) at 99–100.
entered into possession. The vendor brought an action for damages for breach of contract. That action failed in a majority decision of the High Court. Dixon CJ began by stating the general principle that where a restraint is made a condition of a grant it is void on the ground of repugnancy. However, to this he added the alternative ground of public policy. This case, however, was different because the restraint was contained in a separate contract and imposed by way of covenant. Dixon CJ posed the question whether such a restraint should be treated differently as there is 'no fetter upon alienation which does more than sound in damages, that is, unless a doctrine of equity intervenes to make it bind the land.' In the result he held that the provision in question should not be treated differently because it was nevertheless aimed at securing an indefinite, unqualified prohibition on alienation without consent. It was, in short, as effective as a condition subsequent to the grant would have been. However, he limited the ground upon which it was invalid to that of public policy. Fullagar J agreed but specifically said that the provision was repugnant; he did not expressly add the alternative public policy ground. Menzies J, the last member of the majority, agreed with the reasons given by both Dixon CJ and Fullagar J but added that he thought the provision was as contrary to the public interest as a restraint of marriage.

From the facts of the decision itself, whether or not the view is taken that Hall v Busst must be applied to personalty, it is possible to distinguish it from the prohibitions being dealt with in this paper. As was pointed out by Needham J in Reuthlinger v MacDonald, it appears clear that in Hall v Busst the restraint was not operative until the transfer of the land to the purchaser was completed. There is no doubt that it is repugnant or otherwise against public policy for a prohibition on assignment to attempt to deal with the fruits of a contract once those fruits are in the hands of the assignor. But this (and Hall v Busst in terms of its ratio) says nothing about the situation in which the prohibition seeks to prevent the assignment of existing contractual rights where the enjoyment of those rights is contingent on some event or is conditional upon the assignor performing some obligation, and where the right to some performance has unconditionally accrued but the time for performance has not yet arrived.

---

16 The minority (Kitto & Windeyer JJ) construed the agreement as only granting a right of pre-emption.
17 Hall, above n10 at 217.
18 Ibid.
19 Id at 218. Compare Caldy Manor Estate Ltd v Farrell [1974] 1 WLR 1303 (in which the court distinguished a covenant whereby a right of re-entry is provided for upon breach of the covenant so as to destroy the grant from a provision that would not destroy the transfer to the third party but make the covenantor liable to the covenantee in damages; the doctrine of repugnancy applied to the former). For an example of where a restraint was upheld because it was not so indefinite and imposed for a valid collateral purpose; see Reuthlinger (NSWSC), above n13 (affirmed in Reuthlinger (NSWCA), above n13 and see Elton v Cavill (No 2) (1993) 34 NSWLR 289.
20 Hall, above n10 at 223, 225.
21 Reuthlinger (NSWSC), above n13 at 100 (affirmed in Reuthlinger (NSWCA), above n13.
22 The situation is different where the transferor retains title in the subject matter of the contract by way of security or obtains an interest in the subject matter under some other contract. This is more likely to occur, though, in dealings with tangible property.
C. The Restraints on Alienation Doctrine

The above analysis still leaves the issue as to whether upholding such provisions is at odds with the principle prohibiting restraints on alienation. If it is, then such provisions can have one of two effects: either they are completely without effect or, despite an assignment in the face of such a provision being valid, the promise not to assign may still be recognised, thus putting the assignor in breach of contract with the obligor.

The argument that prohibitions on the assignment of contractual rights constitute an improper restraint on alienation was raised and rejected by the House of Lords in Linden Gardens Trust Ltd v Lenesta Sludge Disposals Ltd. Lord Browne-Wilkinson rejected the argument on the basis of both authority and principle. In terms of authority, he simply referred to the line of cases upholding such clauses. As to principle, he took the view that the restraints doctrine was limited to land because it is a finite resource. He thought that there was no public policy reason overriding a prohibition on the alienation of tangible property and no public need for a market in choses in action.

Lord Browne-Wilkinson’s latter statement is problematic. Firstly, as already noted, it would appear that the restraints doctrine is not necessarily limited to interests in land. Moreover, market economies need credit, and there is a very real need for a market in debts. This market is not necessarily limited to debts in the strict sense of unconditional accrued rights to payment but also encompasses conditionally-accrued rights to payment. Thus, international conventions dealing with factoring and receivables financing have express provisions overriding any such prohibitions. In addition, whatever may have been the historical factual scenario for the operation of the restraints on alienation rule, it is clear that part of the thinking behind these Convention provisions is that such prohibitions would be at odds with the principle against restraints on alienation. Likewise, but to a lesser extent, the restraints rule was one factor influencing the insertion of provisions in the Restatement of Contracts and Uniform Commercial Code to restrict the effect of such prohibitions. Admittedly, however, it may be that in these instances, although the restraints rule was referred to, what was really in mind was some broader notions of public policy than the more narrow repugnancy principle that has historically underpinned the restraints doctrine.

---

23 Linden Gardens, above n4 at 106–7.
24 See n76 below. There is a long history of such clauses being upheld in the context of hire-purchase agreements; see, for example, Wickham Holdings Ltd v Brooke House Motors Ltd [1967] 1 WLR 295.
25 Linden Gardens, above n4 at 107.
26 See Caboche, above n9 at 226, 232 (Gummow J).
27 See, for example, UNIDROIT Convention on International Factoring, Article 6; United Nations Convention on Assignment of Receivables in International Finance, Article 11.
28 Section 322 and note official comment (a).
29 Sections 2–210, 9–406.
Perhaps the most powerful statement that the restraints on alienation rule is irrelevant was that of Untermyer J in the New York Court of Appeal in *Sacks v Neptune Meter Co.* He said:

We think both upon principle and upon authority... a covenant against assignment, which in substance provides that the obligation shall be unenforceable in the hands of the assignee, is available to the obligor as a defense... In so holding we do not sanction unlimited restraint upon the alienation of property within the accurate definition of that term. There is a perceptible distinction between the right of a contracting party to impose conditions upon the exercise of a contractual right and the imposition on the owner in fee of undue restraint in respect to the alienation of his property. A grantor cannot transfer complete ownership of tangible property and still control its devolution, because such control is repugnant to the absolute character of the grant... But that rule does not apply where the restraint is upon the alienation of an estate for years and the grantor has received a reversionary interest in the property. Even more conspicuously would the rule seem to be inapplicable where no transfer of title has occurred and the restraint is only of contractual rights.

I have been unable to discover that the rule against restraints on alienation has ever been applied to choses in action... Where the subject-matter is a chose in action, neither public policy nor consistency requires that it be enforceable against the promisor except in accordance with the terms on which his promise was made. The limitation is not so much imposed on the obligee's right of alienation as on the obligor's duty to perform. The restraint then becomes a condition of acquisition.

Despite approval in *Devefi,* the last two sentences in this citation are problematic and are discussed further later in this paper. However, at the end of the first paragraph Untermyer J does point out an important factual distinction. Arguably, the focus of the repugnancy principle is the situation where A absolutely transfers property to B and in that transfer there is a condition aimed at preventing B from transferring the property to a third party. Here, after the transfer to B, A drops out of the picture (it being an absolute transfer) and there are therefore good reasons why A should have no say in what B does with the property. In the case of an assignment of a contractual right, the person in the position of A, that is, the

---

30 The ‘repugnancy’ principle is not without its critics and it has been argued that the law should acknowledge that public policy underlies it: see Glanville Williams, ‘The Doctrine of Repugnancy’ (1943) 59 LQR 343, (1944) 60 LQR 69, 190; Merrill Schnebly, ‘Restrains Upon the Alienation of Legal Interests’ (1935) 44 Yale LJ 961, 1186, 1380. See further Ken Mackie, ‘Contractual Restraints on Alienation’ (1998) 12 JCL 255. See also Hall, above n10 at 217 (Dixon CJ) and *Nullagine Investments Pty Ltd v The Western Australian Club Inc* (1993) 177 CLR 635 at 649 (Brennan J).

31 *Sacks v Neptune Meter Co* 258 NYS 254 (1932).

32 Id at 261–62.

33 *Devefi,* above n1 at 234–7.

34 The doctrine does not apply where what is given is a determinable interest such as a gift ‘to A until he alienates and then to B’, as opposed to a gift to ‘A but if he alienates then to B'; see *Caboche,* above n9 at 227 (Gummow J).
obligor, remains in the picture and therefore does have a vested interest in the identity of the person for whom it performs. Moreover, in the case of a contract, it is only at a very theoretical level that it can be said that formation occurs by one party transferring its promise to the other party.\textsuperscript{35} At a doctrinal level, such a view could have undesirable results.\textsuperscript{36} It follows that there is no initial grant to the assignor which is made the subject of a condition restraining alienation. The only time such a transfer occurs is where one party actually receives the fruits of the contract. Prior to this point, as between the parties to the contract, there only exist personal rights and obligations, albeit that those rights may be considered choses in action for the purposes of transfer. Therefore, the subject matter being dealt with is not the same as that to which the repugnancy principle applies.

Before moving on, it is still necessary to determine whether the decision in \textit{Hall v Busst} is nevertheless at odds with the above argument. It will be recalled that in \textit{Reuthlinger v MacDonald}, Needham J suggested that the High Court in \textit{Hall v Busst} did not limit the restraints rule to restraints that are conditional upon a grant but extended it to restraints imposed by way of covenant. Ultimately, however, the most authoritative interpretation of \textit{Hall v Busst} is that of Brennan J in \textit{Nullagine Investments Pty Ltd v The Western Australian Club Inc},\textsuperscript{37} viz:

\begin{quote}
In this country a court may hold invalid restraints on alienation imposed not only by conditions annexed to a gift or grant of an estate in land but also by covenants and agreements. But the grounds on which a condition subsequent to the gift or grant of an estate of freehold may be held invalid are not necessarily the same as the grounds on which invalidity strikes a covenant or agreement in an instrument which does not itself give or grant the estate the alienation of which is restrained. In the former case, as Dixon CJ said in \textit{Hall v Busst}, “[t]he invalidity may be put on the ground of repugnancy to the grant or upon public policy or for that matter it may conceivably be attributed to an indirect effect of \textit{Quia Emptores}.” In the latter case, the only basis for holding the covenant or agreement invalid is public policy.
\end{quote}

From this we are only left with the argument that, although not repugnant, such provisions may be against public policy. That is, although a contract creates rights between the parties and these rights may be choses in action, those rights are not created by gift, grant or transfer and the issue falls within the ‘latter case’.


\textsuperscript{36} See Yeend, above n10 at 241.

\textsuperscript{37} \textit{Nullagine Investments Pty Ltd}, above n30 at 649 (Brennan J).
D. Public Policy Limitations on Prohibitions

Turning then to public policy, from what was said earlier, there would appear to be an argument that such prohibitions should not be upheld when they relate to debts or at least certain types of debts.\(^{38}\) It is also no doubt at odds with public policy to uphold a provision that seeks to control the power of a promisee to deal with the fruits of a contract once they are in the hands of that promisee.\(^{39}\)

More problematic, in terms of public policy, is the efficacy of a clause that seeks to prohibit the assignment of accrued rights under a contract, that is, where at the time of the assignment the obligation to perform has accrued (conditionally or unconditionally) but where the performance has not been tendered to the assignor.\(^{40}\) The situation may be that the assignor has paid the full contract price and now wishes to assign the right to counter-performance. In some cases, the actual assignment may be entered into before the counter-performance has accrued or been earned. Nonetheless, it is arguable that the prohibition should have no effect once the counter-performance (the right to which is the subject of the assignment) has been earned.\(^{41}\) However, perhaps the better view is that in such a case the prohibition should not be effective if its only purpose was to secure that performance which has now been received. Moreover, clearly, where the obligation to perform involves a payment of money and where that obligation unconditionally accrues, the right to payment is then a debt and any prohibition on the assignment of that debt would be rendered ineffective by reason of the public

---

\(^{38}\) There may also be an argument that public policy should inhibit the efficacy of such clauses (whether relating to debts or not) when the obligor has no legitimate reason for preventing the assignment. If this is the case, then it should be similarly applied to the personal rights rule and perhaps determined by the rules governing restraints of trade – but query the relevance of that doctrine when the right would not exist but for the agreement of the parties and the assignor is not, by agreeing to the prohibition, giving up some pre-existing freedom. See Ockaldale (Richmond) Ltd v National Westminster Bank plc [1997] 1 BCLC 63; Australian Rugby Union Ltd v Hospitality Group Pty Ltd (2000) 173 ALR 702 (affirmed [2001] FCA 1040). At present, no such rule of public policy is applied to provisions which make contractual rights personal.

\(^{39}\) Re Turner Corporation Ltd, above n9 at 767. See further Sir Roy Geode, ‘Inalienable Rights’ (1979) 42 MLR 553. In Linden Gardens, above n4 at 107 Lord Browne-Wilkinson expressed no concluded view on this point, but his statement to the effect that there are no public policy arguments against restrictions on the transfer of tangible property other than land would appear to be against it.

\(^{40}\) In the case of a provision that seeks to forfeit accrued rights upon any attempted assignment, it may be that such a provision would attract the jurisdiction of equity to relieve against forfeiture; see Devey, above n1 at 237. But query the form such relief may take. In some cases, perhaps it could be by way of specific performance whilst still invalidating the attempted assignment. In others, it could be by way of restitution of the payments made which earned the accrued right which is the subject of the forfeiture. However, it is perhaps more problematic where the performance earning the accrued right and which is the subject of the forfeiture is something other than the payment of money. In any case, there is the problem that what is being forfeited are contractual rights which, as between the parties are personal rights and arguably not property for the purposes of relief against forfeiture; see further Bysouth v Shire of Blackburn and Mitcham (No 2) [1928] VLR 562 at 573–75 and compare Hodder & Tolley Ltd v Cornes [1923] NZLR 876. Relief therefore would need to be based on some other head of jurisdiction, particularly if the contract had been terminated; see generally Tanwar Enterprises Pty Ltd v Cauchi (2003) 201 ALR 359.
policy limitations already outlined. Outside of these cases, there may be good reasons for the obligor still wanting to prohibit the assignment even where the assignor has a conditional or unconditional right to the obligor’s counter performance.42

Beyond the above points, it is difficult to come to any firm conclusions. However, if the view were adopted that such clauses are completely at odds with public policy and therefore ineffective, this would appear to be inconsistent with the position that personal contractual rights are not assignable when it is clear law that the distinction between personal and non-personal contractual rights depends on the intention of the parties.43 That is, Anglo-Australian law has not adopted a position that a right should be considered assignable despite the express personalisation of the right by the parties if there is either no material detriment to the obligor or if no reasonable person in the position of the obligor would object to the assignment. Therefore, if prohibitions are not enforceable, this rule could be circumvented by simply stating in the contract that the right in question is personal. Perhaps the better view is that the great weight of public policy limitations relate to debts, perhaps broadly defined,44 rather than other rights to performance under a contract.45 In this way it may be possible to maintain consistency with the personal rights rule which generally applies to obligations to perform acts other than the payment of money.46 Moreover, it may be that in Linden Gardens, Lord Browne-Wilkinson intended his comment that there was no market in choses in action to be limited to the type of transaction before him, namely, an assignment of a right to performance.

3. **Construction and the Efficacy of Prohibitions**

It needs to be kept in mind that an effective prohibition can at most merely render the assignment invalid as between the obligor and assignee. There is still a valid agreement between the assignor and assignee. Where the assignor cannot perform

41 See generally Grismore, above n3 at 306–7.
42 See Goode, above n39 at 555 and see Linden Gardens, above n4 at 105–6 (Lord Browne-Wilkinson). The same reasons may not prevent a declaration of trust of such accrued rights because a trust does not generally provide the beneficiary with the same direct rights that an assignee has against the obligor. See also Restatement of Contracts 2d §322(2)(a) which states that unless a different intention is indicated a prohibition does not prohibit the assignment of a right to damages for breach of the whole contract or the assignment of contractual rights arising from the assignor’s complete performance of the contract.
44 For example, under the United Nations Convention on Assignment of Receivables in International Trade, an amount is fully earned when an invoice is issued; see UNCITRAL Analytical Commentary on the Draft Convention on Assignment of Receivables in International Trade, A/CN.9/489/Add 1, 22 May 2001 at para 27.
that agreement it will be in breach of contract.\textsuperscript{47} Moreover, a prohibition against assignment may not prevent the assignor from declaring itself a trustee of the benefit of the contract.\textsuperscript{48}

In addition, prior to determining the legal effect of a prohibition, it is necessary to determine its extent. This is part of the process of determining its meaning. For example, on construction a prohibition may only prohibit legal assignments, thus still allowing an equitable assignment.\textsuperscript{49} However, unless expressly stated, such a construction is likely to be rare.\textsuperscript{50} In practice an important issue has been whether the prohibition is directed only to future rights and not accrued rights,\textsuperscript{51} or perhaps only prohibits outright assignments and not assignments by way of security.\textsuperscript{52} In some instances a prohibition against assignment may, on construction, only prohibit subcontracting.\textsuperscript{53} Another possibility was suggested in \textit{British Gas Trading Ltd v Eastern Electricity plc},\textsuperscript{54} where Leggatt LJ said of a clause which stated that ‘neither party shall transfer or assign its rights or obligations hereunder without the prior approval of the other party’, that the concern of the clause was not with assignment but with novation.\textsuperscript{55} However, ultimately the meaning of such a clause is a question of construction and the point being made here is simply, if, in terms of meaning, the clause does not extend to the transaction in question there is no need to go on to look at its legal effect.

Another aspect of the meaning of a prohibition concerns its form. Prohibitions against assignment come in three principal forms and more than one may be used in any particular contract. Firstly, the parties may promise that they will not assign

\begin{footnote}
\textsuperscript{45} However, many jurisdictions have statutory provisions prohibiting the assignment of certain unearned wages. Thus, even if a broad view is taken of ‘debts’, there will still exist examples of ‘debts’ that cannot be assigned or which may be the subject of a legitimate prohibition on assignment. There are no doubt other examples of unconditional obligations to pay, whether liquidated or unliquidated, and which may or may not constitute debts, where there exist legitimate reasons for prohibiting the assignment of the right to such payments, for example, bank guarantees; compare ICC Uniform Rules for Demand Guarantees Article 4 with the United Nations Convention on Independent Guarantees and Stand-by Letters of Credit Article 10.
Therefore, the public policy considerations that limit the effectiveness of prohibitions seeking to stop the assignment of debts are probably limited to those trade debts that are used in receivables financing and perhaps debts used in other financing structures such as securitisation.
Statutory provisions may also prohibit the assignment of other rights such as certain tenancies; see \textit{Burton v Camden London Borough Council} [2000] 2 AC 399.

\textsuperscript{46} One public policy consideration that may override a prohibition dealing with the assignability of a right to performance is that if the right to performance can in any way be used as security it may be that effect should not be given to the prohibition to the extent that it would prevent a party using that right as security. There appears to be no authoritative decision on this point. However, the current authorities appear to treat this issue as simply one of construction. Thus, a clause prohibiting assignment is likely to also prohibit a mortgage but perhaps less likely to inhibit charging the right. The current cases are collected and discussed in Angela Flannery, \textit{‘Security Over Contractual Rights and Tripartite Agreements’} (2002) 13 \textit{JBFLP} 179. Anyone wishing to advance this argument would need to keep in mind that a security may need to be enforced and this could result in an absolute assignment of the subject right which would defeat the clause. Moreover, an obliger may have good reasons why it does not want the assignor to use the subject right as security, for example, the obligor may not be set up to deal with such security.
\end{footnote}
their rights under the contract (or only assign them to a limited group).\textsuperscript{56} This is usually expressed as the parties ‘shall not assign’ or the parties ‘shall not assign without the consent of the other party’ or ‘this contract is not to be assigned’.\textsuperscript{57} Secondly, a prohibition may be drafted to negate the power to assign. These may be expressed in terms that the parties ‘cannot assign’ or contract rights are ‘not assignable’ or ‘any assignment is void’ or the parties ‘shall not be entitled to assign’.\textsuperscript{58} Clearly there may be some fine lines between these first two groups and often a provision may combine them.\textsuperscript{59} Moreover, even if the words of promise are used, when read in the context of the contract as a whole the provision may amount to a prohibition.\textsuperscript{60} Third, a provision may simply state that in the event of any assignment the other party has a right to ‘terminate’ the contract. These may be coupled with a forfeiture provision. Where the assigned right is not an unconditionally accrued right, such a clause effectively gives power to the innocent party to stop the assignment by electing to terminate the contract. The forfeiture provision seeks to deal with any unconditionally accrued rights. Such provisions may be drafted in a form that suggests that contractual rights are automatically forfeited upon assignment.\textsuperscript{61} However, it is suggested, that the better construction is that the assignment merely gives the obligor a right to elect to terminate the contract (or to elect to forfeit), otherwise, the law would have provided the assignor (in the case of a termination) with a method of getting out of the contract by taking advantage of its own wrong.\textsuperscript{62}

\textsuperscript{47} However, if the prohibition in question is one that is unenforceable as being against public policy, it would be necessary to have a rule that also prevents the attempted assignment being a repudiation of the contract because if it were a repudiation then the debtor will have an opportunity to terminate the contract and thus defeat the assignment to the extent that the assignment concerned rights that had only conditionally accrued prior to discharge and thus also defeat the public policy that sought to allow for the assignment.

\textsuperscript{48} Even where the contract involves personal, skill or confidence, it has been said that a party may declare itself a trustee of the proceeds for the benefit of a third party; see Don King Productions, above n6. See also Re Turcan (1888) 40 Ch D 5. See further Spellman v Spellman [1961] 1 WLR 921 at 925 (Danckwerts LJ) compare at 928 (Willmer LJ); Pincott v Moorstons Ltd [1937] 1 All ER 513 at 516; Swift v Dairywise Farms Ltd [2000] 1 WLR 1177 at 1184. Ultimately, however, this is an issue of construction. The word ‘assignment’ is a word of wide import and can, in any particular context, be used to refer to any form of alienation; see Norman v FCT (1963) 109 CLR 9 at 26 (Windeyer J). A prohibition that seeks to prevent such a declaration of trust may also be against public policy if it seeks to control the subject matter of the contract when that subject matter is in the hands of the trustee. Given the limited recourse a beneficiary has against an obligor, the public policy restrictions on such prohibitions may be wider than in the case of prohibitions on assignment.

\textsuperscript{49} Spellman, above n48 at 925 (Danckwerts LJ), compare at 928 (Willmer LJ).

\textsuperscript{50} See R v Chester and North Wales Legal Aid Area Office (No 12) [1998] 1 WLR 1406. Anyone taking the view that an equitable assignment only operates between the assignor and assignee is, however, more likely to find such a construction; see Long Leys Co Pty Ltd v Silkdale Pty Ltd (1991) 5 BPR 11,512 at 11,518 (Sheller JA) and compare Gregory Tolhurst, ‘Equitable Assignment of Legal Rights: A Resolution to a Conundrum’ (2002) 118 LQR 98. The position as regards leases is different as prohibitions on the assignment in leases have a history of being restrictively construed; see Peter Butt, Land Law (4th ed, 2001) at para 1592 and most recently Old Papa’s Franchise Systems Pty Ltd v Camisa Nominees Pty Ltd [2003] WASCA 11.
Finally, a prohibition may not wholly prohibit assignment but merely require the consent of the obligor.\textsuperscript{63} Usually such consent is expressed in terms that it is not to be unreasonably withheld\textsuperscript{64} or such a condition may be specified by legislation in some instances. Generally, where the clause states that there is to be no assignment without consent with no reference to consent not being unreasonably withheld, the effect of the attempted assignment without consent will be the same as if the clause was a simple prohibition on assignment.\textsuperscript{65} There may be an exception if, on construction, it may be said that the obligor’s intention was to allow assignment subject to the right in the obligor to later elect to refuse to recognise the assignment. That is, where consent is meant to operate as a condition subsequent to the assignment rather than a condition precedent.\textsuperscript{66} Where consent is not to be unreasonably withheld and the assignor fails to seek consent in circumstances where the obligor could reasonably withhold that consent, the assignment will be ineffective as between the assignee and obligor. Where the assignor fails to seek consent in circumstances where the obligor could not have reasonably withheld that consent, it was suggested by Evans LJ in Hendry v Chartsearch Ltd,\textsuperscript{67} that the assignment may be effective. He distinguished the position with respect to leases of land where the view is taken that consent cannot be considered to be withheld if it is not asked for.\textsuperscript{68} Millett LJ on the other hand was prepared to follow the line of authority dealing with leases although he distinguished a commercial contract from a lease on the basis that as between the parties such a contract is not generally viewed as property but obligation. In his view it was ‘wrong in principle to entertain the hypothetical question whether the defendants had objected to the assignment if they had been asked for it’.


\textsuperscript{52} Inter-Southern Life Insurance Co v Humphrey (1919) 84 So 625; Aetna Insurance Co v Smith, McKinnon & Son (1918) 78 So 289; Stokes v Liverpool & London & Globe Insurance Co (1925) 126 SE 649. As noted above there may be public policy considerations that deny efficacy to provisions that attempt to stop a person using their choses in action as security. See also Uniform Commercial Code §9–318(4). As to restrictive clauses in floating charges which are designed to secure priority, see William Gough, Company Charges (2nd ed, 1996) at ch 10.

\textsuperscript{53} Such a construction is unlikely where there exists a separate provision dealing with subcontracting; see Linden Gardens, above n4 at 103 (Lord Browne-Wilkinson); Helstan Securities Ltd v Hertfordshire County Council [1978] 3 All ER 262. See also Re Griffin [1899] 1 Ch 408 (here the words ‘this receipt is not transferable’ on a deposit receipt did not prevent the fund from being the subject of an equitable assignment) and Anning v Anning (1907) 4 CLR 1049 at 1067 (Isaacs J). See, further, Restatement of Contracts 2nd §322; Uniform Commercial Code §2–210(3).

\textsuperscript{54} Unreported, English Court of Appeal, 18 December 1996.

\textsuperscript{55} Given that a novation requires the consent of the obligor it is doubtful that one would usually need such a clause to provide protection against a novation.

\textsuperscript{56} Often this limitation results from the nature of the subject right, for example, Tolhurst v Associated Portland Cement Manufacturers, above n43.
Henry LJ also took a different view to that of Evans LJ. He suggested that ‘prior consent never applied for is never withheld or refused (whether reasonably or otherwise)’.

4. The Legal Effect of Prohibitions on Assignment and Their Doctrinal Efficacy

If the conclusion made above that prohibitions on the assignment of contractual rights to performance are efficacious in the sense of being legally valid, then, like the meaning of such a provision, its legal effect must be determined by construction. In any given case, an attempted assignment in the face of a prohibition on assignment may give rise to the following effects:

1. The clause may invalidate the assignment but the agreement to immediately assign may still be valid so that the assignor is liable to the assignee for breach of contract in failing in its promise to immediately assign. In addition, the assignor may or may not be in breach of contract with the obligor.

2. The assignment may be valid but the assignor will be liable to pay damages for breach of contract to the obligor.

3. The clause may invalidate the assignment and in the circumstances of the case the agreement to immediately assign may also be invalid so that the assignor is not liable to the assignee for breach of contract in failing in its promise to immediately assign. In addition, the assignor may or may not be in breach of contract with the obligor.

---

57 See, for example, Linden Gardens, above n4; United Dominions Trust (Commercial) Ltd v Parkway Motors Ltd [1955] 1 WLR 719 (the result in this case was overruled in Wickham Holdings, above n25); Belize Motor Supply Company v Cox [1914] 1 KB 244. See also Juric v Taylor (1894) 152 US 634; Portuguese-American Bank of San Francisco v Welles (1916) 242 US 7; Fortunato v Patten (1895) 41 NE 572; City of Omaha, above n51.

58 See, for example, Wickham Holdings, above n24; Hendry v Chartsearch Ltd, The Times, September 16, 1998; Allhusen v Caristo Construction Corp (1952) 103 NE 2d 891.

59 See, for example, Wickham Holdings, above n24.

60 See, for example, in Devefi, above n1, the prohibition was in the form of a promise not to assign but in the context of the contract which involved the performance of various personal obligations, the court (at 237, 239) appeared to conclude that it had the effect of a true prohibition because earlier in their judgment (at 236) the court distinguished between a true prohibition and a promise not to assign. This distinction is discussed further below.

61 See, for example, Hodder & Tolley, above n40 (combination of a promise not to assign and a forfeiture provision).

62 See also Australian Rugby Union Ltd, above n38 at 735 (affirmed sub nom Hospitality Group Pty Ltd v Australian Rugby Union Ltd [2001] FCA 1040); Hodder & Tolley, above n40. See, further, Attwood & Reid Ltd v Stephens Excavators Ltd [1932] NZLR 1332.

63 See, for example, Hospital for Sick Children (Board of Governors) v Walt Disney Productions Inc [1966] 1 WLR 1055.

64 See, for example, Cathedral Place Pty Ltd v Hyatt Australia Ltd [2003] VSC 385.


66 See Australian Rugby Union Ltd, above n38 at para 100 (affirmed (2001) FCR 157).

4. The attempted assignment may amount to a breach of contract or repudiation giving the obligor the right to terminate performance and thus defeat an assignment even if the assignment were otherwise valid.

At present, although prohibitions on assignment have cropped up in a number of cases over time, the legal effect of such clauses has not been the subject of much detailed judicial analysis. Nevertheless, there is a history of such clauses being upheld in the sense of rendering a purported assignment ineffective. Moreover, it is suggested that the presumed intention of the parties in most cases is likely to be reflected in the legal effect one and most of the following discussion focuses on that effect. This was the result in *Linden Gardens Trust Ltd v Lenesta Sludge Disposals Ltd*. The assignment in that case concerned the assignment of certain rights to performance under a contract. The relevant clause stated:

The employer shall not without the written consent of the contractor assign this contract ... [and] the contractor shall not without the written consent of the employer assign this contract, and shall not without the written consent of the architect ... sublet any portion of the works.

It was argued that the clause had the legal effect set out in two above. This was rejected in favour of legal effect one. Lord Browne-Wilkinson (with whom the other Law Lords agreed) said:

[A] prohibition on assignment normally only invalidates the assignment as against the other party to the contract so as to prevent the transfer of the chose in action: in the absence of the clearest words it cannot operate to invalidate the contract as between the assignor and the assignee and even then it may be ineffective on the grounds of public policy .... [T]he existing authorities establish

---

68 See *Eastern Telegraph Co Ltd v Dent* [1899] 1 QB 835; *Barrow v Isaacs & Son* [1891] 1 QB 417; *McMahon v Docker* (1945) 62 WN(NSW) 155. A lease creates an interest in land and it is not possible to wholly deprive the lessee of its ability to assign that property. The lessor may, however, take a covenant against assignment with a right of re-entry for breach. An assignment in breach of the covenant is effective to vest the legal title but that title is defeasible; *Old Grovebury Manor Farm Ltd v W Seymour Plant Sales & Hire Ltd (No 2)* [1979] 1 WLR 1397; *Williams v Freyne* (1937) 58 CLR 710 at 731 (Dixon J) (citing *Williams v Earle* (1868) LR 3 QB 739). If the lessor unreasonably refuses consent, the lessee may assign without it.

69 *Linden Gardens*, above n4; *Bawejem Ltd v MC Fabrications Ltd* [1999] 1 All ER (Comm) 377. Clearly, a prohibition cannot affect an assignment of a right which occurred prior to the parties agreeing to the prohibition; see *Foamcrete (UK) Ltd*, above n51. This list also leaves aside the possibility that in any particular case the assignee may be liable in tort for inducing a breach of contract.

70 Where damages are available, the normal principles governing damages for breach of contract and restitution for payments made would apply and it will be relevant to determine whether the assignment was by way of sale or security. Moreover, except where the assignee has already exercised a right to terminate the contract or perhaps except where the assignee has recovered damages for breach of contract, there may be a possibility that the contract between the assignor and assignee may still be recognised as an agreement to assign in the sense of a bilateral assignment between the assignor and assignee so that any property which is the subject of the assignment coming into the hands of the assignor is held for the benefit of the assignee; see *Devefi*, above n1 at 236. compare *R v Chester*, above n50 (Millett LJ).
that an attempted assignment of contractual rights in breach of a contractual prohibition is ineffective to transfer such contractual rights …. If the law were otherwise, it would defeat the legitimate commercial reason for inserting the contractual prohibition, viz to ensure that the original parties to the contract are not brought into direct contractual relations with third parties.

The result in the case was that the prohibition was intended to invalidate any tripartite assignment, that is, an assignment that does not only create a relationship between the assignor and assignee but also a relationship between the assignor, assignee and obligor. However, the contract between the assignor and assignee was valid and breached because the assignor had failed to perform its promise to immediately assign. Finally, despite Lord Browne-Wilkinson’s reference to clear words being necessary to invalidate the contract between the assignor and assignee, it is difficult to see how the transaction between the assignor and obligor can in any way impact upon the efficacy of the contract between the assignor and assignee except where some public policy issue renders the contract of assignment unenforceable or otherwise illegal. However, Lord Browne-Wilkinson did note that even with clear words such a provision could be against public policy.

What remains unclear from the decision in Linden Gardens is that Lord Browne-Wilkinson made no statement as to whether the assignor was liable to the obligor for breach of contract in attempting to assign its rights under the contract other than what may be implied from his statement that legal effects two and four are unlikely to ever represent the intention of the parties. In practice, of course, clauses can be drafted to ensure that a purported assignment in the face of such a prohibition cannot impact on arrangements between the assignor and assignee and if a prohibition generally exists for the benefit of the obligor and may be overridden by consent, then if the obligor chooses to perform to and for the benefit of the assignee then it may amount to an implied consent and prevent the obligor being in breach of contract in not performing to the assignor, and, in the case of a legal assignment, the obligor would receive a good discharge. However, except where a prohibition has an express or implied consent provision incorporated in it, it may be questioned whether mere consent is sufficient. If the right in question is not transferable by reason of the agreement made between the assignor and obligor, then doctrinally, the only way to change this is by a formal renegotiation of the contract. See further Nokes v Doncaster Amalgamated Collieries Ltd [1940] AC 1014 at 1018 (Viscount Simon LC).

See, for example, Abbott v Philbin [1961] AC 352, [1960] 1 Ch 27, [1959] 2 All ER 270; Siebe Gorman & Co Ltd v Barclays Bank Ltd [1979] 2 Lloyd’s Rep 142 at 160 (Slade J); Showa Shoji Australia Pty Ltd v Oceanic Life Ltd (1994) 34 NSWLR 548. See also notes 9 and 24.

See, for example, In re Turcan (1888) 40 Ch D 5; United Dominions Trust, above n57 (overruled [1967] 1 WLR 295); Hospital for Sick Children, above n63; Helstan Securities Ltd, above n53 at 264–65; Reed Publishing Holdings Ltd v Kings Reach Investments Ltd (unreported 25 May 1993 Court of Appeal England); Linden Gardens, above n4. See also Brice v Bannister [1878] 3 QB 569 at 580–1 (Bramwell L.J).

Id at 108. See also Burch v Taylor, above n57; City of Omaha, above n51. The principal English case to the contrary, Tom Shaw and Co v Moss Empires Ltd (1908) 25 TLR 190, was explained away as an example of a prohibition not being able to invalidate an accounting between assignor and assignee once the fruits of the contract are in the hands of the assignor – if this explanation was not acceptable Lord Browne-Wilkinson (at 108) said that Tom Shaw was wrongly decided. Linden Gardens, above n4 at 104 (Lord Browne-Wilkinson).
prohibition is a breach of contract, however, there is a real issue as to the position where that is not the case. As regards leases, the position appears to be clear, an assignment will amount to a breach of contract. However, in the case of a lease the assignment is effective. This issue was addressed by Millett LJ in *Hendry v Chartsearch Ltd*. That case concerned a clause stating that the relevant party shall not be ‘entitled’ to assign. It was not like the clause in *Linden Gardens* that was in the form of a promise not to assign. Millett LJ said that a clause must take effect according to its tenor. He thought the assignment was effective as between the assignor and assignee but was ineffective to create a breach of contract between the assignor and obligor. As between assignor and obligor it was simply without effect. It followed that there was no tripartite assignment binding the obligor. Earlier in *R v Chester and North Wales Legal Aid Area Office (No 12)*, a case involving a prohibition in the form that the relevant party ‘shall not assign’, (a promise not to assign) Millett LJ said, in concluding that the prohibition also prevented equitable assignments, that ‘equity will not enforce the performance of an obligation [that is, a promise to assign] which constitutes a breach of a prior contract with a third party [that is, the obligor]’. It is clear that Millett LJ was alive to the distinction between a promise not to assign and a clause that negated any power to assign. In *Hendry*, he noted that a prohibition need not take the form of a convenant not to assign or reserve a power to treat an assignment without consent as a repudiatory breach of contract. It was sufficient, he thought, if the clause was in a form that disentitled a party from assigning. However, ultimately it would appear that he was of the view that no matter what form a clause takes its effect is to render any assignment ineffective.

If the clause states that a party ‘shall not assign’ its rights, that is, a promise not to assign, and if it is thought that such a clause prevents any attempted assignment from being efficacious (as was held in *Linden Gardens*), then how can it be said the assignor has breached its obligation to the obligor, because the end result is that it has not assigned the contract? However, perhaps the more logical result is that

---

79 *Linden Gardens*, above n4 at 104. The Restatement of Contracts 2d §322(2)(b) expressly gives the obligor a right to damages but also recognises the efficacy of the assignment.
80 See *Denham Bros Ltd v W Freestone Leasing Pty Ltd* [2002] QSC 307.
81 *MacDonald v Robins* (1954) 90 CLR 515 at 520 (Dixon CJ).
82 See above n68.
83 *Hendry v Chartsearch Ltd*, above n58.
84 *R v Chester*, above n50.
85 *Id* at 1501. See also *Australian Olympic Committee Inc v Australian Rugby Union Ltd*, above n38 at para 100 (affirmed (2001) FCR 157). See further *New Zealand Payroll Software Systems Ltd v Australian Rugby Union Ltd*, above n65 at para 22, suggesting that a purported assignment in the face of a prohibition was a breach of contract and the only question was whether it should be compensated for in damages or it should simply be held that the assignment never occurred. This is a confusing statement: if it constitutes a breach of contract it must give rise to a right to damages.
an attempted assignment in the face of a ‘promise not to assign’ would be a valid assignment but would put the assignor in breach of contract with the obligor. On the other hand, an attempted assignment in the face of a clause stating that the assignor ‘cannot assign’ would render the attempted assignment ineffective, but would not create a breach of contract between the obligor and assignor as there is no breach of promise. The distinction here may be that a clause stating that a party ‘cannot assign’ renders the contractual right personal and not capable of assignment, whereas a mere ‘promise not to assign’ amounts to a personal assumption of obligation not to do that which one would otherwise be empowered to do.

An issue which then arises is whether the law should draw such a fine doctrinal distinction between a ‘promise not to assign’ and a clause preventing assignment, that is, a true prohibition on assignment. Arguably, if the legal effect of a provision is dependent on construction, then the distinction is a real one and may represent the intention of the parties. It is this distinction that in fact would allow for the different legal effects set out in legal effects one and two above. In addition, as noted above, whether or not a right is personal is a matter of intention and arguably there is not much difference between the designation of a right as personal and an express provision negating the power to assign. The designation of a right as personal is not, however, analogous to a promise not to assign. Moreover, there does not appear to be any authoritative decision which has held that an attempted assignment of a right, which was in turn determined to be personal, resulted in the assignor being liable to pay damages to the obligor for breach of contract by reason of the attempted assignment. Thus, arguably, this distinction is in fact already being drawn.

In addition, it is not possible to characterise a chose in action in the form of a contractual right to performance without taking into account the intention of the parties. The terms of the contract mould the characteristics of the chose in action. Therefore, if the parties state that a right is personal or incapable of assignment, then such provisions define the chose in action so that its non-transferability flows from an inherent flaw in the chose. Thus, although it is no doubt correct to suggest that contract operates through promises, it arguably does not follow that any prohibition

---

86 See also Minister for Land and Water Conservation, above n9 at para 29.
87 Fortunato v Patten, above n57. See also Hodder & Tolley, above n40; Attwood & Reid Ltd v Stephens Excavators Ltd, above n62 (query whether these two cases may be explained away on the basis that the prohibition only prohibited legal and not equitable assignments). See Specialised Transport Pty Ltd, above n9 at 662 (Young J); Devefi, above n1 at 236. See further MacDonald v Robins, above n81 at 520 (Dixon CJ).
88 There is a line of American cases that have held that such a clause does negate the power to assign; see Grismore, above n3 at 305–6 and Sacks v Neptune Meter Co (1932) 258 NYS 254 at 262 (Untermyer J).
89 The distinction and its validity are recognised in Devefi, above n1 at 235, but in that case it would appear that the court concluded (at 237) that the clause, which was in the form of a promise not to assign, when read in the context of the contract as a whole was intended to have the effect of a prohibition rendering the assignment ineffective.
90 See further page 183.
on assignment, no matter how it is drafted, must amount to a mere promise not to assign. This would no doubt be the case if a prohibition only operated as a matter of contract, but arguably, it is not the case where the prohibition is intended to characterises the chose.

This is an important point which can easily be overlooked. For example, in *Devefi Pty Ltd v Mateffy Pearl Nagy Pty Ltd*,91 the court approved a statement from Untermyer J in *Sacks v Neptune Meter Co*,92 where he suggested that prohibitions on the assignment of contractual rights did not impeach the rule against restraints on alienation because such provisions merely condition the obligation to perform rather than the right itself.93 That is, because the contractual restraint conditions the obligation to perform it does not affect the alienability of the right. The result of this analysis would appear to be that the right may continue to be assigned but the obligor has a right to elect to continue to perform to the assignor.

One can see the logic in this explanation from a contract perspective. However, what appears to lie at the heart of this explanation is the view that it is possible to characterise a contractual right divorced from its correlative obligation, and that contract may be used to inhibit transferability as a matter of contract law but cannot change the inherent characteristics of a right of property as a matter of property.94 Thus, if a right of property has the characteristic of being inherently transferable, the parties to the contract may inhibit that as between themselves as a matter of contract, but this does not change the fact that the right has this inherent characteristic. It is suggested that this explanation cannot be sustained for the reasons that follow. However, before turning to those reasons, it should be noted that the court in *Devefi*, although citing the decision of Untermyer J, did not appear to follow that analysis through to its logical conclusion as stated above because in the end they concluded that the assignment was ineffective because 'there was no subject matter to which the purported assignment … could attach'.95 This would appear to recognise that the prohibition there robbed the subject right of its transferability.

Turning then to the reasons for the above conclusion: Firstly, as to the right and its correlative obligation, it is not possible to characterise the right to some

---

91 *Devefi*, above n1 at 234–7.
92 *Sacks v Neptune Meter Co*, above n88 at 262 (see above quote).
93 *Devefi*, above n1 at 235. It may be noted that Untermyer J drew a distinction between a provision that seeks to negate the power to assign and a provision containing a promise not to assign. The latter, he held, only gives rise to a right to damages and does not impact upon the efficacy of the assignment. As already noted, the court in *Devefi*, above n1 at 235, also appear to recognise this distinction.
94 Similar reasoning appears to lie behind Millett LJ’s comment in *Hendry v Chartsearch Ltd*, above n58, that as between the parties to it, an ordinary commercial contract is not property but obligation and therefore there is no objection to making the benefit of the contract non-assignable. Although it is true that a contractual right is only property for certain purposes, there seems to be an underlying belief here that individuals cannot tamper with the characteristics of something the law considers property.
95 *Devefi*, above n1 at 239.
contractual obligation divorced from the obligation itself. The fact that the personal nature of a right depends upon the construction of its correlative obligation (taken in the context of the contract as a whole) shows that there is a relationship between the right and the obligation that cannot be separated. That is, whether or not a right is personal depends upon whether the obligor intended to perform the obligation only for the benefit of the other party to the contract. This was in fact admitted by Untermyer J when, after stating that the restriction is imposed on the duty to perform, said that the ‘restriction then becomes a condition of the acquisition’.96

Secondly, as to intention and the characterisation of rights as choses in action, although the issue as to whether or not a right is property is a matter to be determined by those principles and policies that determine the existence of property, it does not follow that the intention of the parties to a contract is irrelevant as regards the characteristics of that right so far as its nature as a chose in action is concerned. When a right owes its existence to an agreement between parties, it must be correct that the contract itself can define the characteristics of that right both in terms of contract and property.97 Therefore, it is not only impossible to divorce a contractual right from its correlative obligation, it is also not possible to divorce the alienability of that right from the correlative obligation.

This can be demonstrated by looking at the genesis of the rule that prevents the assignment of personal contractual rights. If the institution of the assignment of choses in action was in part driven by the needs of commercial people, then why have a rule preventing the assignment of personal contractual rights, and why base the characterisation of a right as personal on the intention of the parties? Clearly, greater assignability would have been promoted by a rule that determined assignability by reference to whether or not a reasonable person in the position of the obligor would object to the assignment, and this in turn could be tested by reference to any material change in the circumstances of the obligor by reason of the assignment.98 It may be that originally, the commercial needs driving assignment were limited to debts which were not considered personal, thus leaving contractual rights to performance in the ‘too personal’ basket. At that time though, the notion that a right was ‘personal’ was different to the notion that we have today that prevents the assignment of a personal contractual right. Historically, attaching the label ‘personal’ to contractual rights was seen as part of the inherent nature of such rights rather than as a matter that falls to be determined by the construction of the contract. It is therefore suggested that the explanation for the personal rights rule does not flow from some vague notion of upholding party autonomy or freedom of contract. The result flows from the institution of assignment itself. An assignment involves a transfer of property. Legal transfers are governed by the

96 Sacks v Neptune Meter Co, above n88 at 262.
98 One commentator has argued for the law to move in that direction; see Larry A. DiMatteo ‘Depersonalization of Personal Service Contracts: The Search for a Modern Approach to Assignability’ (1994) 27 Akron L Rev 407.
nemo dat rule. The reason why a contractual right designated as personal under a contract cannot be assigned is because the result would be that the assignor would be assigning a right different to the one that he or she possesses, which is at odds with the nemo dat rule.99 That is, the intention of the parties as expressed in the contract is not merely shaping contractual rights in their character as personal rights, but also in their character as choses in action. The legal notion of transfer, which is the essence of an assignment of a chose in action, dictates the existence of the personal rights rule with the result that the parties can rob a right of its transferable character.

It may be that the point Untermyer J was really expressing was no more than that, although the parties to a contract may render a right incapable of transfer, this does not prevent the ‘assignor’ dealing with the subject matter of the contract once it is in the hands of the assignor as the assignor would then be transacting in new and different rights.

Finally, a further problem with this explanation is that it appears to suggest that the presence of a prohibition allows the assignor to transfer title to the right to performance without transferring the actual right to performance, that is, the chose or thing as distinct from the title to the chose or thing. This result can be achieved as between the assignor and assignee as a matter of contract, but it cannot be achieved from the perspective of personal property because unlike tangibles, where title and possession can be vested in different people, the nature of contractual rights as choses in action does not allow a transfer of title without a transfer of the actual right to performance. Moreover, if it was possible to separate the title from the thing or chose, then arguably the rule that prevents the assignment of personal rights is misconceived. That is, the existence of a personal contractual right would not prevent the assignor assigning the title to that right, because in all instances he or she could argue that the obligation to perform remained an obligation to the assignor. An assignment of a contractual right involves a transfer of both the title and the thing or chose and this is why an assignee can call for performance. In the case of personal rights to performance, what may be assigned, subject to overriding public policy concerns, is the future secondary right to damages, which is a distinct chose in action. When a right is personal it is simply inalienable and the doctrinal inability to separate the title from the chose cannot be drafted around even by the assignor and obligor.

The above distinction between promises and prohibitions provides an analytical way forward. Clauses that negate the power to assign and contracts that on construction (express or implied) contain personal rights have the effect of robbing the chose of its characteristic of transferability. Clauses in the form of promises not to assign operate merely at the level of contract, leaving the obligor with a remedy of damages if the other party to the contract assigns the subject right.

There are, however, arguments that a distinction between promises not to assign and true prohibitions on assignment should not be drawn.100 Unless very

99 See Minister for Land and Water Conservation, above n9 at para 8 (Mason P).
100 Compare Allhusen, above n58. See also Grismore, above n3.
well advised, contractual parties themselves are unlikely to appreciate this distinction for it to represent their presumed intentions. In addition, in many cases the language of promise and the language of prohibition may be difficult to distinguish and both may appear in a provision. As already noted, in Devefi, where the court clearly recognised the distinction between promises and prohibitions, a clause that on its face encapsulated a promise not to assign appeared to be construed by the court, when read in light of the contract as a whole, as a prohibition.\(^{101}\) Moreover, the issue here is analogous to that arising in the case of options. An option may be explicable either on the basis of an offer coupled with a contract not to revoke or as a conditional contract. Logically, a revocation of an option under the irrevocable offer theory would still be effective (as there is only a promise not to revoke), but would amount to a breach of contract not to revoke, whereas such a revocation under the conditional contract theory would be totally ineffective and would not amount to a breach of contract.\(^{102}\) However, in Goldsbrough Mort v Quinn\(^{103}\) Isaacs J, who appeared to prefer the irrevocable offer theory,\(^{104}\) said that a revocation was nevertheless ineffective and in any case, the court could order specific performance.\(^{105}\) In his view, the revocation was ineffective at law.\(^{106}\) However, at another point, he did state that what was sold was the option and not merely the promise to give an option,\(^{107}\) that is, the subject matter of the contract and not merely the contractual right. O’Connor J was also of the view that no matter what theory was used, the court could order specific performance.\(^{108}\) That is, in his view, whether or not the revocation was valid at law giving rise merely to a right to damages, equity would still order specific performance making the revocation ineffective. This would suggest that fine distinctions, which may be drawn from the express language of a contract, will not be drawn if they would defeat the purpose of the transaction or provision.\(^{109}\)

Another example, but pointing in the opposite direction, is the decision in Cowell v Rosehill Racecourse Co Ltd.\(^{110}\) In this case, on the accepted evidence that a contractual licence contained a promise not to revoke, it was held that a revocation was nevertheless effective as the licence was not coupled with a grant. The result was that the licensor was liable in damages for breach of contract.\(^{111}\) It is unlikely that the High Court would have come to a different conclusion if the licence was expressed as ‘irrevocable’. Again, however, as with options, this result

---

101 See above page 174.
102 See Laybutt v Amoco Australia Pty Ltd (1974) 132 CLR 57 at 76 (Gibbs J).
103 Goldsbrough Mort v Quinn (1910) 10 CLR 674.
104 Id at 690–92.
105 Id at 679 (Griffith CJ).
106 Id at 691.
107 Id at 692.
108 Id at 686.
109 However, it is quite clear, if a party does not simply attempt to revoke an option but rather sells the legal interest in the subject property to a third party, that third party may, if it does not have notice of the option, get good title and the optionee will be left with an action for damages against the optionor. See Blacktown Municipal Council v Doneo [1971] 1 NSWLR 157 at 162 (Taylor AJA). See further, Donald J Farrands, The Law of Options (1992) at 26–34.
is informed by the context of the subject matter. Thus, no one would argue that the ratio of Cowell’s case necessarily applies to documentary credits so as to make an irrevocable documentary credit (which is simply stated to be irrevocable) revocable upon the petition of the buyer.

Finally, it may be noted that the court in Linden Gardens, which was a case involving a promise not to assign, did not appear to draw the distinction between a promise and a prohibition. This must represent the position in England for the time being, and although it is at this point that the decision in Linden Gardens appears to differ from the decision in Devefi, where the distinction was recognised, the result in Devefi shows that in practice it is a distinction that will rarely be drawn as it will rarely represent the intention of the parties. That is, unless there are clear words to the contrary, the intention behind such a provision is to invalidate any attempted assignment. In short, such provisions, whether drafted as promises or true prohibitions, are intended to characterise the contractual right as a chose in action and to rob it of its transferability.

Two questions then arise. Firstly, is it the case that no matter how a prohibition is drafted, from a contractual perspective will it be sufficiently promissory to give rise to a breach of contract by the assignor vis-à-vis the obligor by reason of the attempted assignment? Secondly, if the answer to the first question is in the affirmative, then how is it possible to distinguish prohibitions on assignment from personal contractual rights given that an attempted assignment of the latter does not appear to result in a breach of contract giving rise to an obligation to pay damages? A delegation of a personal obligation is of course different, as that would amount to a defective performance and might evidence a repudiation. Often,

111 This is not the position in England, where it is clearly recognised that equity may grant an injunction to restrain the revocation; see Kevin Gray & Susan Gray Elements of Land Law (3rd ed, 2001) at 181. There has been mention in Australian cases of the possibility of an injunction issuing to prevent revocation; see Bingham v 7-Eleven Stores Pty Ltd [2002] QSC 209; Cathay Developments Pty Ltd v Laser Entertainment Pty Ltd [1998] NSWSC 82. In at least one case the decision in Cowell was explained away on the basis that the result was dictated by the separation existing in 1937 between law and equity in New South Wales; see Leonard George Munday v ACT [1998] SCACT 62 at para 172.

112 Compare Barker v Stickney [1919] 1 KB 121, where a publisher promised not to assign copyright except subject to the conditions of the contract between the author and publisher. It was held in that case that an assignment of that right was valid and the assignee was not required to pay the author royalties. The point was not taken that this provision amounted to a prohibition, and perhaps this was because it was possible to distinguish prohibitions on assignment from personal contractual rights given that an attempted assignment of the latter does not appear to result in a breach of contract giving rise to an obligation to pay damages. A delegation of a personal obligation is of course different, as that would amount to a defective performance and might evidence a repudiation. Often,

113 For American authorities on this point, see Burck v Taylor, above n57; City of Omaha, above n51. See also Grismore, above n3 at 303. Compare Fortunato v Patten, above n57.

114 Compare Australian Olympic Committee Inc, above n9 at paras 119–120, where in answer to a submission that there was an implied term prohibiting assignment, Lindgren J said that if that were correct, then the effect of the prohibition would be to render the subject right personal and any purported assignment would amount to a breach of contract and perhaps a wrongful repudiation. See also Joseph Starke, Assignment of Choses in Action in Australia (1972) at 66, suggesting that in the first instance, it is necessary to determine whether a prohibition rendered the contract personal and if not whether it in any case rendered any assignment inoperative. See further Walter Jaeger (ed), Williston on Contracts (3rd ed, 1993) vol 3 at 46.
attempts to assign personal rights arise in the context of the assignor selling a
business or otherwise ceasing to exist with the result that the obligor is excused from
performance. Any breach or repudiation flows from the assignor’s conduct in no longer performing its obligations under the contract. If these
obligations are fully performed the obligor is simply excused from further
performance by reason of the assignor having no further interest in the contract.

In the case of a promise not to assign, it will not be difficult to imply an
obligation that the parties will not attempt to assign. This would simply flow from
a co-operative construction of the promise. If that is right, then the attempt to
assign gives rise to a breach of contract, making the assignor liable in damages, to
the obligor. Such damages will be nominal because, as the weight of authority
suggests, the assignment will be ineffective so that the obligor has suffered no
substantial loss.

There may also be an argument that a clause in the form that the parties ‘cannot
assign’ (a prohibition) may still imply a promise not to attempt to assign giving rise
to an action for breach if such an attempt is made. If that is not the case then the
law is straightforward: an attempted assignment in the face of such a prohibition
or in the face of a contract containing personal rights is ineffective and does not
give rise to a breach of contract.

If this is not the position then a distinction needs to be drawn between such
prohibitions and personal contractual rights. That is, if an attempted assignment in
the face of a prohibition gives rise to a breach of contract, this needs to be
distinguished from an attempted assignment of a right that is construed to be
personal, as the latter does not appear to give rise to a breach of contract. It may be
that a difference arises in the fact that a right is generally found to be personal by
virtue of construing the entire contract. That is, the personal nature of the right is
usually implied by construction rather than being found in an express provision.
Moreover, it should be noted that this process of construction determines whether
the right is personal rather than whether it is intended to be assignable. This may
dictate that any attempted assignment of such a right would not amount to a breach
of contract as it is difficult to imply from this a promise not to attempt to assign.

What then of an express provision making a right personal? On one view, this
should be subject to the same analysis above as regards personal rights implied by
construction, because it addresses the personal nature of the right rather than

\footnote{See, for example, \textit{Kemp v Baerselman} [1906] 2 KB 606; \textit{Robson v Drummond} (1831) 2 B & AD 303, 109 ER 1156; \textit{Dr Jaeger’s Sanitary Woollen System Co Ltd v Walker and Sons} (1897) 77 LT 180. In the case of a personal right to insurance, the attempted assignment will simply result in the policy lapsing; see \textit{Peters v General Accident & Life Assurance Corp Ltd} [1937] 4 All ER 628.}

\footnote{See \textit{Nokes v Doncaster}, above n72.}

\footnote{See \textit{John Young & Co Kelvinhaugh Ltd v the Rugby Group Plc} (Queens Bench Division, HT 00/337, Judge Richard Seymour QC, 19 December 2000) at para 24.}

\footnote{The reference to ‘implied by construction’ here is used to distinguish the origins of the obligation from one that owes its existence to a specific clause of the contract. Strictly, if an obligation is found by construing the words of a written contract that obligation is an express obligation.
expressly addressing assignability. However, perhaps when an express personal rights provision is incorporated into a contract, it is reasonable to conclude that the contracting parties would see the relationship between this and a true prohibition. What other reason could there be for making a right expressly personal?119 In fact, often a clause will state that rights are personal and then go on to state that rights are not assignable.120 If that is right, then where the right is made personal by reason of an express provision, then a right to damages should flow from an attempted assignment, if such a right to damages also flows from an attempted assignment in the face of a true prohibition. If it is possible to imply a promise not to attempt to assign in the latter, it must also be possible to similarly imply it in the former.

Strictly, however, there is no distinction between the effect of non-compliance with an obligation that is express and one that is implied. That is, even if the existence of an obligation is implied by construction, its existence as a contractual obligation must be because it is a term of the contract which gives rise to a right to damages upon breach. Thus, whether a right is personal by reason of an express provision or implied by construction, the results that flow from an attempted assignment should be the same. If there is to be a difference in the way such personal rights are to be treated, it must ultimately lie in adopting the position that it is easier to imply an obligation not to attempt to assign where the personal nature of the right is the subject of an express term than where the personal nature of a right is implied by construction. In any case, since the assignment is ineffective, the damages of the obligor are only going to be nominal and, in fact, it would be difficult in most cases to identify any loss.

Finally, something must be said about effects two to four. At the outset, since the legal effect of such provisions depends on construction, it is possible to have any number of possible effects. Nevertheless, as regards legal effect two, it would be rare for a simple prohibition to be intended to have this effect — that is, because it would need to both prohibit assignment and at the same time envisage it. Legal effect two generally requires the courts to place much more emphasis on the distinction between prohibitions and promises not to assign than is currently the case. Perhaps one instance where this might occur is where the clause merely prohibits assignment without consent, and where the purposes of the obligor are best served by recognising the initial transferability of the chose subject to a right in the obligor to later elect to deny the assignee the relevant benefit.121 However, if that is the correct construction — that is, if the reference to ‘consent’ in the prohibition is really just a right to elect to not recognise the right vested in the

---

119 One reason may be that there is confusion between the ideas of assignment and vicarious performance, and the intention was merely to prohibit vicarious performance.
120 See, for example, ANC Ltd v Clark, above n51.
121 See Australian Rugby Union Ltd, above n38 at paras 23, 29, 100 (affirmed (2001) FCR 157). A condition of a ticket was that it was inter alia not to be resold at a premium or used for advertising without prior consent — a denial of the use of the ticket only attached to an unauthorised sale leaving the normal remedies for breach of contract to flow from an unauthorised use. See also the apparent approval of the possibility of legal effect two in Broadcast Australia Pty Ltd v Minister for Natural Resources (Lands) (2004) 204 ALR 46.
assignee — then the initial attempted assignment could not amount to a breach of contract. To achieve legal effect two, the contract would need to allow for assignment but at the same time state that certain assignments, though valid, would nevertheless amount to a breach of contract. Another way legal effect two could come about is if a clause merely gives the obligor a right to terminate the contract upon assignment and it does not do so. Here the assignment may be valid but there may still be a breach of contract by the assignor. A similar result may flow from a peculiar fact scenario. For example, if the owner of land upon which there exists an ice works enters into a contract with a neighbour to supply ice for a period of years, there is no doubt that the owner of the land still has a right to sell his or her land. The contract for the supply of ice does not act as a prohibition on such a sale (assignment). However, the effect of the sale may be to repudiate the supply contract, making the owner liable in damages for breach of contract.

Not much needs to be said about legal effect three as most of the issues raised by it have been dealt with above. The only outstanding issue is the efficacy of the contract between the assignor and assignee, and all that needs to be said about that is that in any given case the assignor may be protected from an action in breach of contract if the contract between the assignor and assignee is unenforceable, or if the assignor only promised to assign to the extent that he or she was capable of doing so.

As regards legal effect four, whether or not an attempted assignment may give rise to a breach of contract even though the assignment itself is ineffective has already been covered. However, it may be noted that both Lord Browne-Wilkinson in *Linden Gardens* and Millett LJ in *Hendry v Chartsearch Ltd*, suggested that an assignment or attempted assignment in the face of a prohibition (whether promissory or prohibitory) will not of itself amount to a repudiatory breach of contract. There is some strength in the suggestion that if the assignment is not valid there cannot be a repudiatory breach. Moreover, the fact that damages will always be nominal would point to the conclusion that any breach is not a breach of condition, a sufficiently serious breach of an intermediate term or a repudiation. In addition, it would be difficult to identify a repudiation when the assignor is merely dealing with its rights. Nevertheless, since the issue of repudiation is dependent upon the facts of each case, and given that the terms of a clause may state that an attempted assignment will amount to a repudiation, then this possible effect cannot be simply written off. Finally, the possibility of a court finding a repudiation must be increased if the right in question is one of those rare instances where the burden is also transferred, because here there would be evidence that the

---

122 See *Australian Rugby Union Ltd*, above n38 at para 100 (affirmed (2001) FCR 157 esp at 182–84).
123 *Proctor v Union Coal Co* (1923) 137 NE 659. See also *Roadshow Entertainment Pty Ltd v (ACN 033 006 269) Pty Ltd* (1997) 42 NSWLR 462.
124 See *Bruce v Tyley*, above n5. See also *Barker v Stickney*, above n112.
125 This point was left open in *Westgold Resources NL*, above n9 at 431 (Anderson J) (affirmed [2000] WASCA 85).
126 See further p175.
127 See, for example, *Denham Bros Ltd*, above n80.
assignor was, by its conduct, evincing an intention to be no longer bound by the contract.

A refinement of this issue has been suggested by the authors of *Meagher, Gummow and Lehane’s Equity Doctrines and Remedies*. The authors suggest that in determining the effect of a prohibition it may be important to determine whether the prohibition amounts to a condition, warranty or intermediate term.\(^{128}\) The argument put forward is that if a prohibition merely amounts to a warranty, then although the assignor would be liable in damages for breach of contract to the obligor, this construction would evidence that the obligor did not place sufficient importance on the clause to prevent the assignment itself being upheld.\(^{129}\) In addition, even where the clause is construed as a condition, the efficacy of the assignment will only be impaired if the obligor elects to terminate its contract with the assignor for the breach. With respect, the flaw in this reasoning is simply that the tripartite classification of terms is only relevant in determining whether or not there exists a right to terminate. It has nothing to do with the efficacy of the assignment except to the extent, as noted above, that an election by the obligor to terminate its contract with the assignor for breach or repudiation may extinguish the assignee’s rights if those rights have not unconditionally accrued by the time of discharge. It follows that the tripartite classification would be important to an obligor if a mere promise not to assign did not of itself prevent the assignment and perhaps this is the context in which the authors intended their remarks to operate.

5. **Who Can Enforce a Prohibition?**

The general position appears to be that a prohibition exists for the benefit of the obligor\(^ {130}\) and neither the assignor (or, anyone claiming through the assignor)\(^ {131}\) nor arguably an assignee can raise it to defeat an assignment. The former is uncontroversial. Clearly the assignor and anyone claiming through the assignor cannot raise the prohibition to defeat the assignment because of the valid agreement between the assignor and assignee. The latter, that is, that one assignee cannot raise the presence of the prohibition against another assignee, is more problematic.

Before moving on, it should be noted that the issue here does not concern the efficacy of the assignment as between the assignee and obligor. Clearly, the weight of current authority dictates that an assignment in the face of a prohibition is of no effect as between the assignee and obligor. Unless there are clear words to the

\(^{128}\) See Meagher, Heydon & Leeming, above n78 at 275–278. See also Starke, above n114 at 66–7, para 6.465.

\(^{129}\) See Meagher, Heydon & Leeming, id at 275–278, para 6.465.

\(^{130}\) Portuguese-American Bank of San Francisco, above n57. However, a prohibition may not necessarily be for the benefit of the obligor; see Restatement of Contracts 2nd (1979) 322, Comment (d).

\(^{131}\) Query whether Re Griffin, above n53; Anning v Anning, above n53 and Re Westerton; Public Trustee v Gray [1919] 2 Ch 104 are explicable on this ground; see B Allcock, ‘Restrictions on the Assignment of Contractual Rights’ (1983) 42 CLJ 328 at 332–33 also citing Hodder & Tolley, above n40. See further Attwood & Reid Ltd v Stephens Excavators Ltd, above n62.
contrary, this result is not dependent on the obligor electing to avoid the assignment; the result flows automatically from the prohibition.\(^{132}\) This is its intended effect. It should then follow that generally there can be no true priority dispute between an assignee who takes an assignment with the consent of the obligor and one who takes an assignment in breach of a prohibition, as the latter gets no proprietary interest vis-à-vis the obligor and therefore cannot give a valid notice to the obligor.\(^{133}\) Even if an analysis allows for an assignee, who takes in breach of a prohibition, to obtain an interest under the contract which would therefore allow it to give notice, this will not be worth much if the obligor is not under any obligation to that assignee. The issue that is the concern of this section is one of accounting as between assignees; and the real practical issue arises where an assignee taking with the consent of the obligor is paid by the obligor, and an assignee who has taken in breach of a prohibition wants to claim part of that payment from the first assignee.

There are perhaps three ways to approach giving an answer to this issue. The first is to view a prohibition as purely contractual so that only the party to the contract, for whose benefit the provision is provided, can enforce the provision. This explanation can only be accepted if a prohibition, no matter how it is drafted, only operates as a matter of contract. This was rejected above. However, if this explanation were correct then the answer must be that no assignee can raise the prohibition to defeat another assignee.

Secondly, and related to the first explanation, is that being merely contractual, it characterises the obligation to perform and does not impact on the character of the assignor’s contractual right as a piece of property having the characteristic of being transferable. Therefore, the right remains transferable as a chose in action but gives the obligor a right to elect not to perform for the benefit of the assignee. This explanation of prohibitions was also rejected above. However, if correct, the result as between assignees must again be that each assignment is effective and the prohibition cannot be raised by one assignee against another.\(^{134}\)

The third explanation, and it is suggested the best explanation, is that the intention of the parties as expressed in the contract can define contractual rights in their character as choses in action; and in most cases the obligor is only concerned to deal with assignability as between itself and the assignor, so as to prevent any relationship being formed between the obligor and an assignee without the consent of the obligor. In such a case, if the right would inherently carry the characteristic

\(^{132}\) Of course in any case the obligor may consent to the assignment, compare above n72.

\(^{133}\) It is important to keep this issue distinct from the validity of the bilateral assignments that may exist between the assignor and any assignees that seek to deal with the fruits of the contract. Because of the effect of the prohibition, these contracts are only efficacious as regards those fruits that are in the hands of the assignor. This is an important point because, in respect of equitable assignments of legal property, it is sometimes said (incorrectly it is suggested) that such assignments only operate as between the assignor and assignee; see Tolhurst, above n50.

\(^{134}\) This explanation, like some instances of the third explanation to be discussed, would allow an interest to be vested in the assignee, but for the purposes of giving notice to preserve priority that interest it is suggested is worthless.
of being transferable, it will continue to do so for other purposes and therefore cannot be raised by competing assignees. However, in any given case the obligor may intend to rob the right of transferability for all purposes, in which case it can be raised by competing assignees. These two results cannot be accommodated within a contractual analysis of prohibitions.

There is no reason to see assignability as an all or nothing issue. Such a result is a far too simplistic analysis of property rights. It is suggested that the non-assignability of a chose in action (that would normally be transferable) only goes so far as is intended. This is why a prohibition may only limit assignment for some purposes, or may limit assignment only while the assignor has not earned the relevant counter-performance, or may limit equitable assignments but not legal assignments or vice versa, or may limit assignment without consent.135

One important authority which, it is suggested, clearly reflects the third explanation and shows the weakness in the other explanations is the decision in Burck v Taylor.136 A much simplified version of the facts in this case is that a construction contract (for a new capital building in Texas) contained a promise not to assign without consent. The assignor, with consent, assigned an interest in the contract to an assignee and later, without consent, assigned an interest to a second assignee. The first assignee took over the burden of the contract and completed the work and was paid the contract price. The second assignee then claimed against the first assignee for a part of the contract price. The Court held that the assignor could not assign any right without the consent of the obligor and all the second assignee obtained was a right against the assignor. Thus, only the first assignment was valid and, on completion of the work by the assignee, it alone was entitled to the contract price even though the obligor made no complaint about the second assignment. It is suggested that the contract based explanations cannot explain this result. If the prohibition was only for the benefit of the obligor and the obligor made no protest, the subject right must have been assigned for the purposes of relations between the assignees.

Later in Fortunato v Patten,137 another case involving a promise not to assign without consent, it was held that such a provision is solely for the benefit of the obligor and cannot be raised by competing assignees. The decision in Burck was distinguished on the basis that:138

[It] dealt with a contract made by an individual with the state of Texas, which contained an absolute, unqualified covenant that it should not be assigned .... [whereas in the present case] the covenant is that, if the contract, or any of the moneys due under it, are assigned without consent, no claim can be asserted by virtue thereof .... [Here] no absolute assignment has been made of the contract, but all transfers were of moneys due thereunder as collateral to secure the payment of a debt. There is a wide difference between assigning moneys due

---

135 Compare International Polymers Pty Ltd v Custom Credit Corporation Ltd, above n9 at 75.631.
136 Burck v Taylor, above n57.
137 Fortunato v Patten, above n57.
138 Id at 573.
under a contract, and an absolute assignment of the contract itself, as the latter act disturbs that relation of personal confidence which exists between one desiring work done that requires a high order of skill and intelligence and the contractor he may have selected as possessing these necessary qualifications.

On one view this passage draws a distinction between the assignment of a right to performance and an assignment of the fruits of a contract, and in the case of the latter the prohibition cannot be raised by competing assignees. However, in both cases the right assigned was a right to a payment under the contract and it was assigned prior to it being earned.\textsuperscript{139} Thus, both cases concerned the assignment of rights to performance, the only difference being that the above passage from the decision in \textit{Fortunato} addresses assignments by way of security. Neither case concerned the issue of a prohibition seeking to deal with the fruits of the contract once those fruits are in the hands of the assignor, nor an assignment of only unconditionally accrued rights to payment.

On another view the decision in \textit{Fortunato} appears to suggest that in \textit{Burck}, the first ‘assignee’ was taking over the contract, in the sense of taking over the assignor’s obligations to the State and not the assignor’s obligations to any other ‘assignees’. Thus, what was probably intended was a novation.\textsuperscript{140} It would follow that if the contract was novated to the first assignee, then clearly the second assignee could have no interest in that contract as the assignor had no interest to give the second assignee and therefore it did not matter that the obligor did not protest the second assignment.

The facts in \textit{Burck} were that the original ‘assignment’ to the first assignee was only a three-quarter interest in the contract, and the second assignee was intended to take the one quarter interest left vested with the assignor. The Court made the point that the second assignee may have had a claim against the assignor for any monies it had earned under the contract; but, as it was the first assignee who in fact did all the work and earned the entire contract price, the assignor earned nothing against which the second assignee could make a claim.\textsuperscript{141} The decision of the

\textsuperscript{139} Where there is an assignment of a right to the performance of a contract and that right later unconditionally accrues, the assignee can then call for that performance. This is also the case in the assignment of a right to payment under a contract. That is, where there is an assignment of a conditional right to performance, the assignee can call for payment when the obligation to pay unconditionally accrues. The assignee’s conditional right is replaced with an unconditional right. It does not matter that upon the obligation unconditionally accruing the right to pay becomes a right to a debt which is a distinct chose in action. This is because the obligation to pay is still an obligation constituting the performance of the contract. However, if there was merely an assignment of the right to the performance of a contract (where that performance will involve a payment), and at the time of the assignment there already exists unconditional obligations to make payments under the contract, the assignment of the right to performance will not capture those rights because they are distinct choses in action; see \textit{Geroff v CAPD Enterprises Pty Ltd} [2003] ACLR 110.

\textsuperscript{140} The Court in \textit{Burck} at one point do seem to make this determination; see \textit{Burck v Taylor}, above n57 at 650 (but compare at 651), and this was the understanding of the decision in Portuguese-American Bank of San Francisco, above n57.

\textsuperscript{141} \textit{Burck v Taylor}, above n57 at 652–53.
Court on this point seems akin to the position of a trustee in bankruptcy taking over the contract and appears to be based on principles of assignment rather than novation. The Court stated that:

[W]hen the contract, being wholly executory, is transferred to a third party, who is accepted by the promisor in lieu of the original contractor, such third party enters upon the performance of the contract free from any disposition of the profits made by the original contract before the substitution.\(^\text{142}\)

It is difficult to see how the Court could have had novation in mind if the effect of the transaction was that only three-quarters of the rights in the contract were ‘transferred’ to the first ‘assignee’ but 100% of the burden, and yet at the same time take the view that if the assignor performed then the second assignee may have a right to payment in respect of that performance. If 100% of the burden was novated to the first assignee, then any act of the assignor could not be an act in performance of the contract. In the result, if full effect is to be given to all the statements made in the judgment then assignment and not novation appears to be a better explanation.

More importantly, the Court in *Burck* expressly dealt with the submission that the prohibition was only for the benefit of the obligor and found, on the facts, that that was not the case because the clause in question did not merely involve a promise not to assign, but rather it dealt with the effects of such an assignment and stated that such a transfer shall ‘annul the contract’. Thus, what really occurred in *Burck* was, putting aside the assignment of burdens point, that the prohibition which was incorporated into the contract sought to completely characterise the relevant right both as a matter of contract and as a chose in action, so that unless consent was obtained it did not have its inherent characteristic of transferability, so that an ‘assignee’ could not (as against another assignee) claim it was transferable even where the obligor made no formal objection to the transfer.\(^\text{143}\) Here the obligor did not intend to characterise its obligation (and the correlative right) only as between the obligor and assignor. It intended to maintain that character even as between competing assignees, whereas this was not the intention in *Fortunato*. This is not, however, a case of the obligor crossing the line and seeking to regulate relations between an assignor and any assignees. The obligor was merely seeking to regulate the position between it and any assignees. It is difficult to see how these variations can be achieved by adopting the view that prohibitions merely operate through contract. Contract has a limited ability to benefit or burden third parties, and where this is required or is the result, then property is a legal concept better suited to explain such results.

\(^{142}\) Id at 653.

\(^{143}\) See also id at 651 where the Court says the result would be the same without the provision as the right assigned was a personal right which could not be assigned without consent.
6. Conclusion

This paper has sought to elucidate three main points. Firstly, that there is no general principle of public policy that would inhibit the general efficacy of prohibitions on assignment when the right assigned concerns a right to the performance of a contract. However, legal practitioners should be wary of placing too much reliance on prohibitions when the subject matter relates to debts and particularly trade related debts.

Secondly, it would appear to be the position that prohibitions, whether expressed as true prohibitions or promises not to assign, will usually be construed as intended to prevent the assignment occurring. If parties wish only to contract for a mere promise not to assign they should state this clearly.

Third, it has been suggested that the doctrinal efficacy of these prohibitions lies in recognising the role of the intention of parties as expressed in a contract to shape contract rights as choses in action and rob them of their inherent transferability.
Trade Marks and Language

MEGAN RICHARDSON*

‘… it is true that we live by symbols …’

Abstract

The traditional perception is that trade marks signal the origin of goods and services. The perception is challenged in this article. It is argued that trade marks are an important component of popular culture. Popular culture thrives in a fluid, unregulated environment. Yet the rhetoric of trade mark law is that the great open commons of the English language require the protective mantle of regulation: ‘trade marks’ should be narrowly defined and thresholds for registration set high in order that the language commons should remain in their pristine natural state; while at the other end infringement of a registered trade mark should be narrowly construed to avoid anything that would grant a full ‘proprietary’ right in a trade mark. The article explores the evolution of such ideas and their apparent hold still over the law. It is observed that there is a more sophisticated understanding now of the nature and function of language as a cultural device compared with that which existed at the time the British trade mark registration system was established. And it is suggested that trade marks that develop a cultural dimension should be granted rights accordingly. Regulation in the name of protecting the language commons should be kept to a minimum, targeting cases where trade mark owners seek to use their trade marks as instruments of censorship and control.

* Associate Professor, Law Faculty, the University of Melbourne. Thanks to Michael Arblaster, Graeme Austin, Kathy Bowrey, Helen Dawson, Jason Bosland, Wendy Gordon, David Lindsay, Janice Luck, Thomas McCarthy, Patrick Parkinson, JC Smith, Michael Spence, Kimberly Weatherall and two anonymous referees for the Sydney Law Review for helpful comments and insights; to organisers and participants in the Australian Intellectual Property Academics Conference, Gold Coast, July 2003 and World Trade Organisation Conference, Adelaide, February 2004 for opportunities to workshop this paper in various stages; and to my intellectual property students at the University of Melbourne who over the years acted as sounding boards and inspirations for ideas that finally made their way into this article. References are current as at March 2004.

1 Mishawaki Rubber & Woollen Manufacturing Co v SS Kresge Co 316 US 203, 205 (1941) (Frankfurter J).
1. Introduction

Language is a remarkable instrument of expressive activity. Its capability to reflect and influence human behaviour is still scarcely comprehended. But one thing is clear: words are not merely bi-products of social interchange; they are an essential feature of the way in which we ‘constitute and articulate our world’. And an ever-changing feature. In the last century alone the English language’s corpus of words and meanings has increased by more than a quarter to accommodate the complexity of modern existence, spurred on by the energy of traders looking to capture the public imagination (and no doubt to make a living as well). Expressions like ‘Roller Blades’ and ‘Barbie Doll’ and ‘Star Wars’ never existed in the lexicon until coined by traders. And ‘Just Do It’, ‘Enjoy’, ‘The Real Thing’ and ‘You Know You Want It’, to list but a few of the multitude of expressions chosen for their ‘effervescent qualities’ and ‘injected into the stream of communication with the pressure of a firehose’, had different more limited significations in the pre-advertising age.

The line between trade marks as, at their simplest, ‘signs’ used to distinguish goods or services in the market and trade mark language as a social-cultural phenomenon is a thin one. The imagery of those denoted ‘traders’, who employ meaning as a commercial signalling device and commonly register their trade marks, and those denoted ‘authors’, who create copyright works and rarely register trade marks (except sometimes for character merchandising purposes), can often be basically the same. The Coca-Cola Company may be a trader in soft drinks but 60 per cent of its company value lies in its brands. Why? Because its brands and

2 See Roy Harris, Language, Saussure and Wittgenstein: How To Play Games With Words (1988) at ix and further below n125 and ff.
3 John Ayto, Twentieth Century Words (1999) at 1 referring both to new words and new meanings of old words (the latter by far the larger source of increase).
4 All are currently registered in Australia: see trade marks 166522, 555833, 608914 and 608918 (‘Barbie’ registered for dolls, doll clothing and accessories, games and playthings, audiocassettes and entertainment services); 308508 and 736577 (‘Star Wars’ registered for printed matter, interactive entertainment software, computer games, video discs and dvds); 713303 (‘Roller Blade’ registered for in line skates and related entertainment services). Australian registered trade marks are recorded at <http://pericles.ipaustralia.gov.au/atmoss>.
5 ‘The Real Thing’, first adopted as a trade mark in 1942, is a common law trade mark for the Coca Cola Corporation: see Mark Pendergrast, For God, Country and Coca-Cola: The Definitive History of the Great American Soft Drink and the Company that Makes It (1993). Coca-Cola’s applications to register ‘You Know You Want It’ (for the drink) and ‘Enjoy’ (together with ‘Coca-Cola’) were accepted by the Australian Trade Marks Office in 2003: trade mark 938291 and 959805; and ‘Do What You Want’ was registered (for the drink) in 1997: trade mark 751047. Nike registered ‘Just Do It’ (for clothing) in 1992: trade mark 578679. ‘Nike’, itself was first registered (for shoes) in 1975: trade mark 284351.
6 Alex Kozinski (writing exjudicially), ‘Trademarks Unplugged’ (1993) NYU LR 960 at 973. ‘Trade marks’ are ‘signs’ used or intended to be used to distinguish goods or services in the market: see for instance Trade Marks Act 1995 (Cth) s17.
7 This is a point I made (in a more fledgling form) some years ago: see Megan Richardson, ‘Copyright in Trade Marks? On Understanding Trade Mark Dilution’ [2000] 1 IPQ 66.
the images they conjure are an important part of what it sells. Eminem may be a rap musician but he is also a trader selling his music, his messages, his personality and his brilliant use of language. His repeat phrase, ‘one shot — one opportunity’ used in the semi-biographical film 8 Mile\(^{10}\) have as good a claim to trade mark status as Coca-Cola’s ‘The Real Thing’. They mark him out, they distinguish his work in the market of music and entertainment: they are his signature piece. Mike Leigh’s use of ‘Secrets and Lies’ as the title (signature again), theme and key passage in the film Secrets and Lies\(^{11}\) has fixed those words with his brooding significance. The Star Wars films are as much identified with their ‘Empire Divided’ and ‘Dark Side’ and ‘Star Wars’ language as their plots and characterisations;\(^{12}\) Joanna Rowling’s ‘Hogwarts’\(^{13}\) is fixed in the memories of a generation of 12 year olds;\(^{14}\) and Peter Jackson made ‘Middle Earth’ famous by his faithful depiction in films of Tolkien’s The Lord of the Rings.\(^{15}\) Artists build their reputation around minimalist use of powerful words — as with Colin McCahon’s ‘I Am’,\(^{16}\) Brett Whiteley’s ‘It’,\(^{17}\) and Sam Durant’s ‘Like Man, I’m Tired of Waiting’.\(^{18}\)

In a world where trade mark is replacing copyright ‘as the favoured form for protecting cultural texts’\(^{19}\) it can no longer be said, if ever it could, that trade marks exist only in relation to goods or services and have no meanings in themselves.\(^{20}\) The 20th century has seen popular culture take art, music and poetry into the ‘street’ of film, radio, television,\(^{21}\) and now even more pervasively the internet. Ordinary people (including importantly children) have money and leisure and independence to devote to their entertainment\(^{22}\) — and they are prepared to spend these on things that draw them together, creating

---

10 8 Mile, Director Curtis Hanson, Starring Eminem, Universal Pictures, 2002.
12 See the official Star Wars website at <http://www.starwars.com/>.
13 The trade mark is registered by Time Warner for inter alia paper products and entertainment services: 806048.
14 Is Hogwarts derived from Hogwash? Hogwarts or Hogwash? is the title of a book by Sydney authors Peter Furst & Craig Heilmann (2001) which seeks to help parents respond to the ‘Biblical’ messages of the JK Rowling series.
15 As acknowledged by Tourism New Zealand which has since sold New Zealand as ‘Middle Earth’: <http://www.newzealand.com/homeofmiddleearth/>.
18 Sam Durant, Like Man, I’m Tired of Waiting, Lightbox, California, 2002.
20 As said in Exxon Corporation v Exxon Insurance Consultants International [1982] Ch 119 at 143 (Graham J) (‘Exxon’ not a copyright work); and on appeal Stephenson LJ at 130–1.
22 A radical change from a century ago where 80 per cent of household income was used for basics of food, clothing and shelter: see Robert Fogel, The Fourth Great Awakening and the Future of Egalitarianism (1999).
their own virtual communities based around a network of shared experiences. Now trade marks do more than ‘sell’ goods and services, let alone distinguish their ‘origin’ — still the only true function of trade marks according to trade mark law.23 Like them or not, trade marks tell stories. Their expressiveness is the basis of commercial activity, the trader-author the conduit of meaning, and the market-audience the monitor and arbiter of taste. And the combination of a vivid visual characterisation and aural sound-bite effect makes them easy to remember (and difficult to forget) across the entire base of the population.24 The late Daniel Boorstin of the Smithsonian Institution observed that advertising now shows many of the familiar characteristics of folk culture: ‘… repetition, a basic style, hyperbole and talk talk, folk verse and folk music …[and] a limbo between fact and fantasy’.25 Rosemary Coombe, The Cultural Life of Intellectual Properties author, adds: ‘… mass-media imagery and commodified cultural texts provide important cultural resources for the articulation of identity and community in Western societies …’.26 Popular/folk culture also thrives in a fluid, unregulated environment.27 Yet our trade mark law traditionally operates under the umbrella that the great open commons of the English language require the protective mantle of regulation: ‘trade marks’ should be narrowly defined and thresholds for registration set high in order that the language commons should remain in their pristine natural state; while at the other end infringement of a registered trade mark should be narrowly construed to avoid anything that smacks of copyright style protection. My purpose in this article is to explore how such ideas came to evolve and why they continue to be the rhetorical foundation of the law in a world where so much has changed in the last broad century — especially when, with the opportunity for a fresh start presented by the internet, the opposite approach, one of low-regulation, was adopted with respect to domain names.

23 ‘Selling’ goods by signalling their quality was a ‘true’ trade mark function as remarked upon by Frank Schechter in 1927: ‘The Rational Basis of Trademark Protection’ (1927) 40 Harvard Law Review 813 — challenging the conventional idea that trade marks were simply and properly used to designate trade origin. The conventional idea, however, continues to persist: see, for example, Shell Company of Australia Ltd v Esso Standard Oil Ltd (1963) 109 CLR 407 at 425 (Kitto J); Johnson & Johnson Australia Ltd v Sterling Pharmaceuticals Pty Ltd (1991) 21 IPR 1 at 24–25 (Gummow J).

24 The combination of vivid visual imagery and repetitive sound-bite advertising in an entertaining format combines successful features of language teaching approaches; while the choice of simple word formats and transparent meaning for trade marks makes them ideal subjects for language acquisition: see generally Douglas Reutzel & Robert Cooter Jr, Teaching Children to Read (2000) at 262–3 especially (canvassing recent research on reading instruction approaches) and Eve Clark, The Lexicon in Acquisition (1993) at 109-125 (pointing to transparency and simplicity of language as keys to acquisition).

25 Daniel Boorstin, ‘Advertising and American Civilisation’ in Yale Brozen (ed), Advertising and Society (1974) 11 at 22 – pointing to the change since the 19th Century when the function of advertising was simply to inform. By the early 1920s, Boorstin notes, advertising in the US had taken on the function of persuading – and in a country where ‘a shared, a rising, and a democratised standard of living was the national pride and the national hallmark’ – ‘advertising had become the rhetoric of democracy’: Id at 14.

26 Coombe, above n19 at 57.

27 Boorstin, above n25 at 21, talking particularly of traditional folk culture.
The paper has three main parts. Next, I consider the very different treatment of language when it comes to the regulation of trade marks and domain names. Then, I review the historical context of our registered trade mark system, including the influence of 19th century linguistic theories which treated language as historically pure and fragile to disruptive influences. Finally, I suggest that with the experience of the last 100 and more years, and the benefit of sociolinguistic ideas about the robust flexible character of language development, some rethinking of the fundamentals of trade mark law is in order. In particular, those concerned with freeing the language commons might redirect their attention to the level of regulation in the system. Amid fears that trade mark owners given more rights will naturally seek to exploit expressive components of their trade marks to control and censor popular discourse, the possibility that guiding commercial norms may be a preferable alternative to restrictive regulation has apparently so far been overlooked.

2. Regulation of Trade Marks Versus Domain Names

The internet is a web of interconnected networks, or virtual superhighways, with millions of websites, or virtual shopfronts, which serve as vehicles for trading and dissemination of thoughts and ideas.28 To have a website locus on the internet a domain name must be registered — and the domain name once registered becomes the unique identifier, the signature, or trade mark, of its holder. Registering a domain name is simple. For the most popular domains — those ending in .com, .org and .net — registration is on a first to file basis with no substantive requirements.29 The most important constraint is the practical technical one that a domain name may not be identical to an already registered domain name. (More thresholds apply for registration in the Australian administered .au domains, but even these are not particularly stringent.)30 Also simple is the dispute resolution system if a registration is challenged. Only three grounds of objection may be raised under the uniform dispute resolution process which governs most domain name registrations (including in the .com, .org and .net domains).31 The grounds for objection are cumulative: first, that the domain name is identical or confusingly similar to a trade mark in which the complainant has rights; second, that the registrant has no rights or legitimate interest in the domain name; and third that the domain name was registered and is used in bad faith.32 (The last two may overlap

---

28 Not that it was conceived for this purpose: for a discussion of its origins as a private communication network, see John Naughton, A Brief History of the Future: The Origins of the Internet (1999).
29 See, regarding the .com, .org and .net domains, under the Internet Corporation for Assigned Names and Services (ICANN) <http://www.icann.org>.
30 The Australian .com.au and .net.au domains are administered by the Australian Domain Name Registering Authority (auDA): <http://www.domainregistration.com.au/domain–name–policies–main.php>. The latter requires an Australian connection as well as a prior business name or trade mark registration or an acronym, or an indication of a ‘substantial connection to the registrant’.
31 For a similar dispute resolution system adopted for .au domains see the auDA site ibid.
32 For the full terms of the policy see <http://www.icann.org/udrp/udrp–policy–24oct99.htm>.
Relatively few domain names have been challenged under the UDRP. Over 30 million have been registered since the internet began its commercial life the early 1990s. Of those, approximately 15,000 have been the subject to UDRP disputes since the system was established in 1999; less than 0.05 per cent. The remainder are free to be traded, otherwise used or simply held by the registrant. And since the direct costs of registering and disputing are low (registration in the .com, .org and .net domains set at around $70 (US) for the first two years; and a UDRP single arbitrator dispute costing around $1,500 (US) — and born by the disputant) the result has been an explosion of internet activity since the internet began its public life some 10 years ago.

The domain name system has its flaws. But criticism has tended towards the need for fine-tuning rather than rejection of the basic simple model. By and large, a primary internet policy of the last decade — promoting cheap expeditious trading options for those who wish to exploit the medium commercially — has been realised. And with no great impairment of public debate either. Debate flourishes on the internet and moderately free domain name registration and dispute resolution promotes it. ‘Gripe sites’ like NikeSucks.com, McSpotlight.org, SaveBarbie.com, and SaveHarry.com proliferate on the internet (notwithstanding efforts by some of those named as targets to control their spread) and their evocative names make their purpose immediately apparent to their audiences. In some cases ‘prescient companies’ have taken advantage of the consumer complaints lodged on these sites to change their own practices — enhancing their

---

33 See, for example, Telstra Corporation Limited v Nuclear Marshmallows Case No D2000–0003 (the panel’s conclusion of bad faith rested on an inference from the defendant’s failure to give good cause for the adoption of the claimant’s trade mark ‘Telstra’ as a domain name, that ‘it is not possible to conceive of any plausible actual or contemplated active use of the domain name by the Respondent that would not be illegitimate …’ (Associate Professor Christie at para 7.12); reported at <http://arbiter.wipo.int/domains/decisions/html/2000/d2000–0003.html>.

34 For Network Solutions reporting the 30 millionth registration worldwide (Muzi.com) on 5 October 2000, see Yahoo Finance at <http://uk.biz.yahoo.com/001004/80/ali4g.html> (10 March 2003), numbers in the top five open domains (including .com, .net and .org) totalled 37,383,532: see <http://www.whois.sc/internet–statistics/>.


38 As Rosemary Coome & Andrew Herman also point out with examples: see ‘Trademarks, Property and Property: The Moral Economy of Consumer Politics and Corporate Accountability on the World Wide Web’ (2000) 50 De Paul L. Rev at 609 (‘the Web is certainly used for corporate bashing’) and passim.
reputation for consumer sensitivity in the process.\textsuperscript{39} In other cases they have reacted by engaging in private and public debate.\textsuperscript{40} The sites thus serve an important dialogic purpose in more than one sense, leading potentially sceptical academic observers Rosemary Coombe and Andrew Herman to conclude that:

\begin{quote}
We are seeing a field of power shifting in digital terrain. As systems of univocal proprietary control give way to an interactive ethics that interrogates the claims of property with questions of propriety, we have a unique opportunity to challenge the impositions of privilege and insist upon new forms of responsibility and social accountability …\textsuperscript{41}
\end{quote}

Even the use of such common language terms and expressions as ‘business’, ‘fun’, ‘free’, ‘tired’ and ‘tiredofwaiting’ for domain names\textsuperscript{42} has not emerged as a serious problem on the internet. Contrary to what might have been expected at the beginning of the domain name registration system (based on assumptions carried over from trade mark law that common language must be in public hands if public access is to be maintained), a practice of private ownership and public use has developed. Generic sites are commonly exploited as portals for shared messaging by multiple subscribers.\textsuperscript{43} There are exceptions. For instance, the site tired.com, which carries the words, ‘Are you tired? Tell us why’ has no evident public benefit. However, rather than enclosing the language, tired.com has merely become the subject of public criticism labelled one of the most pointless websites on the internet today.\textsuperscript{44}

The simplicity of the domain name system is in contrast to registered trade mark systems. These nationally based systems are often complex and laborious, even in their most recent incarnations. In the 1990s with the World Trade Organisation’s TRIPs Agreement linking intellectual property rights to international trade\textsuperscript{45} attention turned to harmonisation of national law standards. But barriers to establishment and enforcement of rights have not yet been fully

\textsuperscript{39}Id at 610–611, giving the example of Dunkin Donuts introducing low-fat milk after complaints were lodged on a customer’s \texttt{<http://www.dunkindonuts.org>} website. Another example is Apple’s introduction of expanded warranty for iPod owners and reduction of its iPod battery replacement price after a wrathful ‘gerilla film’ about the company’s practices was published by multimedia artists Van and Casey Neistat on \texttt{<www.ipodsdirtysecret.com>}: ‘The Sonic Bust’ \textit{The Age} (11 January 2004) Agenda at 12 (reprinted from the \textit{Washington Post}).

\textsuperscript{40}See Michael Brush, ‘Got a Gripe: There are Plenty of Places to Vent on the Net’ for examples and advice at \texttt{<http://www.mcspotlight.org/media/press/moneyonline_13jul96.html>}.

\textsuperscript{41}above n38 at 626.

\textsuperscript{42}See, for example, \texttt{<http://www.business.com>} (business search engine and directory); \texttt{<http://www.free.com>} (lists over 8000 internet sites offering free products and services); \texttt{<http://www.fun.com>} (a self-styled ‘comprehensive guide to entertainment and the arts’); \texttt{<http://www.tiredofwaiting.com>} (an ‘all in one web and domain resource’).

\textsuperscript{43}Generic domain names were not even included in the World Intellectual Property Organisation review of the domain name dispute process (nor, unlike indigenous expressions, identified by respondents as an emerging problem): see \textit{Report of the Second WIPO Internet Domain Name Process} at \texttt{<http://wipo2.wipo.int/process2/report/index.html>} (29 March 2004).

\textsuperscript{44}Dubbed ‘Worst of the Web’ Winner, October 2002: \texttt{<http://www.worstoftheweb.com/classic/102202.html>}. 
addressed. A particular regulatory target is the trade marking of language. Legislators and courts in the British mould have traditionally identified ‘inherent distinctiveness’, said to be found in the ‘nature’ of the trade mark irrespective of its use, as an important registration threshold. From a modern linguistic perspective, the notion that any meaning is ‘inherent’ in words is impossible: meaning is a product of social discourse. But in fact it is only use or prospective use ‘as a trade mark’ that may not be considered in assessing inherent distinctiveness, courts have sometimes added. Of course, linguists might also point out that ignoring or marginalising trade mark use leaves out a practical source of language development. And experienced lawyers may note that recognition of common law trade marks, protected from passing off on the basis of their reputation (and without a requirement that the origin of goods or services be indicated) is not so circumscribed. Lack of inherent distinctiveness has not been a bar to common law trade mark status since ‘Camel Hair Belting’ (used for belting made of camel hair) was acknowledged by the House of Lords as giving rise to an action for passing off in the 1896 case of Reddaway v Banham. But in the unreal world of trade mark law inherent distinctiveness is still the preferred standard held to by courts.

The preference is now subtler than before. When inherent distinctiveness decided a trade mark’s ability to be registered — as before the Australian Trade Marks Act 1995 (Cth) and the UK Act of 1994 — results were subjective, unpredictable and sometimes bizarre. For instance, ‘Solio’ was accepted by the House of Lords as inherently distinctive for photographic paper in 1898 (when the sun was employed in the process of developing photographs); as was ‘Tub

---


46 Although some potentially important changes were made, including a new definition of ‘trade mark’ as a sign used to distinguish goods or services in the market without specific reference made to ‘origin’, a broader set of criteria for assessing ‘distinctiveness’, and extension of infringement to cover identical or similar trade marks used for goods or services outside the original scope of registration (where such use would lead to likely confusion): see especially Articles 15 and 16.

47 Registrar of Trade Marks v W & G Du Cros Ltd [1913] AC 624 (hereinafter W & G Du Cros) at 634–5 (Lord Parker); Burger King Corporation v Registrar of Trade Marks (1973) 128 CLR 417 (hereinafter Burger King) at 424 (Gibbs J) (“[i]nherent adaptability” depends on the nature of the trade mark and is ‘not something that can be acquired’). See also (no difference between common language and geographical terms), Clark Equipment v Registrar of Trade Marks (1964 111 CLR 511 (hereinafter Clark Equipment) at 512–4 (Kitto J).

48 As Kitto J makes clearer in Clark Equipment, id (the question is the ‘plaintiff’s chances of distinguishing his goods by means of the mark, apart from the effects of registration’). Compare W & G Du Cross ibid (Lord Parker).

49 For common law purposes reputation is a source of rights: see Reckitt & Colman Products Ltd v Borden Inc [1990] RPC 341 (hereinafter Reckitt & Colman) at 406 (Lord Oliver) (although referring to ‘identifying get-up’ rather than simply trade marks as importantly associated with reputation). Lord Oliver’s classical formulation of passing off’s elements of reputation, misrepresentation and damage has been approved many times, including in ConAgra Inc v McCain Foods (Australia) Pty Ltd (1992) 23 IPR 193, at 246–8 (Gummow J).
Happy’ for cotton clothing (being presumably easy to wash in the tub) by the Australian High Court in 1956, notwithstanding a dissent by Kitto J who pointed out that coining happy expressions was a common form of slang at the time. However ‘Perfection’ and Orlwoola’ were disallowed for soap and cloth respectively in 1909; as were ‘Ribbon’ for toothpaste in 1913, ‘Charm’ for hosiery in 1928, and (in Canada) ‘Superweave’ for textiles in 1949. And (Burger King’s) ‘Whopper’ was deemed not inherently distinctive of hamburgers in Australia in the early 1970s — other traders might ‘without improper motive’ wish to use the term to describe their ‘hamburger sandwiches’, Gibbs J said, although there was no evidence they had or intended to do so. Now, in the post-TRIPs era more jurisdictions follow the US in holding factual distinctiveness sufficient in principle. But with the threshold left to courts less has changed than might be supposed. It has been suggested 90 per cent recognition of the trade mark as a trade mark would be needed in order for the new threshold to be satisfied. The trade mark would have to be ‘almost a household word’. And it would need to

50 Reddaway v Banham [1896] AC 199 (the term accepted as a common law trade mark for industrial belting – setting aside the Court of Appeal’s judgment reported at [1895] 1 QB 286 which in accordance with previous authority held that “[w]hatever is mere description is open to all the world’; at 298 (Rigby LJ) quoting Cheavin v Walker (1876) 5 Ch D 850, at 863 (James LJ)). See also Erven Warnink BV v J Townend [1979] AC 731 (public recognition of Advocaat as a ‘descriptive’ term distinguishing brandy-egg combination liqueur by its quality, without indicating origin or source, was sufficient basis for a common law reputation). As Lord Oliver said in Reckitt & Colman, above n49, ‘even a purely descriptive term consisting of perfectly ordinary English words may, by a course of dealing over many years, become so associated with a particular trader that it acquires a secondary meaning such that it may properly be said to be descriptive of the trader’s goods and of his goods alone’ (at 412).


52 Mark Foys Ltd v Davies Coop & Co Ltd (1956) 95 CLR 190, at 201 (Williams J) (the term conveys only a ‘vague’, ‘covert and skilful’ allusion to the clothing’s washability), see also Dixon J at 193–95. (For a modern comparison see The Procter & Gamble Co v Office for Harmonisation in the Internal Market (Trade Marks and Designs) [2002] Ch 82 (ECJ): ‘Baby Dry’ considered inherently distinctive for nappies.)

53 Joseph Crosfield and Sons’ Application (1909) 26 RPC 837; In particular, Cozens-Hardy MR observed that, notwithstanding 30 years extensive use, ‘[p]erfection’ was used adjectively and was ‘a mere laudatory epithet’ – thus ‘I think it would be wrong to allow a monopoly in the use of such a word’ (at 855). Further, the judge added, as far as ‘Orlwoola’ is concerned the misspelling of ‘all wool’ (properly disclaimed) did not render it distinctive (at 856).

54 In Re Colgate & Co’s Application (1913) 30 RPC 262.

55 In Re Keystone Knitting Mill’s Ltd’s Application (1928) 45 RPC 421.

56 Registrar of Trade Marks v GA Hardie & Co Ltd [1949] 4 DLR 582.

57 See Burger King, above n47.

58 And notwithstanding that by then the Trade Marks Act 1955 (Cth) allowed for registration on the basis of ‘capability to distinguish’, and identified inherent distinctiveness as only a factor to be considered along with secondary reputation (ss 25 and 26): in practice courts continued to use the same negative test (need for trader and public access to common language) to hold the threshold not satisfied as for cases where inherent distinctiveness was the only criterion.

59 For instance, Trade Marks Act 1995 (Cth) s41(6); Trade Marks Act (1994) UK s3(1). The UK Act also implements standards set down in the European Trade Marks Directive No 89/104/EEC.

60 British Sugar Plc v James Robertson & Sons Ltd [1996] RPC 281 at 305 (Jacob J). See also Ocean Spray Inc v Registrar of Trade Marks (2000) 47 IPR 579 (hereinafter Ocean Spray) at 587 (Wilcox J).
be recognised as a ‘trade mark’ denoting the origin of goods or services.\textsuperscript{61} Thus ‘Whopper’ with virtually 100 per cent public recognition as Burger King’s ‘trade mark’ is now registered in Australia.\textsuperscript{62} But ‘Have a Break’ was found insufficiently distinctive of Nestlé’s chocolate when the company recently sought registration in the UK, proffering evidence of common law trade mark recognition.\textsuperscript{63} (The question whether it should be registered is currently before the European Court of Justice).\textsuperscript{64} And Ocean Spray’s inability to show it had used ‘Cranberry Classic’ as more than an identifying descriptor of its fruit drinks barred it from the Australian trade marks register in 2000.\textsuperscript{65} The possibility of secondary meaning establishing distinctiveness is now conceded in principle. But in practice it is inherent distinctiveness that is still a deciding consideration in almost all cases,\textsuperscript{66} albeit registrars and courts may now be more flexible than once they were in applying that standard.\textsuperscript{67} The premise is still that regulating entry into the system is preferable to simply determining whether common law trade mark status has been or will be achieved (or more simply, as with domain name registration, deciding conflicting claims).\textsuperscript{68}

What is the purpose of these limitations? Williams J said in the Tub Happy case that protecting the English language commons is an essential attribute of our trade mark law, quoting Lord Herschell in the Solio case.\textsuperscript{69} The idea seems so at odds with current linguistic thinking that the history of its apparently inexorable hold on the law warrants closer examination.

\textsuperscript{61} See especially Ocean Spray, id at 588–9 (Wilcox J).
\textsuperscript{62} Trade Mark: 751333.
\textsuperscript{63} Societe des Produits Nestlé SA v Mars UK Ltd [2002] All ER (D) 05. There was survey evidence supporting public recognition but Rimmer J placed weight on Nestlé’s past use of ‘have a break’ with ‘have a Kit Kat’ as supporting its merely descriptive use: at para 32.
\textsuperscript{64} The question whether, for registration under the UK Trade Marks Act 1994, distinctiveness could be acquired ‘following or in consequence of the use of that [ie the claimed] mark as part of or in combination with another mark?’ was referred by the Court of Appeal to the European Court of Justice: Societe des Produits Nestlé SA v Mars UK Ltd [2003] EWCA Civ 1072.
\textsuperscript{65} Ocean Spray, above n60.
\textsuperscript{66} For registrations too – for example, a sample of 2,000 of 14,219 Australian registered trade marks sealed in the period 1 July 2003–31 December 2003 reveal two (‘Commonwealth Games’ and ‘Rowing World Cup’) endorsed as registered on the basis of factual distinctiveness (secondary reputation) alone.
\textsuperscript{67} For instance, it is unlikely expressions like ‘Just Do It’ and ‘Do What You Want’ referred to above at n5 would have been considered sufficiently inherently distinctive to warrant registration even 30 years ago when ‘Whopper’ was disallowed (above n47); yet they are registered in Australia. For a survey of 300 recently registered trade marks classified by the Australian Trade Marks Office as inherently distinctive finding up to one third drawing on common language, see Megan Richardson, ‘Trade Marks and Language: Welfare Effects’, WTO Conference, Adelaide, February 2004 (copy on file with author).
\textsuperscript{68} Trade mark law makes provision for conflicting claims as well: see for instance ss44 and 60 Trade Marks Act 1995 (Cth) elaborating on provisions found in earlier Australian and UK statutes; and Southern Cross Refrigerating Co v Townoomba Foundry Pty Ltd (1954) 91 CLR 592 (‘Southern Cross’ for refrigerators and well-drilling equipment) and Berlei Hestia Industries Ltd v The Bali Company Inc (1975) 129 CLR 353 (‘Berlei’ and ‘Bali’ for bras).
\textsuperscript{69} Mark Foys Ltd, above n52 at 199. The reference is to Eastman Photographic, above 51 at 580.
3. Importance of History in the Trade Mark System

Our trade mark system is a product of history.\footnote{70 For a fascinating review, see generally Brad Sherman & Lionel Bently, The Making of Modern Intellectual Property Law (1999).} When the first comprehensive British trade marks registration system, the \textit{Trade Marks Registration Act 1875},\footnote{71 \textit{Trade Marks Registration Act 1875} (UK) 38 & 39 Vict, c91.} was adopted there were designs and patent registration models in place which could easily be adapted to the purpose of determining the character of the system adopted.\footnote{72 See \textit{Report of Select Committee on Trade Marks Bill} (London: Parliament House of Commons, 1862) at 26 (William Smith – Sheffield Chamber of Commerce Secretary and drafter of the ‘Bill No 1’, which subsequently became the 1875 Act).} The fact that trade mark registration was available in some European countries,\footnote{73 The Superintendent of Patents (Woodcroft) testified to the Select Committee in 1862 that Patent Office registered patents, designs and trade marks in Austria, Bavaria, Belgium, France, Hanover, The Netherlands, Portugal, Prussia, Russia, Sardinia (Italy), Saxony, Spain, Sweden, Norway, Prussia, Wurttemberg (Germany) and America: \textit{id} at 71.} and it was hoped granting like rights in Britain would lead to reciprocal rights for British traders abroad, added the weight of public opinion — at least of those traders consulted by the Select Committee on Trade Marks in the lead up to the 1875 Act.\footnote{74 They submitted registration should involve a process of examination of the trade mark; and the registrar should have the authority to reject a claim on the ground of insufficient distinctness from any existing mark, as in Prussia.\footnote{75 See especially \textit{id} at 36 (Ryland). See also \textit{William Smith} at 27 (suggesting an opposition procedure was also desirable to allow for full monitoring).} The Select Committee questioned whether such a bureaucratic system adopted for trade mark registration ‘would suit the habits and feelings of people of this country’.\footnote{76 \textit{id} at 36.} But, ultimately (although it took some 13 years), the Bill was enacted along the general lines sought.\footnote{77 \textit{Trade Marks Registration Act 1875} 38 & 39 Vict. The Bill took some 20 years to be enacted, possibly because its more radical aspects including provision for registration of trade marks, for assignment and for a single trade mark (registration) system applying to ‘all trade marks of all the various and very numerous manufactures of the United Kingdom’ troubled the Committee: see \textit{id} at 27–33 (interviewing William Smith, drafter of the Bill). In the meantime another Bill, which also was considered by the Select Committee (as ‘Bill No 2’) and provided more narrowly for criminal remedies against fraudulent applications of trade marks, was enacted as the \textit{Merchandise Marks Act 1862}: see at 128–139 (Hindmarsh, drafter of Bill No 2, explaining its terms and comparing it with Bill No 1).} The new Act also specified a registered trade mark as ‘belonging to particular goods, or classes of goods’.\footnote{78 \textit{Trade Marks Registration Act 1875} s2 (Characteristics of a registered trade mark).} This was considered necessary to reflect the nature of a ‘trade mark’ as understood and employed at the time, \textit{viz} to denote manufacturing origin.\footnote{79 John Smith, Sheffield Chamber of Commerce President, told the Select Committee that:}
A man might put any label he pleases upon his goods … descriptive of their quality, or recommending them in any way he pleases; but in addition to this description, I think there should be a trade mark that indicated the individual house from which the goods came.\textsuperscript{80}

Further, it was specified that to qualify for registration, a trade mark must consist of certain ‘essential particulars’ (something not provided for in exactly those terms in the Bill as considered by the Select Committee).\textsuperscript{81} Initially this requirement meant that the name, word or device chosen should be represented in a special and particular way. But later, in the \textit{Patents, Designs and Trade Marks Act} of 1883, the list was extended to include fancy or invented words and the like (including newly coined expressions of a non-descriptive character).\textsuperscript{82} Even with the distinctiveness threshold expressed in such embryonic form, comparison could be made with design law.\textsuperscript{83} In both cases more than simply copying of features previously adopted with respect to the articles in question was needed.\textsuperscript{84} As accepted for registration of designs, it would seem the ‘more’ required ‘an exercise of intellectual activity so as to originate, that is to say, suggest for the first time, something that has not occurred to anyone before’.\textsuperscript{85} To give effect to such standards an examination and opposition procedure was adopted for registration of a trade mark, like the procedure adopted for registering a design or patent,\textsuperscript{86} and care and control of the trade marks register was given to the Commissioners of Patents.\textsuperscript{87}

\textsuperscript{80} Select Committee on Trade Marks, above n74 at 57. Compare and contrast John Smith’s later identification of a trade mark as effectively ‘equivalent to a man’s signature to a letter’: id at 58.

\textsuperscript{81} Trade Marks Registration Act 1875 s10. The language of the provision is more restrictive than in the Bill, as drafted by William Smith, which provided that a ‘trade mark’ might ‘include any name, word, letter, mark, device, figure, sign, seal, stamp, label or other thing duly registered as herein provided’: see Select Committee on Trademarks, above n72 at 129 (Hindmarsh) and also at 28 (Smith) (the words ‘trade mark’ should include a trader’s name – and the name and a symbol combined are ‘the best form of trade mark’).

\textsuperscript{82} See Patents, Designs and Trade Marks Act 1883 46 and 47 Vict, c57 s64. In the \textit{Trade Marks Act} 1905 (UK) the latter requirement was slightly modified to allow registration of words having no ‘direct’ reference to the character or quality of the goods: s9 and compare Trade Marks Act 1905 (Ch 161)d).

\textsuperscript{83} As observed by Sherman and Bently, above n70 at 168–70, it was the design system that served as the particular model for the new trade mark registration system – and this (along with the patent system) ‘provided an important point of analogy for the existence of rights, as an aid to interpreting trade mark doctrine, for the shape the system of registration should take, and the language and structure to be used in trade mark legislation’.

\textsuperscript{84} As with trade marks, originally only able to be registered with respect to goods (registration of service trade marks came later), designs were from the beginning registered with respect to ‘articles’ of manufacture: see Designs Registration Act 1839 2 & 3 Vict, c17 s1.

\textsuperscript{85} In Re Clarke’s Design [1896] 2 Ch 38 at 30 (Lindley LJ).

\textsuperscript{86} The opposition procedure was formalised in Patents, Designs and Trade Marks Act 1883. Part IV. But in the case of trade marks under the \textit{Trade Marks Registration Act} 1875 the practice was already adopted of advertising applications for registration in the Trade Marks Journal and of giving consideration to notices of objection before registration: see, for example, In Re Worthington & Co’s Trade-Mark (1880) 14 Ch D 8.

\textsuperscript{87} Trade Marks Registration Act 1875 s1.
There were other common points with the designs and patent systems. In particular, once registration was granted, the 1875 Act provided for a right of ‘exclusive use of such trade-mark’ — analogous to the rights granted in registered designs and patents and quite different from common law trade mark protection which was limited to conduct likely to deceive or confuse. Thus the emergent idea of an intellectual property right as entailing an exclusive right, not simply a more easily recognised and transmissible version of the common law right as previously thought by some during the Select Committee hearings, was given legal effect. Trade marks like patents and designs were classified as industrial property in the Paris Convention of 1883, an event which legal historians Brad Sherman and Lionel Bently observed marked the ‘crystallisation’ of the legal categories. Fifteen years later, the idea of an exclusive property right being obtained in a registered trade mark that was sufficiently innovative to qualify as distinctive for statutory purposes was seen as entirely logical by courts. As Lord Hershell posited in the Solio case:

‘If a trade mark is invented “what harm is done, what wrong is inflicted, if others be prevented from employing it, and its use is limited in relation to any class or classes of goods to the inventor?”’

The question was rhetorical. By then it was simply accepted that the corollary of the innovation threshold adapted across from the designs/patent systems was the granting of an exclusive ‘property’ right in a registered trade mark.

Why was the design/patent model followed so slavishly in trade mark law? We can already posit that, from an administrative perspective, it must have seemed safer and simpler to take the model being evolved for designs and patents and adapt it for trade marks rather than try to come up with an entirely new model — that appears to be a conclusion of Sherman and Bently as well. Certainly, once
the need for a registration system was accepted, there was no consideration of any alternative approach to that proposed in the Select Committee hearings or 13 years later when the Trade Marks Bill 1875 was enacted: it was introduced in June and passed by August with no further discussion.96 And among some of those responsible for the British legislation there did seem to be a rather simplistic view that trade marks, at least those worthy of registration, were rather like designs and inventions and to be protected for similar utilitarian reasons: stimulating innovative activity (seen as important in a country in the midst of the industrial revolution, and appreciative of its benefits). At the same time, conceiving of the ideal trade mark as a creative embodiment enabled the 19th century lawmakers to meet concerns about the costs of restricting competition in a strongly laissez-faire economy. The benefits of some restraints on untrammeled competition in order to reward and promote innovation had already been accepted for patents and designs;97 now the reasoning could be extended to trade marks as well — albeit without much thought being given to the social value of the ‘innovation’ concerned.

Finally, a particular benefit of a distinctiveness-cum-innovation threshold for trade marks was identified in a report issued on the workings of the Patents, Designs and Trade Marks Act 1883,98 where it was said that:

> Words are, undoubtedly, a most popular form of trade-mark, but some limit must obviously be put upon the words which an individual may be permitted to register and claim the exclusive use of … [N]o one ought to be granted the exclusive use of a word descriptive of the quality or character of any goods. Such words of description are the property of all mankind, and it would not be right to allow any individual to monopolise them and exclude others from their use.99

And similar sentiments were expressed in the Solio case, by Lord Herschell: ‘The vocabulary of the English language is common property: it belongs alike to all’.100

---

96 As recorded in Hansard’s Parliamentary Debates (3rd Series, 39 Vic, 1875).
97 Including by leading utilitarians of the day including Jeremy Bentham and John Stuart Mill, as well as (some years earlier) Adam Smith. All favoured a limited term of monopoly in exchange for the ingenuity and effort expended in developing a profitable invention, seen as an engine of the British economy’s growth and development – primarily these writers were talking about patents here. For a careful survey of contemporary opinion, see Harold Dutton, The Patent System and Inventive Activity During the Industrial Revolution 1750–1852 (1984) at 17–29 (concluding ‘inventive activity was associated with progress as well as private profit, and this probably explains why the argument was so popular during the early nineteenth century’); and further M Richardson ‘Owning Secrets: Property in Confidential Information?’ in A Robertson (ed), The Law of Obligations: Principles and Boundaries (2004) at 145.
98 Report of a Commission Appointed by the Board of Trade to Inquire into the Duties, Organisation and Arrangements of the Patent Office under the Patents, Designs and Trade Marks Act 1883, so far as relates to Trade Marks and Designs (81 PP C 5350, 1888).
99 Id at 21. The statement is quoted (with approval) by Halsbury LC in Eastman Photographic, above n51 at 574.
100 Eastman Photographic, above n51 at 580. Compare id at 574 (Lord Halsbury LC) and 583 (Lord MacNaghten). See also Joseph Crosfield, above n53 at 854 (Cozens-Hardy MR), in a passage more openly cynical as to the motives of traders: ‘Wealthy traders are habitually eager to enclose parts of the great commons of the English language and to exclude the general public of the present day and of the future from access to the enclosure’. 
The concern was not just that competition might be hindered if other traders would be denied the ability to employ common language terms to describe and extol their products. Maintaining a language commons that would be freely accessible to the community at large was seen as linguistically essential at the end of the 19th century. Linguistics — the systematic study of language development — had by then become more than a field of academic study: it had captured the public imagination in the same way as philosophy and economics had done so the century before (and continued to throughout the period). Various explanations have been offered of the rise, including the growth of universities and popular education, the establishment of learned societies, the establishment of periodicals, even the publication of the *Oxford English Dictionary* in 1860. Whatever the explanation, it is clear the study of language took a ‘momentous turn’ in the course of the century. And historical comparative linguistics, with its perception of language as a product of refinement taking place over centuries, dominated the discourse. Premising trade mark registration on a degree of real innovation was the way lawmakers had to alleviate fears about the effect of exclusive trade mark rights on the ‘great open commons’ of the English language; a language whose central significance for the process of intelligent thought, the articulation and communication of information and ideas, and the maintenance of a cultural tradition was the subject of serious study. Compare thus the words quoted earlier from the 1883 trade marks report and *Solio* case and those of influential German linguistic scholar Wilhelm Von Humboldt in *Über die Verschiedenheit des Menschlichen Sprachbaues* (1836): ‘A people’s speech is their spirit, and their spirit is their speech.’ Compare also the words of the English language commentator, Archbishop Richard Trench, in his popular set of lectures *On the Study of Words* first published in 1851: ‘[Language] is the embodiment, the incarnation … of the feelings and thoughts and experiences of a nation ….’ With such views commonly held to in the mid-19th century, and with a still limited understanding of the process of language development, it was

101 Albeit this was a primary concern of some judges: see, for example, above n58 and n100.


103 See generally Robert Evans, *The Language of History and the History of Language* (1998) at 21–29. Harris, above n2 at 8 also points to the ‘quasi-mystical’ (religious) approach to language, which had developed by the mid–19th Century.

104 Evans, above n103.


106 This is Robbins’ translation: see Robbins, above n102 at 165. (Von Humboldt’s words ‘[i]hre Sprache ist ihre Geist und ihre Geist ist ihre Sprache’ (Humboldt, above n105 at 7) are more literally ‘[y]our speech is your spirit, and your spirit your speech.’)

107 Richard C Trench, *On the Study of Words* (1851), began as a series of lectures at the Diocesan Training School, Winchester. It was republished in many editions including the one from which the following quotation is taken (1890 ed).
impossible to imagine common language expressions being registered let alone used as a trader’s exclusive proprietary mark. Indeed, even for common law trade marks it took some 50 years more before secondary reputation established in otherwise common terms could be recognised as sufficient to mount a passing off action against deceptive or confusing use — a change which was evidently more difficult still to effect in the case of the more elaborately articulated and conceptually entrenched registration system.

In summary, it would seem that for a multitude of reasons the British registration system was not adopted and was certainly not applied to offer protection to what later were accepted as common law trade marks — those which had become associated with trading goodwill/reputation (which, after some initial uncertainty, were able to be protected under the common law passing off doctrine). Nor was its purpose to protect marks that predictability would emerge as common law trade marks in the future. Rather, as Lord Parker noted in W & G Du Cros Ltd in 1913, registration was permitted for ‘certain kinds of mark (somewhat limited in number) which the Legislature considered appropriate for use as trade marks’. The aim was to foster the generation of ideal trade marks, words or phrases whose creativity was sufficient to take them out of the common pool and avoid any risk of destruction of the natural language. Correlatively, any trade mark that was able to meet trade mark law’s inherent distinctiveness standard was assumed to be safe for and indeed entitled to have full and exclusive ‘proprietary’ rights. It was only later that infringement came to be treated by courts as essentially a hypothetical version of the common law passing off action. And since the right to claim infringement was for a long time restricted to goods, and later services, covered by the trade mark registration (a restriction only abolished in part when required by TRIPs), this led to the illogical position of a trade mark system harder to access yet scarcely broader in protection than its common law counterpart.

108 Id at 29, adding, more floridly, that ‘[l]anguage is the amber in which a thousand precious and subtle thoughts have been safely embedded and preserved. It has arrested ten thousand lightning flashes of genius, which, unless thus fixed and arrested, might have been as bright, but would have also been as quickly passing and perishing, as the lightning’.

109 See Reddaway v Banham, above n50.

110 After some initial hesitation (the Trade Marks Registration Act 1875 was unclear, stating simply that an infringement action could not be brought without registration of the trade mark: s1; and deemed registration to be equivalent to public use: s2). But see, for example, Great Tower Street Tea Co v Langford (1888) 5 RPC 66; Faulder v Rushton (1903) 20 RPC 477, and Reddaway v Banham, above n50; and generally William R Cornish, Intellectual Property (2nd ed, 1989) at 395.

111 W & G Du Cros, above n47.

112 Id at 636.

113 Infringement terms were not spelt out in the early British Acts, but were left to courts, only codified in the Trade Marks Act (1938) UK to require the use of a trade mark identical with the registered mark or one so nearly resembling it as to be likely to deceive or cause confusion (s4). In the case of the Australian legislation statutory codification occurred earlier and ‘substantially identical’ was used in place of ‘identical’: Trade Marks Act 1905 (Cth) s53.
The above position has largely been maintained across the British Commonwealth common law world. Only the United States went a different way in earlier recognising simple status as a common law trade mark as the basis of a trade mark’s entitlement, an approach later codified in the Lanham Act 1946. At the same time the Lanham Act for a long time granted little different protection to registered trade marks than was available at common law. But in 1996 the US legislature endorsed trade mark ‘dilution’ as a basis of a separate claim under an amended Lanham Act (again, without restriction to registered trade marks although protection is limited in another way: viz to trade marks deemed ‘famous’) and before that many US states had their own anti-dilution laws.

---

114 By about the middle of the 20th century courts had generally come to assume that the identical or substantially identical basis for claiming infringement was there for a narrow class of cases where likely confusion could be presumed without the need for this to be proved: see, for example, Shell v Esso, above n23 at 414–5 (Wendeyer J): in assessing substantial identity, a close side by side comparison of similarities and differences should be undertaken – a trade mark had to be virtually literally identical to satisfy the standard. Conversely, the deceptively similar/likely confusion basis for claiming infringement emerged as a flexible impressionistic standard, and a hypothetical one at that: see ibid and further Southern Cross Refrigerating Co v Towoomba Foundry Pty Ltd (1954) 91 CLR 592 at 595 (Kitto J) (the question is what can the registrant do with the trade mark once registered).

115 See above n46.

116 Including in Australia which, even before enactment of the first Commonwealth legislation in 1905, modelled on the UK Act of 1905, found state law counterparts to the UK Acts of 1862 and 1875 – see, for example, Trade Marks Act 1865 (NSW) and Trade Marks Act 1876 (Vic).

117 In part the reasons were constitutional. The US Constitution 1788 did not specify trade marks as ‘intellectual property’; and when the need for federal legislation became apparent the Supreme Court refused to consider trade marks to be analogous to copyrights or patents in the early Trade-Marks Cases 100 US 82 (1879) (concluding ‘the ordinary trade mark has no necessary relation to invention and discovery’), leaving their treatment under the Constitution’s commerce clause open. The Australian Constitution 1901 included trade marks within its intellectual property head of power: s51(xviii). The High Court has noted ‘the efforts in coining a new trade mark’ were sufficient to warrant the general epithet ‘intellectual effort’: Grain Pool of Western Australia v Commonwealth (2000) 46 IPR 515 at 527 (Gleeson CJ, Gaudron, McHugh, Gummow, Hayne & Callinan JJ).

118 Lanham Act § 2, 15 USC § 1052.

119 Thus use is a requirement for trade mark protection: see Lanham Act § 1, 15 USC § 1051 (and although actual use may be delayed for up to 36 months, any delay past 12 months requires ‘good cause’). Also, infringement has the same focus on likely confusion in the case of registered as unregistered trade marks Lanham Act § 32 and 43(a), 15 USC § 1114 and 1125(a) and see Polaroid Corp v Polaroid Electric Corp 287 F 2d 492 (1961) at 495 (factors relevant to assessing likelihood of confusion). See generally J Thomas McCarthy, McCarthy on Trademarks and Unfair Competition (4th ed, 2003).

120 Lanham Act § 43–45, 15 USC §1125–1127; ‘dilution’ defined as ‘lessening of the capacity of a famous mark to identify and distinguish goods or services, regardless of the presence or absence of... likelihood of confusion, mistake or deception’. See J Thomas McCarthy, ‘The 1996 United States Federal Anti-Dilution Statute’ (1998) 9 AIPJ 38. See also (for the domain name field) Anticybersquatting Consumer Protection Act (1999) § 3002, 15 USC § 1125(d)(1A).

Continental European countries have also traditionally recognised a wide range of harms as covered by their trade mark and unfair competition laws, not only those that depend on likelihood of confusion, but their choices could more easily be dismissed as antithetical to the concerns of pragmatic common law jurisdictions (especially outside the UK whose European Community membership imposes obligations to harmonise trade mark law to meet European standards). The latest development put new pressure on British Commonwealth countries to reassess the fundamentals of trade mark law, with courts invited to construe existing trade mark infringement provisions to extend to dilution and the question of legislative law reform openly raised on more than one occasion.

4. Mapping the Future: Towards a Market in Language

In the context of today’s modern world should our trade mark system continue along historical lines? There are several reasons to think it should not.

First, modern sociolinguistic studies of language development have now established language as a far more resilient, adaptable and interesting creature than the historical linguistic scholars imagined. From the early 20th century on, writers such as Franz Boas at Columbia, Edward Sapir at Yale and anthropologist Benjamin Whorf began to observe the dynamic relationship between language and...
There are ‘important interconnections … between language, culture and psychology’, Whorf observed in 1941, referring to his mentor and teacher Sapir. (Whorf himself went even further than Sapir in arguing not just the surface elements of a language, its vocabulary, but its structure, or grammar, are an important shaper of ideas.) In 1929 Sapir had already noted that ‘language is becoming increasingly valuable as a guide to the scientific study of a given culture’. Four years later he added:

Changes in vocabulary are due to a great variety of causes, most of which are cultural rather than of a strictly linguistic nature. The too frequent use of a word, for instance, may reduce it to a commonplace term so that it needs to be replaced by a new word. On the other hand, changes of attitude may make certain words with their traditional overtones of meaning unacceptable to the younger generation, so that they tend to become obsolete. Probably the most important sources of changes in vocabulary is the creation of new words on analogies which have spread from a few specific words … [as well as] borrowings of words across linguistic frontiers. This borrowing naturally goes hand in hand with cultural diffusion.

Such observations identify important features of language not previously well understood: its adaptability in the face of cultural change, its inclination to borrow from myriad sources as well as its influence as a cultural device. They challenge the negative premise of the inherent distinctiveness construct, viz protection of the language commons, pointing to positive as well as negative influences of social discourse and context for language development. And they suggest the need to map not assume in advance. Thus they reinforce the conclusion that, at best, inherent distinctiveness should be a measure of inchoate common law trade mark status, its conclusions easily able to be displaced by showing secondary reputation. Further, more unexpectedly, they raise the prospect of trade marks emerging as a kind of language of their own. At their highest trade marks ‘carry a lot of symbolic freight. They do more than identify … things. They reveal social

---

126 Although sociolinguistics emerged as a formal discipline only in the 1960s, its roots lie further back. In the US they can be traced to the pioneering work of inter alia Franz Boas, his pupil Edward Sapir and Sapir’s pupil Benjamin Lee Whorf who premised their radical insights about the intersection of language and culture on their anthropological studies of American Indian tribes: see generally Hudson, above n125 at 1, 95ff. These scholars were working in the US. For early influences across the Atlantic of Swiss linguistic scholar Ferdinand Saussure and Cambridge philosopher Ludwig Wittgenstein, see Harris, above n2.


128 See, in particular, Whorf’s most well-known article ‘Science and Linguistics’ (1940) 42 Technological Rev 229, reprinted Carroll, above n127 at 227.


hierarchies … [and] acts of identity’ — in short, they satisfy the sociolinguistic
idea of what language is and should be. Current trade marks theories identify the
value of trade marks as lying in their ability to identify ‘things’. These theories
might explain the law’s focus on avoiding likely confusion as the most obvious
threat to the identification function. But at best they offer only a plausible reason
for anti-dilution rights in asserting a trade mark may over time be harmed in its
trade mark capacity — leaving them open to the critique of pandering to owner
interests without clear social benefit. Sociolinguistics reveals trade marks that
cross the threshold from commercial indicators to expressive devices to be not
only more language-like, but more truly intellectual property-like, than before
supposed (and not simply in the impoverished sense of the intellectual effort
expended in achieving inherent distinctiveness, which ultimately became
insufficient to sustain any broad notion of a registered trade mark as an exclusive

131 This need not be seen as violating the bare notion of ‘intellectual effort’ which is referred to in
Grain Pool, above n117 as the guiding principle of the intellectual property head of power in
s51(xviii) the Australian Constitution. As the High Court noted in Grain Pool, that standard
does not set a particularly high threshold for intellectual property protection, and in the case of
trade marks the historical understanding of the concept ‘trade mark’ which may have prevailed
when the Constitution was enacted should not be seen as giving ‘the circumference’ of the
power: id at 520–1 (Gleeson CJ, Gaudron, McHugh, Gummow, Hayne & Callinan JJ), citing
Higgins J (dissenting) in Attorney–General (NSW) v Brewery Employees Union of New South
Wales (1908) 6 CLR 469 at 610 (and revealing a rather sociolinguistic perception of the
Constitution’s language as something whose meaning may change over time). Rather,
intellectual effort for Constitutional purposes can now be understood as accommodating even
the development of secondary reputation as a kind of ‘borrowings of words across linguistic
frontiers’, to use Sapir’s words.

132 Suzanne Romaine said this of language generally: see Romaine, above n125 at 244–5.

133 See, for example, William Landes & Richard Posner, ‘Trademark Law: An Economic
Perspective’ (1987) 30 J Law & Econ 265 (the function of trade marks is to reduce consumer
search costs and provide incentives for quality goods and services. Although acknowledging
trade marks might also foster a market in language, the authors see such benefits however as
‘small’: id at 271) – the revised article is reprinted in Landes & Posner, The Economic Structure
Competition: The Recent History’ (1996) 59(2) L & Cont Prob 13; and (more interestingly in
pointing to the increased demand for trademarks corresponding to consumer demand for
2 Marq Intell Prop L Rev 1.

134 See, for example, Landes & Posner, above n133 at 307 – the first (and main) economic ground
for ‘the extension of the property right in a trademark beyond its use to identify a brand … is
that there is a potential for confusion because … the connotations of the name will blur’. See,
(expressing concern that anti-dilution rights may be construed to grant too much to a trade mark
owner, without necessarily significant benefits). In sum, Posner now appears rather ambivalent
about dilution. For a more sophisticated defence of a trade mark anti-dilution right (still resting
on the damage that may be done to the trade mark function but adopting a broader notion of what
that function is) see Schechter, above n23. But the same criticism may be made (suggesting at
least the need for proof of actual or likely damage in actual cases). Interestingly, Schechter also
appears attracted to the notion that ‘coined, arbitrary or fanciful words or phrases’ – ie,
inherently distinctive trade marks – are particularly deserving of protection, because they have
‘added to … the human vocabulary’: see id at 829.
property right). The utilitarian justification for copyright protection, lying in providing incentives for artistic and cultural development for the broader social benefit, can be extended to these new cultural items. For these trade marks, at least, infringement can be rationally extended to encompass uses that will likely blur or tarnish their expressive associations, reducing them to ‘the commonplace’ or rendering them socially ‘unacceptable’.

Sometimes it is suggested trade mark owners do not require or warrant the mantle of anti-dilution protection, evidenced by the fact that there are expressive trade mark uses even without it. But the question is whether there would be greater efforts by traders generally to develop the expressive dimensions of their trade marks in the knowledge they can retain their benefits for longer than the uncertain leadtime until the first copier comes along. If the evidence from copyright is a guide, the answer should be yes. Rational traders who anticipate rewards flowing from taking on the role of cultural ambassador will do so more readily. The role will not be left, as now, to traders whose resources and standing give control of the public consciousness without great need for legal protection against dilution — those able to select trade marks for their “effervescent qualities” then inject them into the stream of communication “with the pressure of a firehose”. In this regard, US anti-dilution law is not necessarily helpful in pointing to the full extent of benefits that could be reaped through a well-tailored law. For instance, the US federal Act’s requirement for a trade mark to be ‘famous’ to have the benefit of protection seems designed to ensure only those

---


136 See especially Rochelle Dreyfuss, ‘Expressive Genericity: Trademarks as Language in the Pepsi Generation’ (1990) 65 Notre Dame Rev 397 at 408–9. (Dreyfuss goes further in arguing the audience is as much responsible to expressive trade mark meaning as those who claim ownership, and they should share the benefits – and compare Coombe, above n19). See also Landes & Posner above n133 at 293 (“[o]ur society seems not to suffer from such lexical poverty that a more extensive system of property rights in words is needed to alleviate it”).

137 Although reliance on potentially self-interested testimony is risky, Eldred v Ashcroft 537 US 186 (2003) on the Constitutionality of the Copyright Term Extension Act (1998) US reveals many of those generating copyright works do so in hope of “fair” financial rewards: see Ginsburg J at f15 (testimony of Quincy Jones, Bob Dylan, Carlos Santana and others in Senate Hearings on the Act). See also Robert Fry, 20th Century English art historian, quoted Megan Richardson, Joshua Gans, Frances Hanks and Philip Williams, The Benefits and Costs of Copyright: An Economic Perspective (Copyright Studies Centre, 2000) at 7 (“almost all artists who have done anything approaching first-rate work have been thoroughly bourgeois people”).

138 Dreyfuss interestingly acknowledges this: Dreyfuss, above n136 at 424 (“[i]t would be interesting to know whether expressive use of trademarks is expanding’, and suspecting it is ‘true’.

139 See Kozinski, above n6.
Further, as the Supreme Court noted in *Moseley v Victoria's Secret Catalogue*, actual dilution must be proved under the Act and it is questionable whether tarnishment is even covered in statutory language focused on ‘lessening’ of the trade mark function. State laws may be different, however. And with European trade mark and unfair competition laws protecting against dilution in more flexible terms, European-origin labels as diverse as ‘Champagne’ (the prestigious and ‘joyful’ name for wine made by the Champagne method in the Champagne region of France), ‘Clareyn’ (literally ‘clear water’, also a Dutch genever label connoting clarity and purity), and ‘Arthur’ (a name chosen and promoted as ‘drôle et poétique’ for clothing in France) have within Europe developed and maintained a unique expressive *cachet*. Australia has few Australian-origin expressive trade marks of great note. Instead we have a history of attempts to appropriate national icons with their own cultural messages intact for registration as trade marks. Our quintessentially Australian trade marks — including ‘Vegemite’, ‘Freddo’, ‘Qantas’ and perhaps ‘Tim-Tam’ — are

140 ‘Fame’ is not defined but factors which a court may consider are given (including duration and extent of use and of advertising and publicity of the market, geographical extent of the mark’s trading area, and degree of recognition of the mark): see Lanham Act § 43(c)(1), 15 USC §1125(c)(1). It has been suggested only ‘household words’, the ‘best known marks in commerce’, could satisfy the standard (*TCIP Holding Co Inc v Haar Communications Inc* 244 F 3d 88 (2001), 99).

141 *Moseley v Victoria’s Secret Catalogue Inc* 537 US 418 (2003). The facts are noteworthy: the defendants adopted ‘Victor’s Secret’ as the name of their lingerie store, subsequently changing this to ‘Victor’s Little Secret’ after complaints from the Victoria’s Secret owner. Dilution was claimed, but without evidence (circumstantial or otherwise) of blurring or tarnishment.

142 Id at 433–4 (Stevens J) (for the Court) – adding scientific proof is ideal, except where dilution can ‘reliably be proven through circumstantial evidence’. In a concurring opinion, Kennedy J said it should be enough to show ‘diminishment of the mark’s capacity to serve its purpose’ will be the ‘probable consequence’ of the defendant’s actions rather than forcing the plaintiff to wait ‘until the damage is done’: id at 435–6.

143 Id at 432 (Stephens J) (though noting the intention was to include tarnishment).

144 See generally Klieger, above n121 at 812–3 (the laws discussed omit requirement of fame, specifically allow for likelihood of dilution harm, and clearly encompass tarnishment).


146 And the subject of one of the most famous European dilution cases decided under the *Benelux Act* when a sound-alike ‘Klarein’ was adopted for a cleaning product, *Colgate Palmolive BV v NV Koninklijke Distilleerderijen Erven Lucas Bols* [1979] EEC 419 – see Charles Gielen, ‘Harmonisation of Trade Mark Law in Europe’ The First Trade Mark Harmonisation Directive of the European Council” [1992] 8 EIPR 262.

primarily employed as identifiers rather than cultural artefacts. Not surprising when those whose efforts are more ambitious receive, at best, patchy legal support from the law. Most of our popular culture comes from elsewhere. For champions of a locally grown Australian popular culture, more incentives might thus seem better than less.

More incentives, thus broader rights against dilution, need not come at the cost to public access. It is now clearer than perhaps in the 19th century that property rights, especially intellectual property rights, need not be absolute and there are good reasons sometimes for curtailing them. One useful insight from historical linguistics is that ownership of trade mark language may lead to private control over social meaning. If so, this is something that needs to be addressed. Noam Chomsky identifies the issue succinctly:

The principles of people like Von Humboldt ... were that people should be free. They shouldn't be under the control of authoritarian institutions ... [T]he idea of developing ... popular control over all institutions, economic, political and so on ... is the right direction to move. This is not a new idea; in fact, its roots are as old as classical liberalism.

For an interesting review, see Owen Morgan, ‘National Icons and the Trade Marks Act 1995’ forthcoming AIPJ (2004) citing Brenda Lomas’s recent attempt to register ‘Waltzing Matilda’ as a trade mark for a vast range of goods and services. In the Lomas case registration was objected to inter alia on the ground of prior ownership by the Winton Shire Council (home of the Waltzing Matilda Centre), without success: Winton Shire Council v Lomas [2002] FCAFC 413. The court did not consider whether a common expression such as ‘Waltzing Matilda’ could be distinctive in the absence of secondary reputation. In the UK attempts to register common cultural symbols as trade marks have been rejected as lacking distinctiveness: see, for instance Re Elvis Presley Trade Marks [1997] RPC 543.

As well as being established common law trade marks, these are all long registered (in the case of Vegemite for more than 70 years: see trade mark 37184; registered in 1928).

As, for example, in CA Henschke & Co v Rosemount Estates Pty Ltd (2000) 52 IPR 42 (evocative ‘Hill of Grace’ wine label not infringed by adoption of ‘Hill of Gold’ for a cheaper wine and nor was there passing off – in part because the connotations of Hill of Gold (more mercenary) were different from those of Hill of Grace) (more spiritual). That the use appeared tarnishing counted against likelihood of confusion.) Fictional and human characters have sometimes been more generously treated in their claims of confusing misuse of their trade marks and images, especially in Australia, leading some to suggest a de facto anti-dilution standard may be employed for them (see for instance S Ricketson, ‘Character Merchandising in Australia: Its Benefits and Burdens’ (1990) 1 AIPJ 191).

Put simply, if the risk is that trade mark owners allowed unlimited control over the use of their language will seek to exploit this as means of restriction and censorship inhibiting possibilities for valuable expression by others, then control should be separated from ownership. As linguistic expert Roger Shuy observes, in trade mark disputes it often emerges that more important than the question ‘who can own language’ is the question ‘who … [can] determine what can be said or written’.155 There are cases already that show trade mark owners unready to allow uses of their trade marks that they believe could reflect badly on them, or are controversial, or lie too far outside the scope of their activities, no matter their overall social value (and including in cases where those who wish to use will pay).156 Will efforts at restraint simply increase with an expansion of rights against dilution? Does the proliferation of internet gripe sites show the benefit of rights not being expansive? Is it enough perhaps that courts already have the power to tailor remedies to facilitate use on payment terms (as a kind of compulsory license)?

The prospect that trade mark owners, given full anti-dilution rights, might seek to prevent the use of ‘The Pink Panther Patrol’ for a public safety campaign mounted by the gay community,158 or ‘Pee Wee Scouts’ as the title of a series of children’s book,159 or ‘Barbie Girl’ as the name of a pop song mildly lampooning the Barbie image,160 or ‘Star Wars’ to critique a government’s defence policies,161 to name but four US cases where dilution has been rather bizarrely claimed, raises the spectra of unfree speech in a society that generally values free speech and may

156 As in Campomar v Nike International Ltd (2000) 202 CLR 45 – when Campomar in 1985 sought permission for Nike to use its trade marked word for sports perfume, the response was ‘we do not believe perfume and cosmetic products are part of this company’s image’: id at 62. When Campomar commenced using the ‘Nike’ label for its own sports perfume without licence, Nike claimed inter alia passing off. (That Campomar adopted a block-printing style assisted the finding of passing off: see Sheppard J at first instance: Nike v United Pharmaceutical (1996) 35 IPR 385 at 408.)
157 In the intellectual property context, a monetary award granted in lieu of an injunction can be viewed as a kind of compulsory licence; sometimes this may be an efficient response to market failure: see Dennis Karjala, ‘Misappropriation as a Third Intellectual Property Paradigm’ (1994) 94 Colum L Rev 2594 at 2608 and generally Guido Calabresi & A Douglas Melamed, ‘Property Rules, Liability Rules and Inalienability: One View of the Cathedral’ (1972) 85 Harvard Law Review 1089.
158 See MGM–Pathé Communications Co v The Pink Panther Patrol 774 F Supp 869 (1991) – the slogan ‘The Pink Panther Patrol’ used for a gay community based safety patrol held to likely confuse the public about possible sponsorship (which the First Amendment would not exempt) irrespective of the defendant’s political motives and the absence of evidence of actual confusion. The dilution claim under New York law was not considered.
159 See Girl Scouts of United States v Bantam Doubleday Dell Publishing Group Inc 808 F Supp 1112 (1992) – the use of the title considered not to disturb the ‘fame, reputation and selling power’ of the scout movement’s trademark expression ‘Scouts’ for the purposes of New York anti-dilution law. Thus First Amendment objectives could be served
160 See Mattel Inc v MCA Records 296 F 3d 894 (2002). The use was held non-commercial by the US Court of Appeals (9th Cir.), and exempted under the Federal Dilution Act: see n178 ff.
161 See Lucasfilm, Ltd v High Frontier 622 F Supp 931 (1985) – claims of ‘trademark infringement, unfair competition, misappropriation, disparagement’ etc dismissed on the basis that the defendants made no commercial use of ‘Star Wars’ (and in particular, the educational services they provided did not amount to the ‘selling of ideas’).
even provide for it in its Constitution.  

Even if plaintiffs generally fail in legal terms, the worry is their ability to threaten will be enough to stop the practice. In part, though, this depends on the perceived public reaction — since one thing experience of the internet has taught (if not already known) is that traders in the public eye are sensitive to their audience. Laws can help shape attitudes.  

"[S]ocial norms and laws are often mutually supporting", as political scientist Robert Axelrod has observed, giving free speech as a particular example. Ideally, a well framed anti-dilution law would tailor its defences to not only sanction but publicly sanction traders who fail to respond positively to genuine efforts to facilitate public discourse (in the same way as fair dealing/use defences to copyright infringement may be openly used for free speech ends). An exception for ‘good faith’ use of the expressive component of a trade mark could encompass uses that valuably promote public debate on matters of political or social concern — including about the image projected by the trade mark itself.  

And a factor in determining ’good faith’ might be whether genuine effort has been made to obtain agreement, or at least make some reasonable accommodation of an

---

162 As the US does in the Constitution’s First Amendment. As to Australia and the UK, see below n166.  


164 Robert Axelrod, The Complexity of Cooperation: Agent-based Models of Competition and Collaboration (1997) at 61 — adding that the law is particularly likely to be an effective instrument of socialisation where it serves to reinforce values that ‘have already attained strength as a social and political norm’.  

165 As, for example, in the recent case of Suntrust Bank v Houghton Mifflin Co 268 F 3d 1257 (2001) (Wind Done Gone parody of Gone With the Wind held a fair use under US copyright law). In ‘Fair Use as Market Failure: A Structural and Economic Analysis of the Betamax Case and its Predecessors’ (1982) 82 Colum L Rev 1600, Wendy Gordon classically points to market failure as an explanation for at least part of the operation of the US fair use doctrine. Later Gordon adds that not all uses – especially those impacting free speech – need to be analysed in market failure terms, unless ‘market failure’ is taken also to accommodate cases where economic norms are inappropriate to govern: Wendy Gordon, ‘Excuse and Justification in the Law of Fair Use’ in Niva Elkin-Koren & Neil Weinstock Netanel, The Commodification of Information (2002) at 149. In my view the broader understanding of ‘market failure’ should prevail.  

166 Uses that target the original material are particularly appropriate for such a defence (see also Michael Spence, ‘Intellectual Property and the Problem of Parody’ (1998) 114 LQR 594 making the point generally that target parodies are worthy of particular protection). But other expressive uses may have social value as well, as for instance ‘Star Wars’ used as for political purposes: above n161. This suggests benefits should be weighed against costs to determine overall benefit. Indeed, political expression should be constitutionally protected under the implied freedom of political discussion found by the High Court in the Australian Constitution: for this implicating private rights see ABC v Lenah Game Meats Pty Ltd (2000) 208 CLR 199 at 224 (Gleeson CJ). In the UK, given implementation of the European Human Rights Convention 1950 in the Human Rights Act 1998, free speech is provided for in even broader terms (Art 10 of the Convention).
owner’s concerns, before proceeding to use, bearing in mind that accommodation may not always be feasible or appropriate.167

In the Australian case of the Campomar Sociedad Ltd v Nike International Ltd,168 where the issue of dilution under the Trade Marks Act 1995 (Cth) was touched on, it was suggested the infringement action might extend to this on present legislative terms — or at least ‘to some extent’.169 The court at the same time noted countervailing interests in free competition, honest concurrent trade mark use, and ‘free and open communication’.170 The court also noted that dilution implicates trade marks functions going ‘beyond the traditional role as an identifier of origin’.171 It will be interesting to see the court developing these ideas further in an actual case. On its face they offer a second best solution. While it is possible that, on a rather extended reading, the existing provisions of the Trade Marks Act 1995 (Cth) could be stretched further to trade mark dilution a court seeking to do so would face several problems in dealing in a new way with standards previously assumed to be focused on likely confusion.172 And more difficult to overcome is the absence of any defence tailored specifically to fostering public debate. There are partial defences; for instance, the Act allows for comparative advertising, good faith descriptive use, even honest concurrent use covering cases where traders have independently developed similar trade marks.173 The flexibility courts have to limit a remedy for infringement should not be underestimated as a device for de facto compulsory licensing. And infringement

167 As, for example, where a parody is meant to forcefully critique and/or lampoon the original (something Gordon was also concerned with, identifying parody as a fair use for which permission should not be required: Gordon, ‘Excuse and Justification‘ above n165 at 159). In other cases, however, such as a use for broader social comment purposes query whether an attempt to gain permission and address concerns might be considered relevant to the ‘good faith’ of any subsequent unauthorised use. And sometimes even mildly critical or lampooning uses may be the subject of good faith efforts at accommodation of interests – as for instance (on the Moseleys’ part) in Victoria’s Secret, above n141.

168 Campomar, above n156.


170 Id at 53–4, in respect of the last quoting Kozinski, above n6.

171 Id at 53.

172 For a fuller discussion see Richardson, above n8, in particular noting that full anti-dilution protection would require a broader reading of the ‘substantially identical’ bass for claiming infringement on the part of Australian courts (although this is something for which there is now a European precedent: see above n147). And there are other limitations in the language of Trade Marks Act 1995 (Cth) s120(3), the so-called anti-dilution provision in the Australian Act, as well. For instance, the provision only covers cases where the use is for goods or services unrelated to those for which the trade mark is registered. The US Act is not so restricted, and neither the ECJ has held, should anti-dilution rights under the European Directive: see Adidas–Salomon, above n122. In fact it is only a use for goods or services that are not directly within the scope of registration but may be considered ‘related’ to registered goods or services that presents a problem under the Australian Act – if dilution could be found with respect to uses for unrelated goods or services under s120(3) it should equally be available with respect to uses for registered goods or services under the terms of the core infringement provision, s120(1) – the difficulty is s120(2) (extending infringement to related goods or services, but with a statutory defence covering cases of no likely confusion).

173 See Trade Marks Act 1995 (Cth) s122 ‘When is a trade mark not infringed?’.
itself requires that a trade mark must be used ‘as a trade mark’,\(^{174}\) meaning entirely expressive uses are public domain.\(^{175}\) There is considerable scope in the last requirement especially for freedom of speech to be preserved.\(^{176}\) But there is no provision in the Australian Act that specifically covers the case where a trade mark is used as a trade mark and expressively at the same time. On its face the result is users would be forced to either avoid anything that may appear to be ‘trade mark use’ (a term capable of multiple applications)\(^{177}\) or else avoid expressive use.

Leaving courts to read in an exemption to unclear statutory language is not ideal either. In the recent *Barbie Girl* case Kozinski J in the US Court of Appeals, relied on a non-commercial use exemption, bolstered by the Constitution’s First Amendment’s free speech mandate, to hold the pop group Aqua free to use ‘Barbie Girl’ as the name of its song. The use being expressive *could not be* ‘commercial’, was the conclusion.\(^{178}\) Thus a free speech exemption was engrafted onto the non-commercial use exemption specified in the federal anti-dilution Act. Such an approach, if taken literally, would leave little scope for a trade mark dilution claim (certainly one premised on adoption of the expressive elements of the trade mark, as proposed in this article). And the results are perverse if all expressive uses no matter how meagre their contribution to social discourse, how destructive they may be to a trade mark’s integrity or imagery, and how commercial their flavour now find exemption. In the *Barbie Girl* case financial rewards were reaped from the choice of ‘Barbie Girl’ as a feature of the song’s name and theme.\(^{179}\) The outcome in the particular case may still be approved: children could benefit from questioning Barbie’s imagery, promoted by Mattel, of a ‘plastic’ teenage girl-cum-commodity and the extent of real damage done to the Barbie image is debatable. (Query whether a more prescient Mattel might have taken some advantage even of

---

\(^{174}\) The requirement for use ‘as a trade mark’ is expressed in *Trade Marks Act* 1995 (Cth) s120 – unlike s10 of the UK 1994 Act: the omission there has led to debate as to whether use as a trade mark is required under UK law. Contrast *Arsenal* cases, above n123 (the proper question is the effect of the use on the registered owner’s trade mark) and *R v Johnstone* [2003] 3 ALL ER 884 (trade mark use required). In the US trade mark use is not specified in the Federal Anti-Dilution Act, but some think it is inherent in the nature of a diluting (or at least blurring) use that this be as a trade mark: see McCarthy, above n119 at Vol 4, Part II § 24.103.

\(^{175}\) As, for example, in *Top Heavy Pty Ltd v Killin* (1996) 34 IPR 282 (‘Chill out’ used expressively did not infringe the plaintiff’s registered trade mark).

\(^{176}\) The point is well made by Patricia Loughlan, ‘Protecting Culturally Significant Uses of Trade Marks (Without a First Amendment)’ [2000] *EIPR* 328.

\(^{177}\) Loughlan, for example, argues that an Australian court would likely not find Aqua’s use of ‘Barbie Girl’ in its song to be a trade mark use: see id at 331. Some judges might: see, for example, *Coca-Cola Co v All-Fect Distributors Ltd* (1999) 47 IPR 481 (prominent use of Coca-Cola bottle-like shape for sweets held a trade mark use, notwithstanding defendant’s attempt to portray itself as merely acting ‘in fun’).

\(^{178}\) Mattel, above n160 (Kozinski J) ‘[w]ere we to ignore the expressive value that some marks assume, trademark rights would grow to encroach upon the zone protected by the First Amendment … [T]he trademark owner does not have the right to control public discourse’. The Court concluded that Aqua’s song ‘merely lampoons the Barbie image’; therefore was not ‘commercial’: cert denied *Mattel Inc v MCA Records, Inc* 123 S Ct 993 (2003).

Aqua’s critique.)\textsuperscript{180} But a ‘Barbie Enjoys Cocaine’ logo on mass-produced T-shirts,\textsuperscript{181} or a ‘Barbiesplaypen.com’ website employed for adult entertainment trading purposes might be less appealing for public discourse and more damaging to Barbie’s image as a charmingly innocent child-friendly icon.\textsuperscript{182} A legislative solution fine-tuned to the private/public language divide, a tailored free speech defence to infringement that allows for consideration of social benefits and costs of an expressive use as well as the good faith (or not) of the parties concerned, is in my view preferable to simply relying on the blunt instrument of ‘commercial’ or even ‘trade mark’ use.

5. **Concluding Comments**

There is a tendency in modern academic scholarship to argue for even closer regulation of the nexus between trade mark rights and language, extending the high regulation approach for registration to infringement.\textsuperscript{183} I argue that drawing on the expressive power of the law, its capacity as a kind of prestige dialect to foster desirable social attitudes and behaviour, could work better in promoting a healthy relationship between trade marks and language. Traders who seek to negotiate the line between expressive meaning and trade mark use should not be penalised for their contribution to popular culture. Their trade marks should be recognised for registration if they function as common law trade marks or are likely to in the future. Further, protection against dilution by blurring or tarnishing the associations of a registered trade mark should ideally be provided for by legislation, or as a second best solution read into the provisions of the *Trade Marks Act 1995* (Cth). Those recognised as owners should have at least some control over how the expressive components of their trade marks will be used. However, the risk that ownership of language may yet be used as a tool of censorship means a well-tailored free speech infringement exception is fundamental to the modern registered trade mark system.

\begin{footnotesize}
\begin{itemize}
\item\textsuperscript{180} In fact, it might be argued, a ‘prescient’ Mattel might benefit from considering the message of ‘Barbie Girl’ that Barbie needs to be liberated from her pre-feminist roots as a mere commodity, a plaything for others (including her so called ‘playmate’ Ken), without a personality of her own: compare above n39ff (influence of gripe sites on the internet).
\item\textsuperscript{181} See *Coca-Cola Company v Gemini Rising, Inc* 346 F Supp 1183 (1972) (use of ‘Enjoy Cocaine’ in classic Coca-Cola script on T-Shirts enjoined under New York trademark and unfair competition law).
\item\textsuperscript{182} See *Mattel Inc v Internet Dimensions, Inc* 55 USPQ 2d (BNA) 1620 (2000) (defendant’s adoption of ‘Barbiesplaypen.com’ as its domain name found to tarnish and dilute Barbie trade marks and ‘wholesome’ image, in violation of the *Anticybersquatting Consumer Protection Act 1999*).
\item\textsuperscript{183} See, in particular, Brown, above n135 and other references passim.
\end{itemize}
\end{footnotesize}
What Explains Variations in Dissent Rates?: Time Series Evidence from the High Court

RUSSELL SMYTH*

Abstract
This article employs a multiple regression statistical model to examine the determinants of variations in the dissent rate on the High Court using annual time series data between 1904 and 2001. We hypothesize that variations in dissent rates over time depend on both institutional factors internal to the Court and socio-political complexity external to the Court. We find that the main determinants of variations in dissent have been prevailing norms on the Court as to acceptable levels of dissent, caseload, and the Australia Acts which removed the final vestiges of appeal to the Privy Council. We also find evidence of significant variation in dissent under the leadership of different chief justices.

* Professor, Department of Economics, Monash University. I thank Matthew Groves and two anonymous referees for helpful suggestions on an earlier version of this article and Brooke Watson and Olga Ryan for assistance with data collection.
1. Introduction

A large literature has emerged for courts in the United States which examines the reasons why appellate judges dissent. Some studies have employed what has been broadly termed the ‘attitudinal model’ to attempt to explain variations in dissent rates on the basis of the personal characteristics of the judges. The personal characteristics of judges that have been tested to explain differences in dissent rates between judges in the United States include political affiliation, social background (school, religion, class) and demographic and hereditary factors (age and birthplace). Australian legal academics have been reluctant to embrace attitudinal models of decision-making, arguing that they are not apposite to the Australian context. In the United States Supreme Court, which has traditionally been recognised as having a much more overt political role than the High Court, there is greater scope for personal characteristics to come to the fore. Existing attempts to explain dissent in the High Court using the attitudinal model support this conjecture. These studies suggest that the personal characteristics of judges including background have little power in explaining variations in dissent rates between judges.

Another strand of the literature in the United States has focused on institutional and external factors, rather than the judges’ personal characteristics, as the main determinants of variation in dissent. This literature has attempted to explain variations in dissent rates in terms of factors such as changes in caseload, whether the court has a discretion to select the cases it hears, leadership of the chief justice and the proportion of people living in urban areas, which is used to depict growing social complexity. The objective of this article is to draw on the hypotheses that have been tested in the literature for United States courts to examine the influence of changes in the High Court’s institutional arrangements, leadership and external environment on the Court’s annual dissent rate between its first full year of operation in 1904 and 2001. The article is premised on the belief that this exercise is likely to be more fruitful than looking at the personal characteristics of the judges, which has been the focus of past research.

---


3 For the most recent and explicit test of the attitudinal model see Russell Smyth, ‘Explaining Historical Dissent Rates in the High Court of Australia’ (2003) 41 Commonwealth and Comparative Politics 83.
A study such as this is useful because it contributes to our understanding of how the High Court functions and how institutions affect decision-making. To this point, while there has been much casual observation about variations in dissent rates on the High Court over time and the reasons for these fluctuations, there have been few attempts to collect the hard data. The absence of a consistent data set on the High Court’s dissent rate over time has meant that doing studies such as this, which seek to explain variations in dissent using statistical methods, has not been possible. This has been a limitation in our understanding of judicial behaviour because even for those who are otherwise reticent about the use of empirical studies in legal scholarship, quantitative analysis can be used as evidence to support, or disprove, speculation about trends in dissent and its causes.

The balance of the article is set out as follows. The next section proposes several hypotheses about judicial dissent under the broad rubric of external and internal institutional factors, building on the existing literature for the United States. Section 3 contains an overview of the data, including a discussion of how it was collected and a review of the statistical methods employed in the study. The results are presented in Section 4, while the final section concludes with some suggestions for future research.

---


6 Andrew Lynch has made a start in this direction with his methodological piece on how dissent should be measured and an attempt to measure dissent in the Gleeson Court. See Andrew Lynch, ‘Dissent: Towards a Methodology for Measuring Judicial Disagreement in the High Court of Australia’ (2002) 24 Syd LR 470; Andrew Lynch, ‘The Gleeson Court on Constitutional Law: An Empirical Analysis’ (2003) 26 UNSWLJ 32. An example of an earlier attempt at collecting data on dissent rates is Anthony Blackshield, ‘Quantitative Analysis: The High Court of Australia, 1964–1969’ (1972) 3 Lawasia 1. However, there are no studies which have collected data on dissent rates through the history of the High Court.
2. **Hypotheses**

Following the extant literature for the United States we hypothesise that organisational factors, leadership on the Court, and social and political complexity have influenced variations in the High Court’s dissent rate over time. The organisational factors that we consider are the Court’s caseload, whether there is a norm of dissent operating on the Court, the enactment of the *Australia Acts*, which made the High Court the final court of appeal for Australia, and the existence of case selection discretion, defined as whether the Court has a discretion to select the cases that it hears. In order to consider the effects of leadership on the Court we examine the tenure of each chief justice. To capture increasing social and political complexity over time we use the urbanisation rate, which is defined as the proportion of the population which is living in urban areas.

**A. Caseload and Institutions of the Court**

Numerous studies in the United States suggest that courts with discretionary power to control their own caseload are likely to experience higher rates of dissent. One rationale is that with case selection discretion, run-of-the-mill cases are weeded out, leaving only complex cases where there is a higher likelihood that reasonable minds can reach different conclusions. Another explanation is that heavy caseload demands reduce the expression of dissent. The underpinning argument is that in courts with high caseloads not only is there likely to be a higher proportion of relatively easy cases, but there is less time for the reflection and hair splitting which is likely to contribute to dissent.

The requirement of grant of special leave to appeal was introduced in the High Court in 1984 by Section 3(1) of the *Judiciary Amendment Act (No.2)*. This reduced the number of cases the Court heard and increased their complexity. Michael Kirby describes the effect of introducing special leave to appeal: ‘Virtually all of the cases which are chosen involve delicately balanced issues where there are powerful arguments for both sides. Quite frequently they are expressed in the majority and minority opinions of the court under appeal’. This is likely to contribute to increased disagreement over outcome.

---

7 *Australia Act* 1986 (Cth); *Australia (Request and Consent) Act* 1986 (Cth); the *Australia Acts Request Act* 1985 of each state; *Australia Act* 1986 (UK).
This suggests the following hypotheses:

\textbf{H1}: \textit{The dissent rate will be inversely related to the caseload of the Court.}

\textbf{H2}: \textit{The introduction of case selection discretion will result in higher rates of dissent.}

Soon after the introduction of special leave to appeal to the High Court, the High Court became the final court of appeal for Australia. Whilst there had been abolition of appeals to the Privy Council from the High Court through enactments in 1968\textsuperscript{12} and 1975,\textsuperscript{13} until the enactment of the \textit{Australia Acts} in March 1986 it was possible to appeal to the Privy Council directly from State courts exercising State jurisdiction.\textsuperscript{14} The effect of abolishing appeals to the Privy Council on the dissent rate, however, is unclear.

On the one hand, it is well-documented that abolition of appeals to the Privy Council acted as a catalyst for the Mason High Court to be more adventurous in its approach to decision-making. Sir Anthony Mason, himself, has suggested that many of the landmark decisions delivered during his tenure as chief justice would not have occurred if the High Court had still been bound by final review by the Privy Council.\textsuperscript{15} When interpretation of the existing law is moving rapidly as it was in the Mason Court and is being driven by what some commentators regarded as overt judicial activism, there are likely to be strong differences in interpretation. In these circumstances, one might expect there to be higher levels of dissent as some judges hold on to traditional doctrines in the face of change.

On the other hand, when a court becomes a final court of appeal it takes on added responsibility for laying down the law of the jurisdiction. Some argue that with this added responsibility there is a need to ensure consistency in interpretation of the law. For instance, Lord Reid often expressed the view that the writing of dissenting judgments in final courts of appeal should be reserved for very important principles of law, although other judges, such as Michael Kirby, do not agree with this view.\textsuperscript{16}

A related argument, sometimes used to explain the low rate of dissent on the United States Supreme Court before 1940, is that dissent weakens the authority of a final court of appeal. For example, famously Louis Brandeis is known to have withheld many dissenting opinions on the United States Supreme Court because he felt that random dissents would weaken the Court.\textsuperscript{17} William Howard Taft, who was chief justice of the United States Supreme Court from 1921 to 1930, had a similar view. Like Lord Reid, he believed that ‘important questions of law should not be broken anymore than we can help by dissents’.\textsuperscript{18} Taft is reported as stating:

\begin{itemize}
  \item \textsuperscript{12} Privy Council (Limitations of Appeals) Act 1968 (Cth).
  \item \textsuperscript{13} Privy Council (Appeals from the High Court) Act 1975 (Cth).
  \item \textsuperscript{14} See David Jackson, ‘Leave to Appeal’ in Tony Blackshield, Michael Coper & George Williams (eds), \textit{Oxford Companion to the High Court of Australia} (2001).
  \item \textsuperscript{15} Anthony Mason, ‘Reflections on the High Court of Australia’ (1995) 20 \textit{MULR} 273.
  \item \textsuperscript{16} See Michael Kirby, ‘On the Writing of Judgments’ (1990) 64 \textit{ALJ} 691 at 707.
  \item \textsuperscript{17} James P Frank, ‘Review of the Unpublished Opinions of Mr Justice Brandeis’ (1958) 10 \textit{Journal of Legal Education} 401.
  \item \textsuperscript{18} Walter F Murphy, \textit{Elements of Judicial Strategy} (1964) at 47.
\end{itemize}
‘I don’t approve of dissents generally, for I think that in many cases where I differ from the majority, it is more important to stand by the Court and give its judgment weight than merely to record my individual dissent’.19 In contrast, Harlan Stone, who was chief justice of the United States Supreme Court from 1941 to 1946, when there was a sharp increase in the dissent rate, would not control dissent. Stone considered dissent critical because ‘legal principles … never sprang full-fledged from the brains of any man or group of men. They are the ultimate resultant of the abrasive force of the clash of competing and sometimes conflicting ideas’.20

H3: The effect of the Australia Acts on the rate of dissent in the High Court is unclear. On the one hand there might be a proclivity for judges to dissent more in the face of rapid changes in legal interpretation. On the other hand, dissent might be viewed as undermining the authority of a final court of appeal, leading to less dissent.

Judicial decision-making is influenced by the prevailing norms of decision-making on the Court. The sharp increase in the dissent rate on the United States Supreme Court from the 1940s has been identified with a decline in consensual norms.21 Previous research for the High Court suggests that given the long tradition of seriatim judgment writing and the absence of formal conferencing, unlike the United States Supreme Court, the High Court never developed a consensual norm.22 Because the level of dissent will be influenced by norms of behaviour, established custom and precedent in a court’s decision-making practices will be important in explaining judicial behaviour. This suggests that, at any given point in time, accepted levels of consensus and dissent on the Court in the recent past will be an important precedent for current justices. Consistent with this conjecture, in a previous study Brian Goff found that on the United States Supreme Court the dissent rate in any given year had a strong positive effect on the dissent rate the following year.23

This suggests a fourth hypothesis:

H4: Existing approaches and attitudes to a court’s appropriate decision-making norms will be important in explaining current judicial behaviour. The dissent rate in the Court’s recent past will be a good predictor of the Court’s current dissent rate.

19 Id at 61.
23 Brian Goff, ‘Supreme Court Consensus and Dissent: Estimating the Role of the Selection Screen’, Department of Economics, Western Kentucky University, Mimeo, December 2002.
B. Role of the Chief Justice

The leadership of the court has often been identified in studies of United States courts as one of the major factors influencing consensus and dissent. The importance of leadership is reflected in Figure 1, which plots fluctuations in the annual dissent rate on the High Court since its establishment. It shows that while the dissent rate has generally fluctuated between five and 15 judgments per 100 judgments delivered there have been substantial movements under different chief justices. For example, the second half of the 1920s and early 1930s, mid-to-late 1940s and the late 1990s were all periods of relatively high dissent. In contrast, in the early Griffith Court there was little dissent.

With the exception of the early Griffith Court (1903-06) and Gleeson Court, the High Court has not had a formal conference procedure. Latham, Dixon and Barwick all experimented with informal conferences. Latham and Barwick are considered to have been largely unsuccessful in generating consensus through this means, although Dixon had more success. The tenure of Knox, Gavan Duffy and Latham were all marked by strong personality differences on the Court which impeded attempts to build consensus and contributed to relatively high dissent rates in the 1920s–1940s. When Evatt resigned in 1940 personal relations between the judges seemed to improve, but the Court did not function as a cohesive social unit until after Rich and Starke were replaced by Fullagar and Kitto following the defeat of the Chifley Labor government.

Dixon is often regarded as the High Court’s most effective leader. His abilities to forge consensus were founded on the enormous respect he possessed as a jurist. Compared with Dixon, Barwick is regarded as an ineffective chief justice. Barwick attempted to impose greater coordination in judgment writing and this was resisted by the puisne justices as an infringement on their individualism.

---


25 For a more extended discussion of differences in leadership on the High Court and their relevance for building consensual norms see Smyth, above n3 at 88–91.


Figure 1

DISSENT RATE PER ONE HUNDRED JUDGMENTS
The workings of the Court improved under Mason, Brennan and Gleeson. Meetings were held each month in the Mason Court to monitor progress with judgment writing and to exchange views and this practice was continued on the Brennan Court.\textsuperscript{31} The anecdotal evidence which exists suggests that Mason’s approach to managing the Court resulted in a collegial atmosphere.\textsuperscript{32}

\textit{H5: The ability of the chief justice to exercise leadership and foster consensus will influence the dissent rate. The dissent rate will be higher under chief justices who have failed to effectively exercise social leadership or foster consensual norms.}

\textbf{C. Social and Political Complexity}

More complex economic, political and social environments are believed to produce more frequent expressions of disagreement in judicial institutions.\textsuperscript{33} In the studies for United States courts the urbanisation rate is used as a measure of socio-political complexity.\textsuperscript{34} Dean Jaros and Bradley Canon capture the reasoning behind using this variable:

'[The urbanisation rate] is an appropriate indicator of social heterogeneity for as populations become more concentrated in cities, most forms of human activity become more complex. Concentration and industrialization are associated with a more diverse economy and thus with greater specialization. This produces a basis for a large number of relatively specific interests. The resultant configuration of demands upon governmental agencies become more varied. Similarly, in the American historical experience many ethnic and religious minorities have settled primarily in urban areas, thus making such environments more socially diverse.'\textsuperscript{35}

This is true not only of the American historical experience, but also of immigration settlement in Australia. It can be hypothesised that urbanisation produces the conditions for increased demands and conflicts in the political system which will generate more legal conflict.\textsuperscript{36} Increasing rates of urbanisation are also often associated with changing societal values, which have been tied to judicial dissent. In contrast to the common depiction of a dissenter as a judicial innovator, John Schmidhauser has argued that the typical dissenter on the United States Supreme Court is not an innovator, but ‘a tenacious advocate of traditional legal doctrines which were being abandoned during his tenure’.\textsuperscript{37}

\textit{H6: The degree of urbanisation is positively related to the level of dissent}

\begin{itemize}
  \item\textsuperscript{31} Simpson, above n26.
  \item\textsuperscript{32} See, for example, Anthony Mason, ‘Personal Relations: A Personal Reflection’ in Blackshield, Coper & Williams (eds), \textit{Oxford Companion to the High Court of Australia} (2001); Gerard Brennan, ‘A Tribute to Sir Anthony Mason’ in Cheryl Saunders (ed), \textit{Courts of Final Jurisdiction: The Mason Court in Australia} (1996).
  \item\textsuperscript{33} Brace & Hall, above n4 at 57.
  \item\textsuperscript{34} Jaros & Canon, above n1; Henry Glick & Kenneth Vines, \textit{State Court Systems} (1973); Kenneth Vines & Henry Glick, ‘State Courts and Public Policy’ in Herbert Jacob & Kenneth Vines (eds), \textit{Politics in the American States} (3\textsuperscript{rd} edn, 1976).
  \item\textsuperscript{35} Jaros & Canon, above n1 at 332.
  \item\textsuperscript{36} Peterson, above n4 at 420.
  \item\textsuperscript{37} Schmidhauser, above n1 at 209.
\end{itemize}
3. **Empirical Specification and Methodology**

To test the above hypotheses we apply multiple regression analysis to analyze time series data for 1904 to 2001. In multiple regression analysis, one variable, which is called the dependent variable, is expressed as a function of other variables, called the explanatory variables. In this study the dependent variable is the annual dissent rate on the High Court. The explanatory variables are variables designed to measure the effect of consensual norms on the Court, the Court’s caseload, the effect of the *Australia Acts* making the High Court the final court of appeal for Australia, whether the Court had discretion to select the cases it hears and the effect of leadership. The objective is to ascertain to what extent variation in each of these explanatory variables can explain variation in the dissent rate. With multiple regression analysis we can determine which of these explanatory variables has been most important in explaining variation in dissent rates and whether each of these variables has had a positive or negative effect on the dissent rate on the High Court holding the effect of each of the other variables constant.

The definition of each variable and expected sign on the explanatory variables are given in Table 1. The expected sign is either positive or negative or unclear. If the expected sign is positive this means that variation in the variable is expected to lead to an increase in the dissent rate. If the expected sign is negative, it is expected that variation in the variable will result in a lower dissent rate. If the expected sign is unclear the stated hypotheses suggest there could be either a positive or negative relationship.

We now discuss each of the variables in Table 1 in more detail, beginning with the dependent variable. The dependent variable, *Dissent*, is the annual proportion of dissenting judgments per 100 judgments in multi-judge panels in cases reported in the Commonwealth Law Reports decided between 1904 and 2001. The dissent rate was calculated through reading and recording the outcome of each case. A judgment was classified as dissenting if the Justice disagreed with the result proposed by the majority expressed in the orders of the Court. In adopting this approach we use the term ‘dissenting’ in a manner consistent with the recent literature on measuring (dis)agreement on the High Court, particularly the methodology suggested by Andrew Lynch.38 In most cases determining whether a judgment is in dissent is relatively straightforward. However, as Lynch points out, in a few cases there are shifting majority opinions within one case and the researcher has to make choices to determine whether a judgment is in dissent.39

---


Table 1: Definition of Variables and Expected Sign

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
<th>Expected Sign</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissent</td>
<td>The proportion of dissenting judgments per 100 judgments in multi-judge panels delivered in cases, which are reported in the <em>Commonwealth Law Reports</em></td>
<td>Dependent variable</td>
</tr>
<tr>
<td>Dissent(-1)</td>
<td>For any given year, the dissent rate, in the previous year.</td>
<td>Positive</td>
</tr>
<tr>
<td>Urban</td>
<td>The percentage of the population living in urban areas</td>
<td>Positive</td>
</tr>
<tr>
<td>Caseload</td>
<td>Total Caseload</td>
<td>Negative</td>
</tr>
<tr>
<td>Total Caseload</td>
<td>Total number of appeals heard in the High Court as a proportion of the number of justices on the Court</td>
<td>Negative</td>
</tr>
<tr>
<td>Case Per Judge</td>
<td>Number of appeals heard in the High Court as a proportion of the number of justices on the Court</td>
<td>Negative</td>
</tr>
<tr>
<td>Discretion</td>
<td>Dummy variable set equal to 1 when the High Court had discretion over which cases it heard; zero otherwise.</td>
<td>Positive</td>
</tr>
<tr>
<td>Final</td>
<td>Dummy variable set equal to 1 following the enactment of the <em>Australia Acts</em> eliminating all appeals to the Judicial Committee of the Privy Council; zero otherwise.</td>
<td>?</td>
</tr>
<tr>
<td>Griffith</td>
<td>Dummy variable set equal to 1 for the period when Griffith was Chief Justice of the Court; zero otherwise.</td>
<td>?</td>
</tr>
<tr>
<td>Knox</td>
<td>Dummy variable set equal to 1 for the period when Knox was Chief Justice of the Court; zero otherwise.</td>
<td>?</td>
</tr>
<tr>
<td>Isaacs/Duffy</td>
<td>Dummy variable set equal to 1 for the period when Isaacs and Duffy were Chief Justices; zero otherwise.</td>
<td>?</td>
</tr>
<tr>
<td>Latham</td>
<td>Dummy variable set equal to 1 for the period when Latham was Chief Justice of the Court; zero otherwise.</td>
<td>?</td>
</tr>
<tr>
<td>Dixon</td>
<td>Dummy variable set equal to 1 for the period when Dixon was Chief Justice of the Court; zero otherwise.</td>
<td>?</td>
</tr>
<tr>
<td>Barwick</td>
<td>Dummy variable set equal to 1 for the period when Barwick was Chief Justice of the Court; zero otherwise.</td>
<td>?</td>
</tr>
<tr>
<td>Gibbs</td>
<td>Dummy variable set equal to 1 for the period when Gibbs was Chief Justice of the Court; zero otherwise.</td>
<td>Benchmark category</td>
</tr>
<tr>
<td>Mason</td>
<td>Dummy variable set equal to 1 for the period when Mason was Chief Justice of the Court; zero otherwise.</td>
<td>?</td>
</tr>
<tr>
<td>Brennan</td>
<td>Dummy variable set equal to 1 for the period when Brennan was Chief Justice of the Court; zero otherwise.</td>
<td>?</td>
</tr>
<tr>
<td>Gleeson</td>
<td>Dummy variable set equal to 1 for the period when Gleeson was Chief Justice of the Court; zero otherwise.</td>
<td>?</td>
</tr>
</tbody>
</table>
Where there are multiple issues in the case, one approach would be to record a
dissent if Justice X dissented on any issue, but this tends to exaggerate the level of
dissent if Justice X agreed in the orders for the other issues in the case. Therefore,
in such cases we decided on which was the most important issue or issues before
the Court and recorded whether Justice X dissented on this issue or issues.

Turning to the first of the explanatory variables shown in Table 1, for any given
year, \( \text{Dissent}(\text{-}1) \) represents the dissent rate in the previous year. For example, for
1905, \( \text{Dissent}(\text{-}1) \) denotes the dissent rate in 1904; for 1906, \( \text{Dissent}(\text{-}1) \) denotes
the dissent rate in 1905 and so on. \( \text{Dissent}(\text{-}1) \) is used to test Hypothesis 4 that the
dissent rate in the Court’s recent past will be a good predictor of the Court’s current
dissent rate. If Hypothesis 4 holds, the sign will be positive meaning that an
increase in the dissent rate in any given year is expected to have a positive effect
on the dissent rate the following year.

The next explanatory variable given in Table 1 is \( \text{Urban} \), which is defined as the
percentage of the population that is living in urban areas in Australia. The
expected sign on the \( \text{Urban} \) variable is positive, meaning that an increase in the
urbanisation rate is expected to result in a higher rate of dissent on the High Court
(Hypothesis 6). The urbanisation rate from 1904 to 1959 is sourced from Mukherjee et al\(^{40}\) and the urbanisation rate from 1960 to 2001 is taken from the
World Bank World Tables.\(^{41}\)

To measure the effect of caseload on the dissent rate we use two alternative
variables. \( \text{Total Caseload} \) is the total number of appeals heard in the High Court
for each year between 1904 and 2001. \( \text{Case Per Judge} \) is the total number of
appeals heard as a proportion of the number of judges on the Court for each year
between 1904 and 2001. \( \text{Case Per Judge} \) is designed to correct for the fact that the
number of justices has not been constant throughout the Court’s history and that an
increase in caseload might reflect more sitting justices. The expected sign on both
variables is negative, which implies that an increase in the Court’s caseload is
expected to have a negative effect on the dissent rate (Hypothesis 1). The sources
for the caseload figures are the Commonwealth of Australia Yearbooks and the
High Court of Australiia Annual Reports.

The remaining explanatory variables are depicted as what are known as
dummy variables. Data may either be quantitative or qualitative. Quantitative data
is that which takes a number for each year. For example, in any given year we can
assign a number to the Court’s caseload or the urbanization rate. Qualitative data
refers to a characteristic which either exists or does not exist at a given point in
time. Examples are: Did the Court have case selection discretion? Was the Court
the final court of appeal for Australia? Was Griffith chief justice? These are

---

\(^{40}\) Satyanshu Mukherjee, Anita Scandia, Dianne Dagger & Wendy Matthews, \textit{Sourcebook of

\(^{41}\) Available through the DX database.
questions to which the answer is either ‘yes’ or ‘no’ for any given year. Dummy variables are used to operationalise these characteristics. In other words, dummy variables are used to quantify the effect of specific institutional characteristics of the Court at particular periods of time on variations in the dissent rate.

Each dummy variable takes the value of one or zero for each year from 1904 to 2001 corresponding to whether a specific institutional characteristic of the Court was in place. For each year, if the answer is ‘yes’ the variable is assigned the value one. If not, the variable is assigned the value zero. The first dummy variable described in Table 1 is *Discretion* which depicts whether the Court has a discretion to select its own cases. It takes the value zero for each year from 1904 to 1983 when the Court did not have a discretion to select its own cases and one for each year from 1984 to 2001 when the Court did have discretion to select its own cases. The expected sign on *Discretion* is positive, meaning that the introduction of case selection discretion is expected to have a positive effect on the dissent rate. To put it differently, holding the effect of all the other explanatory factors constant, the dissent rate is expected to be higher after 1984 when the High Court was given the discretion to select the cases that it can hear (Hypothesis 2).

The next dummy variable in Table 1 is *Final*, which is designed to measure the effects of the *Australia Acts* on the dissent rate. It takes the value zero for each year from 1904 to 1985, prior to the enactment of the *Australia Acts* and the value one for each of the years 1986 to 2001. The expected sign on *Final* is unclear, meaning that the *Australia Acts* could have either a positive or negative effect on the dissent rate (Hypothesis 3).

The other dummy variables denote the tenure of the chief justices and are designed to capture the effect of leadership on the Court. For instance, the dummy variable for the Griffith Court takes the value one for each year from 1903 to 1919 and the value zero for the years 1920 to 2001. The dummy variable for the Knox Court takes the value one for each year from 1920 to 1929 and the value zero for every other year. The dummy variables for the other chief justices were constructed in a similar fashion. When examining the effects of leadership on the Court we are interested in whether the dissent rate under each chief justice is higher or lower than under the other chief justices. Multiple regression analysis allows us to ascertain whether the dissent rate under each chief justice is higher or lower relative to the term of a specific chief justice, which is treated as a benchmark. In the analysis below we treat Gibbs’ period as chief justice as the benchmark. Thus, the findings for the signs for the dummy variables for the other chief justices have to be interpreted relative to Gibbs’ period as chief justice.

To illustrate, if the sign on the dummy variable for Barwick’s term as chief justice is positive this means that the dissent rate was higher under Barwick than Gibbs, holding the other explanatory variables constant. To take another example, if the sign on the dummy variable for Dixon is negative, this implies that the

---

42 When a new chief justice assumed office mid-year, the chief justice who presided for the majority of the year was coded 1. Because Isaacs was chief justice for a period shorter than one year (April 1930 to January 1931) his term was coupled with that of Gavan Duffy.
dissent rate when Dixon was chief justice was lower relative to when Gibbs was chief justice holding the other explanatory variables constant. If this is the case, in terms of Hypothesis 6 we could conclude that Dixon was more effective than Gibbs at forging consensus. In Table 1 the expected sign for each Chief Justice’s term relative to Gibbs is given as unclear because while anecdotal evidence coupled with visual inspection of Figure 1 might suggest that the expected sign should be positive or negative for specific chief justices there is no theoretical reason to expect positive or negative signs on each of the variables.

Table 2 presents the mean (or average), standard deviation and minimum and maximum values for each variable. Some interesting points are worth noting. The average dissent rate over the period of the study was 9.8 per 100 judgments with a minimum of 1.05 per 100 judgments, which was in 1905, and a maximum of 23.46 per 100 judgments, which was in 1944. The average total caseload of the Court has been 86.07 cases per annum with a low of 29 cases and a high of 134 cases per year. Given that the dummy variables can either take the value zero or one, the minimum value for each dummy variable is zero and the maximum is one. The mean value for most of the dummy variables is close to zero which reflects the fact that there are more ‘zeros’ than ‘ones’ for each of the variables. To illustrate, consider the dummy variable Final denoting the Australia Acts. It takes the value one for 16 years (1986–2001) and zero for 82 years (1904–1985).

Table 3 shows the correlation coefficients between the dissent rate and each of the explanatory variables. As a first approximation, this is a measure of the degree of association between the two variables. The correlation coefficient can either have a positive or negative sign and can take values between -1 and +1. Values close to -1 suggest a strong negative relationship. Values close to +1 suggest a strong positive relationship. Most of the correlation coefficients in Table 3 are in the range -0.3 to +0.3 suggesting a weak association between each of the variables and the dissent rate.

43 There is no statistical rule for choosing the benchmark and equally one of the other chief justices could have been chosen. If one of the other chief justices had been used as the benchmark the results would be interpreted relative to the chosen chief justice. The Gibbs Court was chosen as the benchmark in this study because visual inspection of Figure 1 suggests that dissent under Gibbs was ‘average’, meaning that it was neither very high or very low. This should make it easier to document variations from the average under other specific chief justices such as Griffiths & Dixon (below average dissent) or Latham & Barwick (above average dissent) if they are measured relative to Gibbs’ term.
### Table 2: Descriptive Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>Minimum</th>
<th>Maximum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissent</td>
<td>9.80</td>
<td>3.73</td>
<td>1.05</td>
<td>23.46</td>
</tr>
<tr>
<td>Urban</td>
<td>74.21</td>
<td>11.23</td>
<td>57.20</td>
<td>91.10</td>
</tr>
<tr>
<td>Caseload</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Caseload</td>
<td>86.07</td>
<td>23.74</td>
<td>29.00</td>
<td>134.00</td>
</tr>
<tr>
<td>Case Per Judge</td>
<td>13.26</td>
<td>4.08</td>
<td>4.10</td>
<td>29.70</td>
</tr>
<tr>
<td>Discretion</td>
<td>0.18</td>
<td>0.39</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Final</td>
<td>0.16</td>
<td>0.37</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Griffith</td>
<td>0.16</td>
<td>0.37</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Knox</td>
<td>0.10</td>
<td>0.30</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Isaacs/Duffy</td>
<td>0.06</td>
<td>0.24</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Latham</td>
<td>0.16</td>
<td>0.37</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Dixon</td>
<td>0.12</td>
<td>0.33</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Barwick</td>
<td>0.17</td>
<td>0.38</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Gibbs</td>
<td>0.08</td>
<td>0.24</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Mason</td>
<td>0.08</td>
<td>0.28</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Brennan</td>
<td>0.03</td>
<td>0.17</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Gleeson</td>
<td>0.04</td>
<td>0.20</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

### Table 3: Pairwise Correlation Coefficients with the Dissent Rate

<table>
<thead>
<tr>
<th>Variable</th>
<th>Pairwise Correlation Coefficient With Dissent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Urban</td>
<td>0.29</td>
</tr>
<tr>
<td>Caseload</td>
<td></td>
</tr>
<tr>
<td>Total Caseload</td>
<td>-0.15</td>
</tr>
<tr>
<td>Case Per Judge</td>
<td>-0.37</td>
</tr>
<tr>
<td>Discretion</td>
<td>0.20</td>
</tr>
<tr>
<td>Final</td>
<td>0.22</td>
</tr>
<tr>
<td>Griffith</td>
<td>-0.49</td>
</tr>
<tr>
<td>Knox</td>
<td>0.19</td>
</tr>
<tr>
<td>Isaacs/Duffy</td>
<td>0.01</td>
</tr>
<tr>
<td>Latham</td>
<td>0.12</td>
</tr>
<tr>
<td>Dixon</td>
<td>-0.10</td>
</tr>
<tr>
<td>Barwick</td>
<td>0.11</td>
</tr>
<tr>
<td>Gibbs</td>
<td>-0.09</td>
</tr>
<tr>
<td>Mason</td>
<td>0.13</td>
</tr>
<tr>
<td>Brennan</td>
<td>0.13</td>
</tr>
<tr>
<td>Gleeson</td>
<td>0.16</td>
</tr>
</tbody>
</table>
4. Results

Table 4 presents the results from the multiple regression of the dissent rate on the explanatory variables in Table 1 and a constant term.\textsuperscript{44} We begin by discussing how to interpret the results in Table 4 and then proceed to discussing the hypotheses. The first column lists each of the explanatory variables. The columns headed I and II present the results of two alternative sets of regressions. In the column headed I \textit{Case Per Judge} is used to measure caseload and in the column headed II \textit{Total Caseload} is used to measure caseload. Note that as these variables are both being used to measure caseload they are not included in the same regression. \textit{Total Caseload} is excluded from the first set of regression results and \textit{Case Per Judge} is excluded from the second set of regression results. The other variables are included in both sets of regressions reported in I and II. The constant term, which is also reported in Table 4, can be interpreted as the average effect on dissent rates of all variables omitted from the regression model.

For each explanatory variable, for each regression (reported under I and II) we provide the value and sign (positive or negative) on the coefficient and the t-value, which is reported in parenthesis next to the coefficient. To illustrate how to read Table 4, consider \textit{Dissent(-1)}. In the first regression under column I, the coefficient on this variable is 0.33 and the t-value is 3.08. In the second regression under column II, the coefficient on this variable is 0.38 and the t-value is 2.81. The value of the coefficient is an indicator of the responsiveness of the dissent rate to a change in that variable, holding all the other explanatory variables constant. The coefficient on \textit{Dissent(-1)} can be interpreted as follows. Because the coefficient on \textit{Dissent(-1)} has a positive sign in both sets of results an increase in the dissent rate last year will have a positive effect on the dissent rate this year. We can say that for any given year, a 1 per cent increase in the dissent rate last year will result in a 0.33 per cent increase (according to the results in I) or a 0.38 per cent increase (according to the results in II) in the dissent rate this year, holding all the other variables constant.\textsuperscript{45} Note that this is slightly smaller than the corresponding rate for the United States Supreme Court between 1800 and 1994 where Goff found

\textsuperscript{44} Because we are using time series data it is necessary to pre-test each of the variables other than the dummy variables for stationarity. Time series data are often non-stationary. This means that if there is a sudden shock to the data, such as, in the case of the dissent rate, a change in leadership or change in the institutional arrangements on the court, it will tend to drift aimlessly and not return to its equilibrium level. Non-stationary data is sometimes likened to a drunkard’s walk where a drunk, having left a bar, will walk in a certain direction for a distance, stop and then walk in a different direction. If one or more variables are non-stationary the results will not be reliable unless the variable is made stationary. If one or more variables are non-stationary, they can be made stationary by taking the first difference. To investigate whether the variables are stationary we used the Augmented Dickey-Fuller and Phillips-Perron unit root tests with and without a trend, which are the most common tests for this purpose. The two unit root tests gave the same results; namely, \textit{Dissent}, \textit{Dissent(-1)}, \textit{Case Per Judge} and \textit{Total Caseload} are each stationary, while \textit{Urban} is non-stationary. Therefore, in the analysis below \textit{Urban} is estimated in first difference form. The results of the unit root tests are available on request.

\textsuperscript{45} The variables, other than the dummy variables, were expressed in natural logs which has the advantage that the coefficients on the explanatory variables can be interpreted as elasticities.
that, for any given year, a 1 per cent increase in the dissent rate last year results in about a 0.50 per cent increase in the dissent rate this year.\footnote{46} The coefficients on the other explanatory variables in Table 4 for both sets of regressions can be interpreted in a similar fashion.

The t-value, which is given in parenthesis, measures the statistical significance of the coefficient or, to put it differently, the reliability of the estimate of the coefficient.\footnote{47} Standard levels of statistical significance are 1 per cent, 5 per cent and 10 per cent. These are denoted in Table 4 by one, two and three asterisks respectively, next to the t-values.\footnote{48} If a coefficient is statistically significant at 1 per cent, it can be regarded as highly significant, at 5 per cent moderately significant and at 10 per cent weakly significant.\footnote{49} If a coefficient is not statistically significant we can conclude that the variable does not have any effect, which can be reliably measured, on the dissent rate. Thus, the variables which can be regarded as having an important effect on the dissent rate, either positive or negative, in Table 4 are those variables which are statistically significant.

Overall, the fit of the model is reasonably good. The adjusted coefficient of determination (which is depicted in Table 4 as Adjusted R\(^2\)) is a measure of how much of the variation in the dissent rate can be explained by variation in the explanatory variables. The adjusted coefficient of determination is 0.41 in column I and 0.39 in column II, which suggests that about 40 per cent of the variation in the dissent rate can be explained by the specified model. Table 4 also presents the F-statistic, which is a measure of the overall significance of the regression. It is statistically significant at the 1 per cent level in the regression results reported in both columns I and II, which suggests that the explanatory variables are jointly important in explaining variations in the dissent rate.\footnote{50}

---

\footnote{46} Goff, above n23.  
\footnote{47} The reported t-values are adjusted for heteroskedasticity using Newey-West robust estimator.  
\footnote{48} The asymptotic critical values are 1.64 (10 per cent), 1.96 (5 per cent) and 2.32 (1 per cent).  
\footnote{49} Technically, each coefficient represents a point estimate of an unknown population. Statistical significance is an indicator of the reliability of the estimate. Instead of relying on the point estimate alone we can construct an interval around the point estimator, such that this interval has \(\alpha\) per cent probability of including the true parameter value. If the coefficient is significant at 1 per cent, there is a 99 per cent probability that the interval contains the true value; if the coefficient is significant at 5 per cent, there is a 95 per cent probability that the interval contains the true value, and if the coefficient is significant at 10 per cent, there is a 90 per cent probability that the interval contains the true value.  
\footnote{50} Formally, the results from the F test mean we can reject the null hypothesis that the true slope coefficients are zero. In addition standard tests for serial correlation and stationarity of the residuals, not reported in Table 4, indicate that there are no statistical problems with the model.
We now discuss each of the hypotheses set out above in light of the empirical results. The first two hypotheses are that the dissent rate will be inversely related to the caseload of the Court (Hypothesis 1) and that the introduction of case selection discretion will result in higher rates of dissent (Hypothesis 2). Beginning with Hypothesis 1, both measures of caseload have the expected negative sign. Case Per Judge is statistically significant at the 10 per cent level in the first specification, while Total Caseload is statistically insignificant in the second specification. This suggests that an increase in caseload does have a reliably measurable negative effect on the dissent rate when we control for changes in the number of justices on the Court. A 1 per cent increase in annual caseload, controlling for the number of justices on the Court, results in a 0.33 per cent reduction in the dissent rate. There is, however, no support for Hypothesis 2. While the sign on Discretion is negative, it is statistically insignificant in both specifications. The introduction of case selection discretion has not had any reliably measurable effect on the High Court’s dissent rate, holding each of the other explanatory variables constant.

Table 4: Multiple Regression Results (Disent is Dependent Variable)

<table>
<thead>
<tr>
<th>Variable</th>
<th>I</th>
<th>II</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>2.26 (4.67)*</td>
<td>2.19 (3.23)*</td>
</tr>
<tr>
<td>Dissent(-1)</td>
<td>0.33 (3.08)*</td>
<td>0.38 (2.81)*</td>
</tr>
<tr>
<td>Urban</td>
<td>-9.47 (1.30)</td>
<td>-8.37 (1.21)</td>
</tr>
<tr>
<td>Total Caseload</td>
<td>_</td>
<td>-0.19 (1.36)</td>
</tr>
<tr>
<td>Case Per Judge</td>
<td>-0.33 (1.75)***</td>
<td>_</td>
</tr>
<tr>
<td>Discretion</td>
<td>-0.03 (0.36)</td>
<td>-0.03 (0.33)</td>
</tr>
<tr>
<td>Final</td>
<td>-0.08 (3.08)*</td>
<td>-0.09 (3.53)*</td>
</tr>
<tr>
<td>Griffith</td>
<td>-0.27 (2.04)**</td>
<td>-0.28 (1.80)***</td>
</tr>
<tr>
<td>Knox</td>
<td>0.20 (1.67)***</td>
<td>0.20 (1.69)***</td>
</tr>
<tr>
<td>Isaacs/Duffy</td>
<td>0.15 (1.35)</td>
<td>0.09 (0.78)</td>
</tr>
<tr>
<td>Latham</td>
<td>0.20 (1.50)</td>
<td>0.16 (1.15)</td>
</tr>
<tr>
<td>Dixon</td>
<td>0.15 (0.94)</td>
<td>0.12 (0.79)</td>
</tr>
<tr>
<td>Barwick</td>
<td>0.13 (1.52)</td>
<td>0.14 (1.60)***</td>
</tr>
<tr>
<td>Mason</td>
<td>0.26 (2.75)*</td>
<td>0.30 (3.45)*</td>
</tr>
<tr>
<td>Brennan</td>
<td>0.20 (1.01)</td>
<td>0.28 (1.62)***</td>
</tr>
<tr>
<td>Gleeson</td>
<td>0.22 (1.32)</td>
<td>0.26 (1.61)***</td>
</tr>
<tr>
<td>R²</td>
<td>0.50</td>
<td>0.47</td>
</tr>
<tr>
<td>Adjusted R²</td>
<td>0.41</td>
<td>0.39</td>
</tr>
<tr>
<td>F-statistic</td>
<td>5.76 [0.01]</td>
<td>5.30 [0.01]</td>
</tr>
</tbody>
</table>

Notes: Gibbs was treated as the reference point for the leadership dummy variables. Figures in round parenthesis are Newey-West Heteroskedastic-consistent t-values. Figures in square parenthesis are probability values. *(**)(***) indicates coefficients are statistically significant at the 1%(5%(10%) statistical significance level.
The third hypothesis concerns the effect of the *Australia Acts* on the dissent rate on the High Court. We hypothesised that there are competing considerations suggesting that the *Australia Acts* could have either a positive or negative effect on the dissent rate. Final is statistically significant at 1 per cent with a negative coefficient in both specifications. This means that the introduction of the *Australia Acts* has had a strong negative effect on the dissent rate, holding each of the other explanatory variables constant.

The fourth hypothesis is that judicial decision-making is influenced by the prevailing norms of decision-making on the Court. Therefore, for any given year, an increase in the dissent rate last year will have a positive effect on the dissent rate this year. There is strong support for this hypothesis. As noted earlier a 1 per cent increase in the dissent rate last year will result in a 0.3–0.4 per cent increase in the dissent rate this year. This result is statistically significant at the 1 per cent level in columns I and II making it reliable.

The fifth hypothesis examines the effect of leadership on the Court. The coefficient on Griffith has a negative sign, while the coefficient for the other chief justice variables have a positive sign. If the tenure of a particular chief justice has an important effect on dissent rates, relative to the benchmark, which is Gibbs’ term, this will be reflected in a statistically significant coefficient for that chief justice’s tenure. The dummy variables for Griffith, Knox and Mason are statistically significant in the results in both columns I and II, while the dummy variables for Barwick, Brennan and Gleeson are statistically significant in the results in column II. This suggests that, holding the effect of the other explanatory variables on the dissent rate constant, we can reliably conclude in a statistical sense that the dissent rate was lower when Griffith was chief justice relative to when Gibbs was chief justice, while the dissent rate was higher under the leadership of Knox, Barwick, Mason, Brennan and Gleeson, relative to Gibbs’ term as chief justice.

The findings for the Griffith, Knox and Gleeson dummy variables are expected. The dissent rate on the Griffith Court, particularly the early Griffith Court, was very low relative to later years, while Knox presided over the first sharp increase in the dissent rate in the second half of the 1920s. The dissent rate also spikes in the late 1990s. This seems to reflect a ‘Kirby effect’ with just under one-third of Kirby J’s judgments being in dissent, which makes him the highest dissenter in the history of the Court.51 Interestingly, the spike in the dissent rate in the first half of the 1930s under Gavan Duffy and in the mid-to-late 1940s under Latham does not show up in the form of statistically significant effects. Similarly, the leadership of Dixon does not have a statistically significant negative effect on the dissent rate relative to the leadership of Gibbs.

The final hypothesis is that an increase in the urbanisation rate will have a positive effect on the dissent rate. There is no support for this hypothesis with the

---

urbanisation rate statistically insignificant in both specifications. Thus, in a statistical sense, the urbanisation rate has no measurable effect on the dissent rate. This finding is consistent with the most recent study of the determinants of dissent on State Supreme Courts in the United States by Paul Brace and Melinda Hall,\(^\text{52}\) which also found that the urbanisation rate had no reliably measurable effect. Earlier studies which have found that urbanisation has a positive effect on dissent have used cross-sectional data. Brace and Hall suggest: ‘It is likely that whatever relationships, which were found between [urbanisation] and dissent rates in the past were time bound’,\(^\text{53}\) meaning that they do not take account of the time series dimension. Another possibility is that while the urbanisation rate has been widely used, it might be too crude to capture socio-political complexity.

5. Conclusion

A clear picture that emerges from this study is that the Court’s institutions are better predictors of the dissent rate than the external environment. While the urbanisation rate does not have a statistically significant effect on the dissent rate, several factors internal to the Court are important predictors. These are norms on the Court as to acceptable dissent, with an increase in the dissent rate last year resulting in a higher dissent rate this year for any given year, caseload after controlling for changes in the number of justices throughout the Court’s history, and the enactment of the *Australia Acts* removing all appeals from Australian courts to the Privy Council. There are also statistically significant differences in the dissent rate under the leadership of different chief justices.

This study should be seen as a first step in unravelling the role that environmental and institutional factors have had on dissent in Australian courts. There are at least two avenues open to future research. First, future analyses could incorporate more complex and varied environmental factors. A potential problem is that some of the external factors which have been found to have the most predictive power in the United States, such as partisan competition, are clearly not relevant to Australia where judges are not elected. However, other factors which might be used to examine increased socio-political complexity and changing social attitudes in lieu of urbanisation could include indicators such as immigration levels reflecting ethnic diversity or possibly even divorce rates. Second, more sophisticated modelling could be employed to combine institutional variables with judges’ personal characteristics. While existing evidence on the application of the attitudinal model to the High Court suggests it might not have much power, the modelling was fairly crude.\(^\text{54}\) Future studies could exploit the cross-sectional and time series dimensions of larger data sets to examine how personal characteristics, institutions and external factors interact in influencing variations in the dissent rate.

\(^{52}\) Brace & Hall, above n4.

\(^{53}\) Brace & Hall, above n4 at 65.

\(^{54}\) See Smyth, ‘Explaining Historical Dissent Rates in the High Court of Australia’, above n3.
Abstract

The Amcor Case currently before the High Court raises the question of the approach which should be taken to the interpretation of agreements made under the federal system of enterprise-based collective bargaining. The immediate issue is whether employees of a company whose business had been transferred to a related company became entitled to redundancy payments when their employer finally decided to cease employing them, even though the employees were immediately hired by the new business operator with no loss of entitlements. The case is therefore centrally concerned with the understanding of the term ‘redundancy’ under the certified agreement in question. This issue depends on the method used in determining the meaning of the agreement: whether it should be construed more like a contract or a statutory instrument. The court will be faced with the appellant employer’s argument that the agreement should be read according to the generally assumed meaning of redundancy as a form of termination resulting in unemployment. Against this is the respondent union’s view (accepted by the Full Federal Court) that the meaning of the agreement is limited to its strict terms, whether or not the employees became unemployed.
1. Introduction

‘Yet again it is necessary to determine the proper construction of an industrial agreement.’\(^1\) So wrote Finkelstein J of the Federal Court, with perhaps a touch of weariness, when the redundancy provisions of a certified agreement came before him for interpretation in May 2002. His decision was received by many employment lawyers with surprise, as it appeared to provide that several hundred employees who had simply been transferred between two related companies without losing their jobs had each been made redundant and were accordingly entitled to payments of many thousands of dollars (the first individual applicant was awarded over $88,000). Did the decision mean that henceforth corporate restructurings which resulted in a change in the direct employer were subject to significant payments for mass redundancies? What exactly was a redundancy if this situation was classed as one, even though the employees continued in employment performing the same tasks?\(^2\) The surprise continued when the decision was unanimously upheld on appeal in March 2003.\(^3\)

The matter has now found its way to the High Court.\(^4\) The specific question at issue is whether Amcor’s treatment of the employees in question constituted ‘redundancy’. This is not, however, simply to be answered by determining the general usage of a common industrial term. As the dispute revolves around the effect of a clause contained in a certified agreement made under the Workplace Relations Act 1996 (Cth) (the WR Act), the case provides an opportunity for the court to examine the approach which should be applied to the interpretation of such agreements. Under the award system of industrial regulation, questions of interpretation and application were frequently resolved by the award-making tribunal making a determination or variation which clarified the issue in dispute. The tribunal usually had the benefit of knowing the background to the dispute and, as the framer of the award, the purpose which the award had been designed to achieve. Now that agreements have replaced awards as the predominant form of regulation at the workplace, questions of their interpretation are an increasingly common task for the courts. The functions of drafting and interpretation have become separated, creating a need for consistent principles to be used for determining the meaning of agreements.

\(^1\) Construction, Forestry, Mining and Energy Union v Amcor Ltd [2002] FCA 610; (2002) 113 IR 112 (hereinafter Amcor) at [1].

\(^2\) Peter Punch, ‘Redundancies Cost Amcor’ (2002) 6 Australian Industrial Law News (June): ‘Clearly the outcome is unfair. The employees had become entitled to a benefit that was designed to compensate them for losses arising out of termination of their employment by reason of redundancy, but those losses had not actually been sustained by them.’ See also Joe Catanzariti, ‘Who is Liable for Redundancy following Transmission of Business?’ (2002) 40(6) LSJ (July) 36.


2. Background to the Dispute

Until 1998, Amcor Ltd owned and operated paper manufacturing mills in four locations in New South Wales, Queensland and Tasmania. As a result of a corporate restructuring in June of that year, Amcor sold its two mainland paper mills to a wholly owned subsidiary, Paper Australia Pty Ltd. While Paper Australia owned and operated the mills, Amcor remained the employer of the employees who worked there. Under a subsequent agreement which took effect from the date of the sale, Paper Australia undertook to carry out all Amcor’s obligations in relation to the employees.

Then in February 2000, Amcor announced that it would separate its paper making business from its other manufacturing operations. Amcor’s shares in Paper Australia were transferred to a new company, Paperlinx Ltd (which was not wholly owned by Amcor). Amcor’s two remaining paper mills in Tasmania were sold to Paper Australia, which was by now a wholly owned subsidiary of Paperlinx. In order to complete the transfer, in February Amcor wrote to each of the paper mill employees notifying them that their employment with Amcor would end in six weeks’ time, and enclosing a letter offering employment with Paper Australia ‘on the same terms and conditions as you currently enjoy.’ The letter included a statement that ‘all benefits will be preserved, including continuity of service for all employment-related purposes, salary/wage, superannuation and accrued leave entitlements.’ The letter stated that the offer could be accepted by the employees turning up for work as normal from 1 April.

In 1997 a certified agreement covering the paper mills had been concluded between Amcor and the Construction, Forestry, Mining and Energy Union (CFMEU). The agreement, made under the WR Act, commenced effect on its certification by the Australian Industrial Relations Commission in June 1998, with a nominal expiry date of two years. Under the agreement (clause 55.4), Amcor undertook to give the maximum possible notice to the union of changes which permanently affected employment, and at least one month’s notice to each employee whose employment would be affected. Amcor was clearly concerned to minimise disruption and labour turnover which might result after major changes had been announced. While there was no provision for consultation with the union over the implementation of change, Amcor and the CFMEU agreed to co-operate in finding alternative employment and retraining for retrenched workers (clause 55.7).

The relevant provision of the redundancy clause of the agreement provided that ‘should a position become redundant and an employee subsequently be retrenched,’ the employee would be entitled to payments for accumulated sick leave, annual leave and long service leave, as well as redundancy pay based on the length of the employee’s service (clause 55.1). The redundancy clause went on to provide for the situation where an employee became redundant and was transferred to a lower paid job (clause 55.2). There was a further provision for ‘an employee who has opted for transfer to another classification in lieu of retrenchment’ (clause 55.3). So the agreement seemed to contemplate a two-stage redundancy process: a
position was first declared redundant, and then an employee occupying that position would be retrenched or could alternatively be transferred with their consent to another position (which might be lower-paid than their former position or involve a change of location). There was nothing in the agreement which expressly covered the situation where an employee’s engagement with Amcor was terminated yet they continued to work for another employer at the same job with the same pay and conditions as before.

Once it heard of the letters, the union began proceedings to enforce the agreement, claiming that the employees had been made redundant by Amcor and should be paid their entitlements in accordance with the plain meaning of the agreement. At least part of the union’s motivation for taking this action seems to have been a concern that it had not been consulted prior to the changes. It claimed that neither the consultation nor the dispute resolution provisions contained in the agreement were used by Amcor in the redundancy process. Amcor’s response was that the redundancy provisions were never intended to apply to situations where employees continued in employment with no real change of status or loss of access to their accrued entitlements. It argued that as a matter of reality the positions of the employees were not made redundant: they continued to be required, though not by Amcor.

3. The Decision at First Instance and On Appeal

The matter came before the Federal Court in its original jurisdiction by way of a summons for imposition of a penalty against Amcor for breaching the certified agreement by failing to pay the redundancy entitlements. In his decision, Finkelstein J began by considering the meaning of the word ‘redundant’. Based on previous authority, he found that the usual understanding is that an employee becomes redundant when their employer no longer requires their services for reasons unconnected to the individual employee’s performance. The formula which has frequently been applied is that ‘a job becomes redundant when the employer no longer desires to have it performed by anyone’.5 Previous decisions had treated an employee as redundant even where the employment had terminated because of a transfer of the business and the employee had been given assistance in finding a new job or an offer of employment had been made by the new owner of the business.6 The crucial issue was whether the employment contract with the existing employer had been terminated. Once the existing relationship had ended for ‘economic’ rather than performance reasons there was a redundancy. Whether the employee subsequently obtained employment, immediately or not, was irrelevant.

5 The Queen v Industrial Commission of South Australia; Ex Parte Adelaide Milk Supply Co-Operative Ltd (1977) 16 SASR 6 at 8 (Bray CJ). This is also the common definition in awards: see, for example Hospitality Industry, Accommodation, Hotels, Resorts and Gaming Award 1998 (Australian Industrial Relations Commission – hereinafter AIRC), AW783479 cl16.1.
Looking at the redundancy provisions of the agreement itself, Finkelstein J saw the two-stage process as involving two separate conditions: first, a determination by the employer that the employee’s position was redundant, and secondly the retrenchment of the employee because of the redundancy. If the act of retrenchment or termination occurred by reason of redundancy, both these conditions were met, so the employee became entitled to a redundancy payment including accrued leave entitlements. According to Finkelstein J this process indicated that the question of redundancy was to be determined at the time of the termination of the employee’s existing employment.7 Presumably this was because the drafting of the redundancy provisions in this way indicated that the intentions of the drafters was to focus on the situation at a particular time, the time of dismissal. This approach gained further support, according to Finkelstein J, from the fact that retrenchment was envisaged as meaning termination of employment with the particular employer who was party to the agreement, namely Amcor. This was clear from the agreement’s use of employment as meaning employment with ‘the Company’, ie Amcor, as well as the use of ‘position’ as referring specifically to employment with Amcor.8

Finkelstein J agreed that a decision in favour of the employees and the CFMEU might be regarded as unfair and ‘an affront to commonsense,’ since in practical terms the situation of the employees was unchanged.9 Nevertheless his Honour felt bound to make it. The terms of the agreement itself, considered as a whole, impelled him to conclude that a redundancy had occurred and the employees were entitled to claim their payments. In a case like this, when the circumstances giving rise to the dispute seem not to have been contemplated by the drafting parties, the court was limited in the extent to which it could provide the missing piece to the puzzle. The judge noted that when it comes to the interpretation of statutes, it is nowadays accepted that courts have some scope to remedy an omission provided that the legislature’s intention is clear.10 However a similar power to fix a hiatus in an agreement could not be applied in this case as Finkelstein J considered it by no means certain what the parties to the agreement would have agreed to if they had turned their minds to the issue at hand.

Amcor’s appeal against the decision was dismissed unanimously by the Full Court consisting of Moore, Marshall and Merkel JJ. All judges agreed that the proper interpretation of the agreement led to the conclusion that it was the termination of employment with Amcor which caused the employees to be considered redundant under the agreement, entitling them to be paid their accrued benefits and severance pay. In the words of Marshall and Merkel JJ, Amcor’s obligations to its employees under the agreement ‘crystallised upon their dismissal’.11

---

7 Amcor, above n1 at [16] (Finkelstein J).
8 Ibid.
9 Id at [1], [17] (Finkelstein J).
10 The extent and justifiability of such a power remain controversial: see Derek Auchie, ‘The Undignified Death of the Casus Omissus Rule’ (2004) 25 Statute LR 40.
11 Amcor FC, above n3 at [33].
The joint judgment of Marshall and Merkel JJ (with whom Moore J agreed) advanced several clear reasons supporting the appropriateness of the employees being entitled to the redundancy payments, and challenging the argument that such a result was unfair. Firstly, the employees were not claiming entitlements which they had given up. In choosing to take up employment with Paper Australia, the employees had not waived their redundancy entitlements under the agreement with Amcor. Such an approach was inconsistent with the status of a certified agreement, which obtains its force from statute, independent of any individual agreement. Employees cannot consent to receive less than their entitlements under an award or certified agreement, a principle recently reinforced by the Metropolitan Health Service Board Case.

Furthermore, in taking up their entitlements the employees were not gaining a benefit unjustly, as there was no possibility of ‘double dipping.’ Employees who obtained payment for accrued long service and sick leave from Amcor could not later claim exactly the same entitlements from their new employer, Paper Australia, under that company’s promise to preserve employees’ benefits. In their Honours’ opinion, the benefits which were preserved would only include outstanding entitlements not already paid by Amcor. No doubt any later claim for benefits already paid by the previous employer would be defeated by the argument that the entitlements had been discharged by payment. Under this approach, by invoking the redundancy provisions the employees were simply obtaining accelerated access to their accrued entitlements, and obtaining them from the employer in whose service they earned them.

On the meaning of ‘redundancy,’ Marshall and Merkel JJ considered the crucial issue to be whether Amcor as employer had decided that particular work was no longer required to be done for it by any of its employees. If the work was subsequently required by another employer, this was irrelevant to the operation of the agreement. Their Honours treated the employment relationships regulated by the agreement as entirely separate from any subsequent work relationship which might arise. The only employment relationships relevant to the operation of the agreement were those between Amcor, as the employer party to and bound by the agreement, and its employees. Amcor’s obligations under the agreement could not be affected by the employees entering into an employment relationship with another employer. Although they did not express it in this way, their Honours’ approach is a contractual one, redolent of the doctrine of privity of contract.

A slightly different approach was taken by Moore J, for whom a key issue was the nature and source of the entitlements. Most of the redundancy payments under the agreement related to ‘unrealised benefits that the employee has accumulated in

---

12 Josephson v Walker (1914) 18 CLR 691 at 700 (Isaacs J); Gapes v Commercial Bank of Australia Ltd (1979) 41 FLR 27 at 29 (Smithers & Evatt JJ).
13 Amcor FC, above n3 at [34]; Metropolitan Health Service Board v Australian Nursing Federation (2000) 99 FCR 95 at 104 (French J).
15 Id at [41]–[42].
his or her employment with Amcor.\textsuperscript{16} The fact that they originated in the employment relationship with a particular employer, and accrued as that relationship continued, indicated that the parties had intended the benefits to become payable once that relationship had ended.

All the judges made it clear that Amcor’s action in obtaining a promise of continued employment from Paper Australia could have no bearing on how the agreement was to be interpreted. Consistent with contract interpretation, subsequent conduct could not be used as evidence of the intention of the parties at the time they entered into the agreement, which was assumed to be the relevant time for ascertaining intention. This was particularly the case as only one party’s conduct was involved.\textsuperscript{17}

One argument relied on by Amcor was that when the standard award redundancy clause was devised by the then Australian Conciliation and Arbitration Commission in the \textit{Termination, Change and Redundancy Case} (the \textit{TCR Case}), it had not been intended to apply in the case of a transmission of business.\textsuperscript{18} This intention, Amcor argued, should govern the interpretation of the agreement currently before the court. This approach was rejected as no such intention was found in the terms of the agreement itself. There was nothing in the agreement to indicate that the \textit{TCR Case} standard was being adopted. Nor could any intention be derived from the subsequent conduct of the parties to the agreement. The court did not discuss whether Full Bench principles should govern the interpretation of a particular agreement as an independent standard; rather, as in contract law, generally it was the mutual intentions of the parties at the time of the agreement which mattered.

Finally, Marshall and Merkel JJ denied that the interpretation of the agreement should be guided by parties’ perceptions of the fairness of the outcome when the agreement was applied to particular situations. Their Honours noted that other ways of affecting the operation of the agreement had been open to Amcor. The company could have sought to vary the agreement by negotiation with the unions party to it and with the approval of the Commission. The \textit{WR Act} also allowed the Commission to remove any ambiguity or uncertainty in the Agreement.\textsuperscript{19} That Amcor had not taken advantage of these opportunities negated, in their Honours’ view, the argument that it should be relieved of the particular obligations under the agreement in this case on the ground of unfairness.

4. Legal Issues

A. Redundancy

Several rationales have been advanced in support of redundancy or severance pay. The early approach taken by the NSW Industrial Commission, in the immediate context of the severe 1982 recession, was to consider such pay as a cushion against

\begin{footnotesize}
\begin{itemize}
\item \textsuperscript{16} Id at [5] (Moore J).
\item \textsuperscript{17} Id at [6] (Moore J), [48] (Marshall & Merkel JJ).
\item \textsuperscript{18} \textit{Termination, Change and Redundancy Case} (1984) 8 IR 34 (hereinafter \textit{TCR Case}) at 75.
\item \textsuperscript{19} \textit{WR Act} s170MD.
\end{itemize}
\end{footnotesize}
the immediate financial effects of unemployment while the employee looks for other work, rather than as a compensation for lost job security based on past service.\textsuperscript{20} A somewhat different view was preferred by the federal Commission in the \textit{TCR Case}: severance pay was primarily directed at ameliorating the ‘inconvenience and hardship’ of sudden job loss rather than as direct income support in the search for new employment. Based on previous decisions, the Commission also justified it as providing compensation for loss of non-transferable credits such as sick and annual leave.\textsuperscript{21} The Commission accepted a qualification to this entitlement where the employee’s income was not completely severed. The Commission emphasised that it did not ‘envisage severance payments being made in cases of succession, assignment or transmission of a business.’ Accordingly, a standard clause was inserted which provided that in the case of transmission an employee’s service was carried over to the new employer for the purpose of calculating severance pay.\textsuperscript{22} However the former employer was not automatically relieved from the redundancy pay obligation. If they obtained suitable alternative employment for their redundant employees, they could apply for relief from such obligations in part or full. Whether the new employer recognised all the employee’s previous service would be a significant consideration for exercise of discretion in such a case.

The approach to redundancy pay taken in the \textit{TCR Case} has continued to this day, and was reiterated in the recent \textit{Redundancy Test Case}.\textsuperscript{23} The current standard federal award provisions for redundancy continue to allow an employer in a specific situation to apply for a variation of the general severance pay amounts ‘if the employer obtains acceptable alternative employment for an employee.’\textsuperscript{24} The Commission routinely grants an exemption to an employer who has been instrumental in obtaining work for an employee which is reasonably suitable for their personal situation. The Full Bench of the Commission has held that it is inappropriate to make an order for severance pay in the case of the sale of a business and transfer of employees when the employees had agreed to the transfer and the new employer recognised their accrued entitlements. Where there is a simple transmission of business the Commission has recently stated that the TCR severance pay standard should not apply.\textsuperscript{25} However the Commission takes this approach in exercise of its discretion: the result is not automatic and depends on an assessment of the particular redundancy situation in question.\textsuperscript{26}

\textsuperscript{20} Shop, Distributive and Allied Employees’ Association (NSW) v Countdown Stores (1983) 7 IR 273 at 273, 293 (Fisher P).
\textsuperscript{21} \textit{TCR Case}, above n18 at 71–73.
\textsuperscript{22} Id at 75, 76; \textit{Termination, Change and Redundancy Case, Supplementary Decision} (1984) 9 IR 115 at 129.
\textsuperscript{23} \textit{Redundancy Case} (AIRC, Full Bench, 26 March 2004) PR032004 at [133]. The Commission restated the position that redundancy pay was not designed to supplement income loss, but was for hardship and loss of credits. It was these factors which justified an increase in pay entitlements for those with more than four years’ service: id at [138], [154].
\textsuperscript{24} Hospitality Industry, Accommodation, Hotels, Resorts and Gaming Award 1998 (AIRC) AW783479 cl16.5.
It would appear, therefore, that industrial practice is against severance pay being available in the case of a simple transmission of business when there is no loss of continuity in work, especially if the new employer recognises leave entitlements accrued in service with the former employer. That is so when standard redundancy provisions apply. But in this case the redundancy provisions in the Amcor agreement were rather unusual. They were briefer and less qualified than is the norm. The agreement did not contain the standard clause deeming prior service to be carried over in a transmission of business. There was no statement of objects or even an express definition of redundancy. It also seems that there was no regulation of redundancy in the underlying industry award. If this is so, then any entitlements on redundancy (apart from general notice periods for termination) depended solely on the agreement. This situation also tends to negate the view that the parties must have negotiated with the TCR Case award standard in mind. Is there anything in this particular agreement to indicate that the severance payment was intended to compensate only for lost entitlements and the hardship of unemployment? Unless we are to apply standards and notions extraneous to the agreement itself, it is difficult to ascribe any particular purpose to the terms. All that can be relied on is their plain meaning in context.

Finkelstein J’s view was that redundancy means the termination of the existing employment relationship. The fact that in several cases, including the TCR Case itself, an exception was allowed where there was a transfer of the business with no loss of employment, merely confirmed this point in his opinion, since there would be no need to make such an exception if it did not otherwise result in redundancy. This view is supported by the difference between making and interpreting. In such cases the tribunals concerned were acting in their capacity as arbitrators deciding what future legal relations should pertain between parties to a wide-ranging interests dispute. They were not acting as courts of construction. In the present situation, the court’s task was strictly that of an interpreter, charged with finding meaning according to the presumed intention of the framers of the agreement.

Against this, it can be argued that there is a significant difference between a situation where an employer makes an employee redundant while offering to assist them in finding work, and the position (as here) where the employer sells the business but takes steps to ensure that all employees will be offered jobs with the new employer. In this case, Paper Australia had not only assumed liability for the employees’ accrued entitlements, but made a clear offer of employment to all

25 Automotive, Food, Metals, Engineering, Printing and Kindred Industries Union v United Milk Tasmania Ltd (AIRC, Full Bench, 23 June 2000) Print S7351; Re Steppes t/as The Beaumont Darwin v Australian Liquor, Hospitality and Miscellaneous Workers Union (1998) 86 IR 337. While this approach is reached in exercise of the Commission’s discretion, a similar result has been reached in the interpretation of a certified agreement on the basis that redundancy pay in the case of business transmission was not within the contemplation of the parties: see Stones v Simplot Australia Pty Ltd (Industrial Relations Court of Australia, Ryan JR, 30 June 1997).

26 Commonwealth Bank of Australia v Finance Sector Union of Australia (AIRC, Duncan SDP, 4 December 2002) PR925304 at [85].

27 The current award does not contain any specific severance provisions, apart from the standard termination clauses: Pulp and Paper Industry – Production Award 1999 (AIRC) AW792323.
Amcor’s employees once they had been made redundant by Amcor. The assumption of responsibility for the existing entitlements was contained in a contract between the two companies, which the employees could not enforce under the doctrine of privity of contract. However the employees’ new employment contracts with Paper Australia were made on the express promise of recognition of all benefits and continuity of service. The employees’ right to claim redundancy pay against Paper Australia for their prior service with Amcor now simply rested on a contractual basis rather than by virtue of the certified agreement.

B. Successorship

It is here that the successorship provisions of the WR Act become relevant. It is not altogether clear whether they would protect the employees’ entitlements when there is a time delay between the transmission of business and the change in employer. The provisions in the WR Act dealing with successor employers appear to contemplate that these two events are simultaneous, since it states that the new employer becomes bound by a certified agreement only from the time that the employer becomes the successor or transmitee of the business, and then only to the extent that the agreement relates to the transmitted business.28 At the time Paper Australia took over the paper manufacturing business of Amcor it did not yet employ any workers covered by the agreement, so the agreement did not apply to it as an employer. Arguably, however, once it became the employer, Paper Australia became bound by the agreement, which by then related to the business which Paper Australia was operating. Amcor continued to be bound by the agreement as long as it remained the employer, even though it no longer operated the business. Presumably Amcor also continued to be liable under the agreement for entitlements accrued prior to the transfer of the business. The issue was raised but not decided by Finkelstein J.29

This case also raises a broader issue in relation to successorship. Recent decisions have highlighted the limitations of the current provisions to outsourcing arrangements. An employer which takes over the outsourced work may not be bound by an award or agreement which covered the work if the character of its business as a whole is found not to be substantially identical to that of the outsourcing employer.30 This will commonly occur in outsourcing arrangements, since one of the principal purposes of such moves is to hive off ‘non-core’ functions to an external supplier which will ordinarily be engaged in a different range of business activities from the outsourcing corporation.

The separation of the business owner and the employer, as happened in Amcor, is of concern because if the successorship provisions do not apply (or apply in a limited way), a corporation might be free to divest its business while remaining the employer, leaving the employees with no effective recourse for recovery of future

28 WR Act s170MB; see also s149(1)(d) in relation to awards.
29 Amcor, above n1 at [6].
entitlements. There would be a similar situation to the waterfront dispute of 1998,31 when assets owned by Patrick’s stevedoring companies were transferred to a parent company within the same group without the employees’ knowledge, leaving the employer companies as ‘empty shells.’ In the present case Amcor retained substantial assets so the employee’s entitlements were never in danger; however there was an attempt to transfer future obligations under the certified agreement to another corporation within the same group. Should this be considered any differently from the everyday situation where a business is sold and the employees are at the same time more or less automatically transferred to the new business owner?

C. Interpretation of Agreements

All the arguments from common industrial practice in relation to redundancy provide support for Amcor’s position, based on principle and policy. But are they relevant? The development of enterprise bargaining over the last decade has resulted in working conditions being predominantly regulated by various kinds of industrial agreements at federal and state level. Such instruments are hybrid in form, bearing elements of both contract and statute. In most cases such agreements are enforceable only by way of civil penalty, with the possible support of an injunction to cease conduct in breach of the agreement.32 Their dual character raises significant questions of interpretation. Should such agreements be construed in the same way as contracts, giving strong weight to the presumed intentions of the parties and the factual background; or should they be treated more like statutes, with primary emphasis on the ordinary meaning of the words in their context, bearing in mind the purpose of the enactment?33

The approach used in the interpretation of awards and agreements is anything but uniform, but has tended to adopt a similar approach to that used in statutory interpretation. It has long been accepted that narrow literalism is not appropriate, and that the interpreter should read an award ‘broadly’ with the aim of giving effect to its intended operation as far as the ‘general intention of the parties can be gathered from the whole award.’34 This approach is probably not very far from that which is now used in the interpretation of statutes,35 although it is not truly purposive since awards tend to be more specific in effect and not so concerned with a wide range of public policy issues. Awards and agreements are also treated

---

32 WR Act ss178, 170NG. However injunctions and declarations under the general law are not available: ACTEW Corporation Ltd v Pangallo [2002] FCAFC 325.
33 It has, however, been argued that the task of interpretation is essentially the same for statutes and contracts: see Michael Kirby, ‘Towards a Grand Theory of Interpretation: The Case of Statutes and Contracts’ (2003) 24 Statute LR 95.
34 George A Bond & Co Ltd (In Liquidation) v McKenzie [1929] AR(NSW) 498 at 503–504 (Street J); Re Hospital Employees (Administrative and Clerical) (State) Award (1982) 2 IR 123 at 124–125 (Glynn J).
‘generously’ since they are often framed by non-lawyers with their practical operation in mind. The common approach to awards has been described as ‘partially statutory and partially contractual,’36 a description which is particularly apt as many awards are drafted by the parties and made by consent.

At the heart of the Federal Court decisions in Amcor was the approach which the judges took to the interpretation of the agreement. The method used was similar to the objective approach conventionally adopted to the construction of contracts, the aim being to determine the mutual intention of the parties as evidenced by the document as a whole and in its factual context.37 As Finkelstein J put it:

The object to be achieved is to discover what was meant by the parties to the agreement. The task does not involve deciding what the parties actually intended … . What must be done is to discover what the parties intended from the meaning that is conveyed by their words, construed in the context in which those words are used.38

In recent years the Federal Court has been developing a consistent approach to the interpretation of certified agreements, one which emphasises their character as consensual documents rather than statutory instruments, and therefore focuses on the objectively determined intentions of the parties.39 Finkelstein J has been a particularly strong exponent of a contractual approach to the interpretation of certified agreements, focussing on the parties’ presumed intentions. Hence his view of the limits to the court’s interpretive function in Amcor:

there comes a point when a court of construction must resist the temptation of forcing a meaning to a bargain which the parties did not intend and to substitute for the arrangements actually made, an arrangement which the court believes is a better one.40

While it is well accepted that the Federal Court is restricted to declaring the true meaning of the award rather than its intended or appropriate effect,41 in Amcor this distinction was given a contractual basis. The contractual approach seems to depart from the more traditional purposive approach taken to awards, and to certified agreements as well.42 The difference in approaches is one of emphasis and there may be no difference in practical outcome in particular cases. Amcor seems to be a case where the difference between statutory and contractual approaches matters a great deal.

37 Codelfa Construction Pty Ltd v State Rail Authority of New South Wales (1982) 149 CLR 337 at 348, 352 (Mason J); Darlington Futures Ltd v Delco Australia Pty Ltd (1986) 161 CLR 500 at 510 (Mason, Wilson, Brennan, Deane & Dawson JJ).
38 Amcor, above n1 at [1] (Finkelstein J).
40 Amcor, above n1 at [18] (Finkelstein J); approved in Commonwealth Bank of Australia v Finance Sector Union of Australia (2002) 125 FCR 9 at 29 (Full Court).
41 City of Waneroo v Holmes (1989) 30 IR 362 at 379 (French J).
It may be that the contractual approach is more suitable to enterprise bargaining under the *WR Act*. If the terms of a certified agreement are to be construed according to judicially identified purposes, and departures from award standards only allowed if the agreement shows a manifest intention to do so, this would undermine the object of the *WR Act* to ensure that ‘the primary responsibility for determining matters affecting the relationship between employers and employees rests with the employer and employees at the workplace or enterprise level.’

5. **Conclusion**

In view of the frequency of disputes over interpretation of agreements and the apparent differences in approach to them, it is hoped that the High Court will give close consideration to the wider interpretation issues raised by the case. Guidance in this area will be significant not only to the Federal Court, but to the federal Commission as well as the state courts and tribunals – not to mention the parties themselves. All are faced with the daily task of trying to derive the legal effect in particular circumstances from statutory collective agreements of varying precision. Inconsistency and uncertainty in their application can only harm co-operative industrial relations and reliance on workplace bargaining.

In some respects, though, industrial parties are moving away from the courts for resolution of differences over the application of agreements. Recent years have seen an explosion in the use of private arbitration clauses for the settlement of such disputes. In most such situations the parties give the Commission the power, in accordance with the *WR Act*, to decide contentious issues. The High Court has held that when it does so, the Commission acts as a private arbitrator and may make binding decisions in exercise of powers and procedures conferred on it by agreement between the parties. Since this position was clarified, private arbitration has become a major aspect of the Commission’s business, providing a significant alternative to the Federal Court in matters involving the application of agreements. The Commission has much greater flexibility in procedure and approach than the Federal Court can exercise. It is also a more appropriate venue for taking into account the industrial background and consequences of decisions. As part of the dispute resolution procedure required in certified agreements, the Commission may be empowered to conciliate disputes with a view to obtaining a mutually satisfactory negotiated outcome. That avenue was not taken by Amcor in this case, although it appears to have been available under the terms of the agreement.

---

42 *Kucks v CSR Ltd* (1996) 66 IR 182 at 184 (Madgwick J); applied in relation to certified agreements in *Australasian Meat Industry Employees Union v Coles Supermarkets Australia Pty Ltd* (1998) 80 IR 208 at 212 (Northrop J); *Hawkins v Commonwealth Bank of Australia (No 2)* (1996) 70 IR 213 at 218. All are cases relied on by Amcor.

43 *WR Act* s3(b).


45 *Amcor* HCA Trans, above n4 at 6–7.
Incorporation of the equitable notion of unconscionability in the *Trade Practices Act* 1974 (Cth) has been fraught with uncertainty. When *Australian Competition and Consumer Commission v Bertatis* came before the High Court for consideration, there were high hopes that this uncertainty would be resolved. In this paper, an outline of how unconscionability is codified in s51AA of the *Trade Practices Act* is followed by a detailed discussion of the *Bertatis* litigation and its implications. The author concludes that the High Court in fact failed to fully delineate the parameters of s51AA, other than to confirm that an inequality of bargaining power between commercial parties is not, in and of itself, sufficient to establish s51AA unconscionability. The paper also questions whether the outcome in *Bertatis* would have been different if s51AC had been in force at the relevant time.

* LLB(Hons), BA(Hons). The author sincerely thanks Joellen Riley for her comments and suggestions on the draft of this article. The views expressed in this article are those of the author.
1. Introduction

The equitable notion of unconscionability is not easily defined. One means of understanding the concept is to separate it into two categories. The broad principle of unconscionability refers to equity’s jurisdiction to grant relief where a stronger party acts against equity and good conscience to enforce, or retain the benefit of, a transaction. This ‘central informing idea of Equity’ underpins numerous equitable doctrines like equitable estoppel, undue influence, misrepresentation, duress and, of course, the narrow doctrine of unconscionability.

The narrow doctrine of unconscionability, a ground for equitable relief in and of itself, is typically characterised by one party being at a ‘special disadvantage or disability’ vis-à-vis another and exploitation by the other party of their superior bargaining position to the disadvantaged party’s detriment.

Unconscionability is a concept now codified in the Trade Practices Act 1974 (Cth) (hereinafter ‘TPA’). This paper begins with an examination of the type of unconscionability incorporated in s51AA. In particular, consideration of that provision in Australian Competition and Consumer Commission v Berbatis (hereinafter Berbatis) is examined in detail. While this litigation provided the first in-depth judicial analysis of s51AA’s scope and operation, discussion of the implications of that analysis demonstrates that critical questions concerning s51AA’s reach remain unresolved. The High Court’s judgment neither directly counters the idea that s51AA unconscionable conduct can arise between commercial parties as a result of ‘situational’ disadvantage, nor does it definitively reject the proposition that s51AA may embrace not only the narrow doctrine, but also the broad principle, of unconscionability. Finally, consideration of whether Berbatis may have been decided differently under s51AC of the TPA highlights that even though s51AC’s full reach remains unclear, its legislative history and judicial interpretation thus far indicate its scope is wider than s51AA. Consequently, s51AC appears to have greater potential to provide small business with better recourse to relief against unconscionable conduct encountered in commercial transactions.

2. Section 51AA

Following much debate, s51AA was introduced in 1992. It provides that ‘a corporation must not, in trade or commerce, engage in conduct that is

---

2 Jeffrey Hackney, Understanding Equity and Trusts (1987) at 17.
4 Commercial Bank of Australia Ltd v Amadio (1983) 151 CLR 447 at 461 (Mason J).
unconscionable within the meaning of the unwritten law, from time to time, of the States and Territories.’

Extrinsic material indicates s51AA was not intended to create new legal rights, but rather to make available to commercial transactions characterised by unconscionability the TPA’s flexible remedies and the possibility of the Australian Competition and Consumer Commission (ACCC) bringing representative actions on behalf of relevant parties. Additionally, such a legislative prohibition was intended to have an ‘educative and deterrent effect’. The legislature envisaged s51AA as codifying the narrow doctrine of unconscionable conduct, as articulated by the High Court in *Blomley v Ryan* and *Commercial Bank of Australia v Amadio*. As such, s51AA was designed to provide relief where ‘one party by reason of some condition [or] circumstance is placed at a special disadvantage vis-à-vis another and unfair or unconscientious advantage is then taken of the opportunity thereby created.’ Traditionally, factors that establish the necessary special disadvantage or disability have been ‘constitutional’ in nature and include, but are not limited to, personal factors like poverty, illness, ignorance, impaired faculties, age, financial need, illiteracy or lack of education.

Despite the legislature’s comments in respect of s51AA’s scope, it became clear following its introduction that s51AA’s reach was in fact not so clear-cut. Uncertainty arose because s51AA’s phrasing does not seem entirely consistent with the legislature’s comments as to its scope and purpose. Reference in s51AA to the ‘unwritten law’ has proved to be a particular concern because, as Baxt and Mahemoff observe, ‘the current limits of the unwritten law are not defined.’ Dal Pont and Chambers elaborate on this point, remarking:

> It cannot be said that that which is *unconscionable within the unwritten law* is limited to the equitable doctrine of unconscionability. Numerous equitable doctrines are grounded, at least in part, in the notion of unconscionable conduct, including estoppel, unilateral mistake, undue influence, economic duress, constructive trusts and relief against forfeiture.

---

7 Ibid.
8 Id at [44].
9 (1956) 99 CLR 362 (hereinafter *Blomley v Ryan*).
10 *Amadio*, above n4 (hereinafter *Amadio*). Explanatory Memorandum accompanying the Trade Practices Legislation Amendment Bill 1992 (Cth) at [41].
11 *Amadio*, above n4 at 462 (Mason J), 423 (Deane J).
12 *Blomley v Ryan*, above n9 at 405 (Fullagar J), 415 (Kitto J). A similar sentiment was expressed in *Amadio*: ‘The adverse circumstances which may constitute a special disability for the purposes of the principles relating to relief against unconscionable dealing may take a wide variety of forms and are not susceptible to being comprehensively catalogued’: *Amadio*, above n4 at 474 (Deane J).
Uncertainty has thus arisen as to whether s51AA is in fact limited to the narrow doctrine of unconscionability articulated in Blomley v Ryan and Amadio. The result, as Horrigan notes, is that ‘[o]ngoing debate about the scope of section 51AA has troubled courts as well as the ACCC,’15 not to mention left those engaging in trade and commerce unclear as to their legal rights.

3. The Berbatis Litigation

A. The Facts

In 1996 a Western Australian couple, Mr and Mrs Roberts, found themselves in a difficult predicament. The lease on their fish and chip business in a shopping centre was to expire in February 1997 and contained no renewal option. The Roberts wished to sell their business, principally due to their daughter’s ill health. However, the sale would not be financially viable unless a new lease could be negotiated with the centre’s owners and managers, CG Berbatis Holdings Pty Ltd and others (hereinafter ‘Landlord’). Further complicating matters, the Roberts were engaged in legal proceedings against the Landlord.16 Something, it seemed, would have to give.

Although prospective purchasers, the Hollands, were found, the sale initially fell through. The Hollands were only willing to buy if a new lease was negotiated. However, the Roberts refused to sign a new lease when the Landlord made it conditional upon the Roberts dropping legal proceedings against the Landlord. The Roberts eventually gained a seven year extension of their lease and, as a result, negotiations with the Hollands resumed. The purchase was again conditional upon assignment of the Roberts’ lease to the Hollands. The Landlord agreed to this assignment, but made it contingent upon the Roberts discontinuing legal proceedings against the Landlord, a clearly non-negotiable condition. The Roberts sought legal advice and were advised against signing the lease on this term. However, feeling they had little other choice, the Roberts did in fact reluctantly sign.

The ACCC commenced proceedings against the Landlord, alleging unconscionable conduct under s51AA of the TPA. Testing s51AA’s scope, the ACCC argued, inter alia, that s51AA embraces special disadvantage based on ‘situational’, as well as ‘constitutional’, factors.17

16 These legal proceedings arose in 1996, when the Roberts and numerous other tenants instituted proceedings against the Landlord in the Commercial Tribunal. They alleged that the Landlord had levied charges in excess of those allowed under the terms of the tenants’ respective leases.
17 Although s51AC also provides protection against unconscionable conduct in commercial transactions, as will be discussed in greater detail in Section 4.B., it was not in force when the ACCC commenced legal action against the Landlord.
Given the uncertainty surrounding s51AA’s scope, as already noted in Section 2, much expectation surrounded *Berbatis*, particularly once it reached the High Court. Speculation centred on two main issues:

i) whether s51AA embraced the broad principle of unconscionability or only the narrow doctrine of unconscionability; and

ii) if limited to the narrow doctrine, whether both ‘situational’ and ‘constitutional’ factors were relevant to the determination of ‘special disadvantage or disability’.

**B. The Decision at First Instance**

In a preliminary ruling on the constitutional validity of s51AA, French J reviewed both s51AA and the general principles of unconscionable conduct.\(^\text{18}\) It was observed that unconscionability in the unwritten law operates at both a generic and a specific level.\(^\text{19}\) At the generic level, unconscionability is the central principle according to which equity operates (the broad principle of unconscionability). At the specific level, unconscionability is a distinct ground of equitable relief, usually associated with notions of unconscionable exploitation and parties being ‘specially disadvantaged’ (the narrow doctrine of unconscionability).\(^\text{20}\) French J noted that the legislature intended ‘the meaning of the unwritten law’ in s51AA to embrace unconscionability as articulated in *Blomley v Ryan* and *Amadio*,\(^\text{21}\) that is, the narrow doctrine of, or specific level, unconscionability. Yet, having pointed to this legislative intention, his Honour cautioned that this ‘may turn out to have been an unduly narrow selection of case law involving the application of the concept of unconscionable conduct.’\(^\text{22}\)

Justice French regarded s51AA as having no settled technical meaning\(^\text{23}\) and characterised ‘the unwritten law from time to time’ as being governed by ‘judge-made rules that can change from time to time.’\(^\text{24}\) French J thus seemed to regard s51AA as capable of embracing the broad principle of unconscionability, rather than simply the narrow doctrine of unconscionability.

In his judgment addressing the facts,\(^\text{25}\) French J first considered whether the Roberts had been at a special disadvantage or disability in the landlord/tenant relationship. Of primary significance was French J’s broad interpretation of the narrow doctrine of unconscionability. His Honour found that ‘special disadvantage or disability’ can be based not only on ‘constitutional’ but also ‘situational’ factors, that is, factors arising from the circumstances relevant to the transaction. Indeed,

\(^{18}\) *Australian Competition and Consumer Commission v CG Berbatis Holdings Pty Ltd* (2000) 96 FCR 492.

\(^{19}\) Id at 502.

\(^{20}\) Ibid.

\(^{21}\) Id at 495, 502–503.

\(^{22}\) Id at 495.

\(^{23}\) Id at 502.

\(^{24}\) Id at 504.

\(^{25}\) *Australian Competition and Consumer Commission v CG Berbatis Holdings Pty Ltd (No 2)* (2000) ATPR 41,165.
it was on the basis of situational factors, in particular the intersection of legal and commercial circumstances in which the Roberts found themselves, that French J found the Roberts to have been specially disadvantaged.26 His Honour added that because the Roberts’ disadvantage was not constitutional in nature, it ‘was not able to be mitigated by the fact of legal representation which they had available to them at all material times.’27

Justice French’s next main concern was whether the Landlord ‘unfairly exploited the tenant’s disadvantage in a manner that would be regarded in equity as unconscionable.28 According to his Honour, such unfair exploitation can arise where an owner uses ‘its bargaining power to extract a concession from the tenant that is commercially irrelevant to the terms and conditions of any proposed new lease.’29 Emphasising that the Roberts’ legal claims against the Landlord were not frivolous or vexatious, French J characterised the condition to drop legal proceedings against the Landlord as ‘commercially irrelevant’ to the Landlord’s lease offer.30 Unsurprisingly, the Landlord appealed French J’s decision.

C. Appeal to the Full Court of the Federal Court
On appeal to the Full Federal Court,31 the relevant principles were not disputed. All parties accepted it was necessary to demonstrate that the Roberts were at a ‘special disadvantage’ in accordance with Amadio unconscionability. The appeal centred on the correctness of French J’s finding that the Roberts had been specially disadvantaged and that the Landlord took unconscientious advantage of that disadvantage.32 In a cross-appeal, the ACCC sought the Court’s opinion on whether the narrow or broad view of unconscionability applies under s51AA, urging that the phrase ‘unwritten law’ in s51AA need not embrace only the narrow doctrine of unconscionability as articulated in Amadio.33

Considering the issue of special disadvantage, the Full Court highlighted that French J did not regard the emotional strain suffered by the Roberts as a result of their daughter’s illness as relevant to the Roberts being deemed specially disadvantaged.34 Instead, the Full Court noted that French J’s finding of special disadvantage was based on the fact that the Roberts’ ability to sell their business had depended on the renewal or extension of their lease. In the Full Court’s opinion, it was inappropriate to characterise as a ‘special disadvantage’ the detriment a tenant suffers as a result of a lease’s imminent expiration,35 explaining

26 Id at 41,197.
27 Ibid.
28 Ibid.
29 Ibid. [Emphasis added.]
30 Ibid.
31 CG Berbatis Holdings Pty Ltd v Australian Competition and Consumer Commission (2001) ATPR 43, 179 (hereinafter Berbatis (No 3)).
32 Id at 43,191.
34 Berbatis (No 3), above n31 at 43,191.
35 Id at 43,192.
that ‘[a]ny proprietor of a business carried on in leased premises, where the goodwill of that business depends on its location, is in precisely the same position as the Roberts.’

The Full Court then examined French J’s decision that the Landlord acted unconscionably by making the lease’s extension or renewal conditional upon the Roberts discontinuing bona fide and serious legal proceedings against the Landlord. Their Honours questioned French J’s distinction between seeking a release from frivolous or vexatious litigation and seeking a release from bona fide and serious claims, finding no difference between the two. The Full Court disagreed with French J’s finding that the Landlord’s condition was commercially irrelevant to the proposed lease’s terms, instead characterising the Landlord’s lease offer as ‘a lifeline’. The Roberts were given a choice where previously none had existed: either continue legal proceedings against the Landlord and lose the opportunity to sell their business (initially the Roberts’ only option), or abandon legal claims and sell the business. Hill, Tamberlin and Emmett JJ thus drew a distinction between the opportunistic striking of a hard bargain and s51AA unconscionable conduct. Highlighting that the Roberts were financially better off for having accepted the renewed lease on the Landlord’s terms, the Full Court overturned French J’s decision.

Of significance is that the Full Court upheld the appeal by reaching a different conclusion on the facts. Unlike French J, their Honours did not rely (or indeed even comment) on whether ‘situational’ factors can establish special disadvantage.

D. The High Court’s Verdict

The ACCC appealed against the Full Court’s findings that the Roberts were not specially disadvantaged and that the Landlord’s insistence upon legal proceedings being discontinued was not a commercially irrelevant condition of the proposed lease. Ultimately, however, the High Court upheld the Full Court’s decision in favour of the Landlord by a 4:1 majority (Kirby J dissenting). Although the majority issued no joint judgment, certain common threads in their respective judgments are identifiable.

The majority agreed with the Full Court that the Roberts were not specially disadvantaged, however the High Court majority’s reasoning was somewhat different. Emphasising the difference between a ‘special disadvantage or disability’ and a ‘hard bargain’, the majority held that the Roberts were at a distinct commercial disadvantage, but that this disadvantage was not ‘special’ because it did not impair their ability to make judgments in their own interests. The majority thus adopted the conventional position that mere inequality of bargaining power is not, in and of itself, sufficient to establish special disadvantage.

---

36 Id at 43,191.
37 Id at 43,192.
38 Ibid.
39 Ibid.
40 Australian Competition and Consumer Commission v CG Berbatis Holdings Pty Ltd (2003) 197 ALR 153 at 168 (Gummow and Hayne JJ) (hereinafter Berbatis (No 4)).
While the majority agreed with the Full Court’s finding that the Roberts had not been specially disadvantaged, the High Court justices voiced different opinions as to whether the Full Court erred in construing special disadvantage as requiring the weaker party’s will to have been so overborne that they acted neither independently nor voluntarily. Gummow and Hayne JJ, together with the otherwise dissenting Kirby J, held that the Full Court erred in construing s51AA against that criterion, noting that questions of quality of consent are relevant to duress, but not to the narrow doctrine of unconscionability. Gleeson CJ and Callinan J took a different view and sought to clarify what the Full Court must have meant. Gleeson CJ regarded the Full Court’s remark as ‘simply an observation of fact as to part of the context in which the issue of unconscionability arose.’ In Callinan J’s opinion, the Full Court had actually meant that ‘whether a person’s will had in fact been overborne was a relevant, but not necessarily an essential element in many cases of unconscionability’.

The majority of the High Court agreed with the Full Court that the Landlord did not take unconscientious advantage of the Roberts. Given there was no option to renew, the majority found that the Landlord was legally entitled to make discontinuance of legal proceedings a condition of any new or extended lease. As Gleeson CJ highlighted, ‘[p]arties to commercial negotiations frequently use their bargaining power to “extract” concessions from other parties’, this being ‘the stuff of ordinary commercial dealing.’ A relevant, but not determinative, factor was that the Roberts would have suffered a far greater financial loss had they not renegotiated the lease for their business premises.

The lone dissenting voice in the High Court was Kirby J. His Honour interpreted s51AA broadly, finding that s51AA has the potential to extend beyond the narrow doctrine of unconscionability, as articulated in Amadio. Although noting the appeal before him was substantially argued by reference to the narrow doctrine of unconscionability, Kirby J maintained that ‘the reach of the section … goes further’, with its full scope still ‘to be elaborated in this and future cases’. Given the facts in Berbatis, Kirby J concluded it was open to French J to regard the Roberts as specially disadvantaged given their lease had been due to expire, they had pressing reasons to sell their business, and the business was connected to

41 National Westminster Bank plc v Morgan [1985] 1 All ER 821 at 830 (Lord Scarman); Amadio, above n4 at 462 (Mason J); Louth v Diprose (1992) 175 CLR 621 at 654.
42 Berbatis (No 4), above n40 at 162.
43 Id at 174.
44 Gummow and Hayne JJ commented that what ‘was said by the Full Court reflects notions associated with common law duress and the defence of non est factum rather than unconscionable conduct’: Berbatis (No 4), above n40 at 162. Kirby J similarly noted that ‘the question of whether the will of the party was overborne, so that it cannot be said that that party acted voluntarily, is a consideration relevant to the doctrine of common law duress’: Berbatis (No 4), above n40 at174.
45 Berbatis (No 4), above n40 at 158.
46 Id at 196.
47 Id at 166–167 (Gummow and Hayne JJ).
48 Id at 158.
49 Id at 174.
the value of their lease renewal or extension of the lease. Additionally, Kirby J concluded it was open to French J to find that the Landlord had exploited the Roberts’ special disadvantage.

4. Murky Waters: Implications of the Berbatis Litigation

A. The Scope of Unconscionability under s51AA

Following the handing down of the High Court’s judgment in Berbatis, the ACCC stated that ‘although the High Court appears to adopt a restrictive interpretation, the decision has helped to clarify the application of s 51AA of the [TPA].’ However, close analysis of the High Court’s decision in fact reveals that many questions in respect of s51AA’s scope remain unresolved.

(i) The Broad Principle of Unconscionability

In considering the implications of Berbatis, it is important to note that the s51AA claim brought by the ACCC, as well as the subsequent appeals, were based on the narrow doctrine of unconscionability. A passage from Gummow and Hayne JJ’s joint judgment highlights this. Their Honours note that although s51AA has been judicially considered both to extend and not to extend to the broad principle of unconscionability, they conclude that the appeal can ‘be decided without choosing between the differing emphases in the[se] views,’ given that the facts of Berbatis fall within the narrow doctrine of unconscionability. As a result, the High Court was not obliged in Berbatis to consider whether s51AA also encompasses the broad principle of unconscionability, thereby embracing equitable doctrines other than the narrow doctrine of unconscionability. Given this, it is interesting to briefly refer to the Full Federal Court’s recent observations on this issue in ACCC v Samton Holdings Pty Ltd.

In Samton Holdings, the Full Court noted that reference in extrinsic materials to the narrow doctrine of unconscionability, as articulated in Blomley v Ryan and Amadio, does not limit s51AA to that doctrine. In the Full Court’s opinion,

52 Australian Competition and Consumer Commission v CG Berbatis Holdings Pty Ltd [2000] FCA 2 (French J).
53 GPG (Australia Trading) Pty Ltd v GIO Australia Holdings Ltd (2001) 117 FCR 23 at 77 (Gyles J).
54 Berbatis (No 4), above n40 at 165.
55 Id at 163.
56 (2002) 189 ALR 76 (hereinafter Samton Holdings). A similar approach had been taken even earlier in Pritchard v Raccavage Pty Ltd (1997) 72 FCR 203 and Olex Focus Pty Ltd v Skodaexport Co Ltd (1997) 142 ALR 527. However, both were interlocutory proceedings only.
57 Samton Holdings, above n56 at 93.
unconscionable conduct under s51AA ‘is that which supports the grant of relief on the principles set out in specific equitable doctrines.’ This reference to ‘specific equitable doctrines’ indicates that the Court did in fact regard s51AA’s language as sufficiently wide to embrace the broad principle of unconscionability. Yet, despite this broad interpretation, the Full Court held that it was not unconscionable under s51AA for a landlord to demand a $70,000 payment from a tenant who accidentally failed to exercise an option to renew the lease on time. Like in Berbatis, it was held that in cases involving parties with commercial experience, mere commercial vulnerability is not enough to constitute ‘special disadvantage or disability’. Thus, while Samton Holdings suggests s51AA is wide enough to embrace the broad principle of unconscionability, it sheds little light on what circumstances will enable that principle to be invoked when s51AA is relied upon.

Ultimately, both Berbatis and Samton Holdings highlight the continuing uncertainty over s51AA’s limits. Certainly, concerns have been voiced in respect of s51AA being judicially interpreted as extending to the broad principle of unconscionability. Tucker, for instance, argues that such an interpretation would be unconstitutional, commenting:

It could hardly be argued that a section that encapsulates all equitable, and a growing proportion of common law, doctrines now and into the future does not confer upon the judiciary a legislative role of enormous import and impact. The offence to Chapter III of the Constitution is obvious.

O’Brien raises another concern. He asserts that an expansive reading of s51AA would blur the distinction between that provision and ss51AB and 51AC, thereby questioning ‘the utility of having three separate provisions in the TPA which regulate unconscionable conduct.’ Judicial opinion in respect of such concerns has yet to be expressed. Noteworthy, however, is that in the event of the High Court interpreting s51AA as embracing the broad principle of unconscionability, there would also be implications for other areas of the law. For example, Riley argues that if s51AA is interpreted as embracing an equitable doctrine like estoppel, s51AA may be able to be used in addition to, or instead of, s52 in employment cases, because such cases often satisfy the elements of equitable estoppel. Moreover, such a development would have implications for cognate

---

58 Id at 92.
59 The ‘specific equitable doctrines’ to which the Court refers are set out in Samton Holdings, above n56 at 92.
60 Samton Holdings, above n56 at 97. Also relevant seems to have been the fact that the lessors were under no legal obligation to negotiate a new lease at all.
62 Tucker, above n1 at 86.
unconscionability provisions under other statutes like s12CA of the *Australian Securities and Investments Commission Act 1989* (Cth). Ultimately, however, without a High Court pronouncement on the matter, it presently remains unclear whether s51AA of the TPA even embraces the broad principle of unconscionability.

(ii) The Narrow Doctrine of Unconscionability

Despite having the opportunity, the High Court did not definitively establish whether special disadvantage for the purposes of the narrow doctrine of unconscionability can be based on ‘situational’ as well as ‘constitutional’ factors. Gleeson CJ was the only High Court justice to expressly address the notion of ‘situational’ disadvantage. His Honour did not entirely rule out such an expansion of the narrow doctrine of unconscionability, noting that ‘emotional dependence of the kind illustrated by *Louth v Diprose* as a form of special disadvantage described as “situational” rather than “constitutional” is understandable and acceptable.’ Gleeson CJ did, however, add a significant caveat, namely that terms like ‘situational disadvantage’ and ‘constitutional disadvantage’ should not take on a life of their own and become substitutes for the language and content of the statute.

It is interesting to note that prior to *Berbatis*, the Full Federal Court considered the notion of ‘situational’ disadvantage in *Samton Holdings*. While the Full Court clearly established that mere differences in bargaining power between commercially experienced parties are insufficient, the Court did acknowledge that the categories of ‘special disadvantage’ may extend beyond ‘constitutional disabilities’ to also embrace cases involving ‘situational disadvantage’. In light of *Samton Holdings*, and the High Court’s failure in *Berbatis* to rule one way or the other in respect of the notion of ‘situational’ disadvantage, it is likely this issue will come before the courts again in the future.

What the High Court’s decision in *Berbatis* does clarify is that mere inequality of bargaining power is insufficient to establish special disadvantage for the purposes of the narrow doctrine of unconscionability in s51AA. As Gleeson CJ emphasised:

> A person is not in a position of relevant disadvantage, constitutional, situational or otherwise, simply because of inequality of bargaining power. Many, perhaps even most, contracts are made between parties of unequal bargaining power, and good conscience does not require parties to contractual negotiations to forfeit their advantages, or neglect their own interests … Unconscientious exploitation of another’s inability, or diminished ability, to conserve his or her own interests is not to be confused with taking advantage of a superior bargaining position. There may be cases where both elements are involved but, in such cases, it is the first, not the second, element that is of legal consequence.

---

65 Horrigan, above n15.
66 *Berbatis (No 4)*, above n40 at 156.
67 Ibid.
69 *Berbatis (No 4)*, above n40 at 157.
Even the otherwise dissenting Kirby J agreed that an inequality of bargaining power, or the striking of a hard bargain, is not, in itself, sufficient to justify concluding that the stronger party acted unconscionably. Such comments clearly reflect the ongoing importance of contractual considerations and the well-known maxim ‘Chancery mends no man’s bargain’. However, while *Berbatis* may reiterate that driving a hard bargain does not, in and of itself, place a party at a special disadvantage so as to attract s51AA’s protection, it remains uncertain what additional factors – in particular, whether situational, as well as constitutional, factors – are required to turn a hard bargain into unconscionable conduct.

(iii) Section 51AC Unconscionability – Would its Application Have Led to a Different Outcome in *Berbatis*?

Like s51AA, s51AC applies to commercial transactions. However, aside from this, the sections differ significantly. Section 51AC was introduced only in 1998, following extensive lobbying and a number of reports which concluded that the TPA was still largely ineffective at remedying unconscionable conduct in commercial transactions involving small business. Under s51AC(1), a corporation must not, in trade or commerce, engage in unconscionable conduct in connection with the supply or acquisition, or possible supply or acquisition, of goods or services to a person (other than a publicly listed company). The opposite is also prohibited under s51AC(2), which stipulates that a person must not, in trade or commerce, engage in unconscionable conduct in connection with the supply to or acquisition from, or possible supply to or acquisition from, a corporation (other than a publicly listed company) of goods or services. There are, however, constraints on the size of businesses to whom s51AC applies, the business cannot be a publicly listed company, and s51AC applies only in respect of goods or services worth up to $3 million.

70 Id at 176.
71 *Maynard v Moseley* (1676) 3 Swabs 651 at 655 (Lord Nottingham LC).
74 Underscoring s51AC’s advent was thus a desire to afford small business better recourse to unconscionability actions than was available in equity and under s51AA: Second Reading Speech of the Trade Practices Amendment (Fair Trading) Bill 1997 (Cth), Commonwealth of Australia, House of Representatives, *Parliamentary Debates (Hansard)*, 30 September 1997 at 8801; Explanatory Memorandum accompanying the Trade Practices Amendment (Fair Trading) Bill 1997 (Cth).
75 TPA s51AC(1).
76 TPA s51AC(2).
77 TPA s51AC(9) and (10). Until the *Trade Practices Amendment Act (No 1) 2001* (Cth) this figure was only $1 million.
A principal difference between s51AC and s51AA is that s51AC(3) provides a list of indicia to which a court may have regard when determining whether s51AC unconscionable conduct has occurred. Beale maintains that these factors indicate s51AC ‘is intended to include conduct already covered by equitable doctrines and, in addition, to ‘extend’ to other conduct that is, in all the circumstances, unconscionable.’78 Indeed, s51AC even appears to encompass ‘situational’ disadvantage, at least insofar as allowed under the s51AC(3) indicia, which include considerations such as inequality of bargaining power, unfair tactics, compliance with industry codes, and the good faith of parties. These indicia clearly signal that s51AC escapes the constraints of s51AA and equity’s narrow doctrine of unconscionability. Consequently, whereas s51AA unconscionability principally concerns the means by which parties transact, s51AC unconscionability examines both the means of transacting and the substance of a transaction.

Section 51AC’s parameters received significant attention in a recent case involving a franchisor/franchisee relationship.79 In Simply No-Knead, Sundberg J referred to a Full Federal Court decision in which it was noted both that ss51AB and 51AC unconscionable conduct require evidence of ‘serious misconduct or something clearly unfair or unreasonable’ and that the ‘various synonyms used in relation to the term ‘unconscionable’ import a pejorative moral judgment.’80 Sundberg J then himself commented:

… in my view ‘unconscionable’ in s 51AC is not limited to the cases of equitable or unwritten law unconscionability the subject of s 51AA. The principal pointer to an enlarged notion of unconscionability in s 51AC lies in the factors to which subsection (3) permits the Court to have regard. Some of them describe conduct that goes beyond what would constitute unconscionability in equity.81

Simply No-Knead thus confirms s51AC is far broader than s51AA. The result, as noted by the ACCC Commissioner, is that it does not appear necessary ‘for a person wanting to establish a contravention of … [s]51AC to show that the weaker party took unfair advantage of that disadvantage (which is the requirement for unconscionable conduct in equity, or unwritten law).’82 Wilson takes this observation even further, commenting:

---


81 Simply No-Knead, above n79 at 315.

… there is no doubt that the introduction of s 51AC was intended to unleash a less restrictive concept of unconscionability where taking unconscientious advantage will not require a special disadvantage on the part of tenants, but will open up questions about integrity, bad faith, unfair bargaining based on disparities of size and resources, the pursuit of legitimate business interests and the general idea of fairness in transactions between supplier corporations and business consumers.83

In other recent cases, further guidance on s51AC’s scope has been provided. For instance, in the trial decision of *Berbatis*, French J asserted that the s51AC(3) indicia are not exhaustive, meaning the categories of unconscionable conduct falling under s51AC ‘will never be closed albeit the circumstances of the application of the standard prescribed … is confined by the language’ of s51AA.84 Section 51AC was thus regarded as of wider import than s51AA, allowing the court to consider undue influence, duress and/or other issues falling outside s51AA’s relatively narrow scope.85 Interestingly, however, neither the Full Federal Court nor the High Court made any comment on s51AC in the *Berbatis* litigation, presumably because it was not in fact a point directly in issue.

In a different case, *ACCC v 4WD Systems Pty Ltd & Ors*,86 Selway J implied that s51AC is not limited to the narrow doctrine of unconscionability by suggesting that while Kirby J’s dissenting approach in *Berbatis* does not represent the law on s51AA, it may be relevant to s51AC’s interpretation.87 Selway J concurred with the opinions of French J in *Berbatis* and Sundberg J in *Simply No-Knead*, namely that s51AC unconscionability is not limited to the meaning of unconscionability at common law or in equity.88 Moreover, Selway J noted that to find s51AC unconscionability, more is required than simply establishing that ‘behaviour is misleading or deceptive, or otherwise in breach of some other provision of the TPA.’89 Instead, it must be shown ‘that the conduct is so unacceptable that it can properly be described as ‘unconscionable’.’90

Despite these judicial reflections on s51AC, its full reach has yet to be definitively established. The Fair Trading Coalition points to the fact that ‘there has not been a substantial amount of successful litigation on s51AC’ and that actions which have been successful involved only ‘the most egregious behaviour.’91 Gardini maintains that the ACCC’s lack of legal action under s51AC is attributable to s51AC falling well short of the ‘unfairness’ concept that the 1997 Reid Report recommended be incorporated into the TPA instead of

---

83 Wilson, above n30, 81. Webb similarly contends that the list of factors in s51AC will encourage the courts to adopt an expanded concept of unconscionability: Eileen Webb, ‘Unconscionable Conduct and Retail Shop Wars – The Lessees Strike Back’ (2000) 8 APLJ 67.
84 Ibid.
85 Id at 335–336.
86 (2003) 200 ALR 491 (hereinafter, *4WD Systems*)
87 Id at 543 (Selway J).
88 Ibid.
89 Id at 544 (Selway J).
90 Ibid.
‘unconscionability’. Moreover, Gardini argues the threshold test for unconscionability is too difficult to establish, except in cases involving the most blatant forms of unconscionable conduct. In any case, what is clear is that s51AC’s ambit has not yet been clearly delineated. What Simply No-Knead and 4WD Systems do indicate is that the s51AC(3) indicia render s51AC wider than s51AA. In light of this, there has emerged the suggestion in relation to Berbatis that ‘if similar conduct was to occur again and be assessed against s51AC, the result may well be different.’

Given the lack of clarity as to s51AC’s full reach, consideration of whether Berbatis would have been decided differently had s51AC been in force at the relevant time remains speculative only. Nevertheless, looking at the facts in Berbatis and the considerations to which a court may have regard in s51AC(3), there are strong signals that the Roberts’ would have been more successful in gaining a TPA remedy under s51AC than they were under s51AA. Certainly, s51AC prima facie applies to their case given it is directed at commercial dealings of small business. Turning to consider s51AC(3), the facts also appear to prima facie satisfy various indicia. For instance:

i) Section 51AC(3)(a) allows a court to consider inequality of bargaining power. An inequality of bargaining power certainly existed in Berbatis. However, it remains unclear whether it is the type of inequality of bargaining power envisaged by s51AC(3)(a).

ii) Section 51AC(3)(b) directs the court to consider whether, as a result of conduct engaged in by the supplier, the business consumer was required to comply with conditions that were not reasonably necessary for the protection of the supplier’s legitimate interests. It seems at least questionable whether making the extension of a lease conditional upon discontinuance of legal proceedings is reasonable and in the supplier’s legitimate interests. Indeed, Webb maintains that the Landlord in Berbatis would ‘likely … have fallen foul of this provision’ because ‘[s]urely it is not reasonable for a party to prevent the valid exercise of a tenant’s legal rights through the threat of the lease not being extended’.

---
92 Robert Gardini, ‘The Dawson Review – A Small Business Perspective’ (2003) 9(1) UNSWLJ Forum. By this Gardini seems to be implying that ‘unconscionable conduct’ in s51AC is narrower than the ‘unfairness’ concept canvassed in the Reid Report. The main difference between the two appears to be the indicia to which a court may have regard. In this regard, the ‘unfairness’ provision recommended in the Reid Report appears potentially broader, as it includes an additional factor for consideration, namely ‘the harshness of the result’: Reid Report, above n73, Recommendation 6.1.

93 Gardini, above n92.

94 Mullins & Mullins, ‘Business Update: Unconscionable Conduct on the Rise’ at 2: <http://www.mullins-mullins.com.au/files/HP275.PDF> (10 October 2003). Indeed, French J himself in a ‘Summary’ that preceded (but did not form part of) his judgment commented: ‘It may be that a different result could have been obtained under the later and wider provisions of s 51AC. That question will have to await another day and another case’: Australian Competition and Consumer Commission v CB Berbatis Holdings Pty Ltd (ACN 008 799 040) (2000) ATPR 41–778.

95 Webb, above n30 at 109.
iii) Section 51AC(3)(d) enables the court to consider whether any undue influence or pressure was exerted on, or any unfair tactics were used by, the supplier against the business consumer. With the facts of Berbatis in mind, Webb argues that ‘a threat to refuse to extend a tenant’s lease, and therefore lose their business, unless legal action was halted would appear to fit within this section.’

iv) Section 51AC(3)(k) allows the court to examine the good faith of both parties. Arguably, making the renewal of the Roberts’ lease contingent on discontinuance of legal proceedings against the Landlord may be interpreted as a condition not negotiated in good faith. This is particularly so given the Roberts’ primary reason for needing to sell their business was their daughter’s illness, a factor the Landlord was aware of, and the condition was clearly non-negotiable. However, while good faith considerations have been applied in Australia previously in relation to the exercise of contractual powers, whether the Roberts’ would have succeeded on this basis largely depends on how ‘good faith’ in s51AC(3)(k) is judicially defined, a fact which remains unclear.

The s51AC indicia clearly reflect the two central concerns that Parkinson identifies as being at ‘the heart of all the different applications of the conscience of equity’, that is, protection of the vulnerable and protection of people’s reasonable expectations. There thus seems to be strong reason to conclude that the Roberts’ may have had greater success if their claim had been brought under s51AC. Ultimately, however, this depends on what sort of conduct the courts will regard as sufficient to satisfy the indicia listed in s51AC.

5. Conclusion

As s51AA’s reach has been unclear since its inception, there were high hopes that in Berbatis the High Court would firmly delineate the parameters of that provision. The High Court’s decision has arguably made it harder for landlord-tenant conduct to be characterised as unconscionable under s51AC by emphasising that inequality of bargaining power is not, in and of itself, enough to constitute unconscionable conduct. However, the High Court did not directly reject the idea that s51AA unconscionable conduct can arise between commercial parties as a result of ‘situational’ disadvantage. Similarly, the High Court did not altogether reject the proposition that s51AA has the potential to embrace the broad principle of unconscionability. Exploitative conduct between unequal commercial parties, including extraction of irrelevant commercial concessions, might be enough to ground a finding of unconscionability, yet as Horrigan points out, it is ‘[o]f course,

96 Id at 110.
97 Good faith has been referred to as an obligation that ‘underwrites the spirit of the contract and supports the integrity of its character’: Overlook v Foxtel [2002] NSWSC 17 at [67] (Barrett J).
recognizing and drawing that line in practice [that] is the key.\(^9^9\) It is precisely where this line should be drawn that the High Court’s decision in *Berbatis* leaves at large, thus the only firm guidance offered by the Court is that an inequality of bargaining power between commercial parties is not, in and of itself, enough to establish s51AA unconscionability.

Horrigan maintains that the judicial and legislative development of ss51AA and 51AC mirrors the TPA’s ‘transformation from an Act primarily regulating anti-competitive conduct and abuse of market power to one which equally regulates commercially unfair, self-interested, and opportunistic conduct, regardless of its impact on competition and markets.’\(^1^0^0\) There is undeniably an element of truth in this comment, and it is true that there is now required a greater level of good conscience in business and other relations than ever before. Yet just how great this move has been remains unclear, even after *Berbatis*. Horrigan is, however, correct, in concluding that the ‘gap between sharp conduct and unconscionable conduct has narrowed’\(^1^0^1\) as a result of s51AC’s introduction and judicial interpretation of both ss51AA and 51AC. Only future litigation will reveal just how narrow this gap has in fact become.

---

\(^9^9\) Horrigan, above n15.

\(^1^0^0\) Ibid.

\(^1^0^1\) Ibid.
Renvoi Rejected? The Meaning of ‘the lex loci delicti’ after Zhang

ROBERT YEZERSKI*

Abstract

This article considers the scope of the doctrine of renvoi in private international law, and specifically, whether it has any operation in the field of tort. It has generally been assumed that the common law ‘rejects’ renvoi in tort and that where the forum’s choice of law rules require the application of foreign law, only the foreign legal system’s ‘internal’ laws are applied, absent any consideration of the foreign choice of law rules. However, it is argued that the merits of rejecting renvoi in tort can only be assessed by reference to the relevant choice of law rule. For this reason, and in light of the High Court’s recent adoption of the lex loci delicti rule for international torts in Zhang, it is submitted that renvoi should be given operation in tort because it is consistent with the policies underlying the acceptance of that choice of law rule. The analysis of this question proceeds upon a consideration of the recent decision of the Full Court of the Supreme Court of Western Australia in Mercantile Mutual Insurance (Australia) Ltd v Neilson in which the Court adhered to the traditional view that renvoi is rejected in tort.

* BA(Hons) LLB(Hons) (University of Sydney). The author wishes to thank Ross Anderson for his valuable comments on the draft of this paper. Thanks also to Ben Doyle, James Emmett & Sarah Knuckey. All opinions, and any errors, remain my own.
1. Introduction

Since the High Court’s decision in *Regie Nationale des Renault Usines SA v Zhang* *(Zhang)*,1 the Australian common law has provided a single choice of law rule for international torts. In that case, the High Court extended the choice of law rule for intranational torts from *John Pfeiffer Pty Ltd v Rogerson* *(Pfeiffer)*2 by establishing that the lex loci delicti is the applicable system of law for all foreign tort claims. What *Zhang* left unresolved, however, was the question of what is meant by the ‘lex loci delicti’ — is it simply the domestic or ‘internal’ law of a foreign country, or does that phrase incorporate that country’s choice of law rules as well?3 Those rules might conflict with choice of law rules of the forum by requiring the application of either the lex fori (remission) or the laws of a third state (transmission).4 This is the problem of renvoi, which arises where there is a conflict between conflicts of law rules, thereby creating the potential for a perpetual series of references from one system of law to another.

This was precisely the problem that was faced by the Full Court of the Supreme Court of Western Australia (McLure and Johnson JJ, and Wallwork AJ) in *Mercantile Mutual Insurance (Australia) Ltd v Neilson* *(Neilson)*.5 That case was an appeal from the judgment of McKechnie J at first instance.6 The genesis of the litigation was an action in negligence commenced by an Australian national and Western Australian resident, Mrs Barbara Neilson, against a corporation incorporated in, and owned by, the state of Victoria, Overseas Projects Corporation of Victoria Ltd (OPC). Mrs Neilson alleged that she had suffered tortious injury while living in China as a result of the negligent omission of OPC. The lex loci delicti was the law of China; however Chinese law included a provision allowing a claim for damages between two foreign nationals of common nationality to be resolved according to the law of that nationality. In other words, although *Zhang* required that Chinese law be applied (the lex loci delicti), the lex loci delicti here arguably required, or at least permitted, the application of Australian law.

2. The Facts in Neilson

The facts in *Neilson* were straightforward. The plaintiff, Mrs Neilson, accompanied her husband to China in 1991 when he went to teach at Wuhan Iron and Steel University in the city of Wuhan, in Hubei Province, central China. Her husband, Mr George Neilson, was employed by OPC.7

---

5 [2004] WASCA 60 (‘Appeal decision’).
6 *Neilson v Overseas Projects Corporation of Victoria Ltd* [2002] WASC 231 (hereinafter ‘Trial decision’).
7 Trial decision at [1].
Under Mr Neilson’s employment contract, the Neilsons were to move to China for a period of up to two years. The contract expressly provided that Mrs Neilson would accompany her husband to Wuhan and that OPC would supply the couple with an apartment to live in while in China. Mrs Neilson was also employed by OPC. However, her employment was not pursuant to a written contract and did not include any terms relating to accommodation.

In October 1991, while in Wuhan, Mrs Neilson woke in the middle of the night and tried to make her way to the downstairs section of the apartment. While attempting to feel her way in the dark, Mrs Neilson fell down the stairs, suffering head and back injuries. At the date of the trial, she suffered ongoing back and neck pain, as well as headaches, tinnitus and vertigo.

3. The Trial Decision

Mrs Neilson commenced proceedings against OPC in the Supreme Court of Western Australia claiming damages in tort and contract on the basis that her injuries resulted from OPC’s failure to provide safe accommodation. The claim in contract was dismissed on the basis that her contract with OPC did not contain any promise to provide her with accommodation in Wuhan and because she was not entitled to sue under her husband’s contract because of considerations of privity. For present purposes, only the action in tort is relevant.

In relation to the tort grounds, Mrs Neilson alleged that the failure of OPC to install a balustrade alongside the stairs in the apartment was the direct cause of her injuries. The evidence established that the Neilsons had complained to OPC’s representatives about the lack of a balustrade several times prior to the accident and that OPC had taken no steps to rectify the problem.

At trial, McKechnie J applied Distillers Co (Biochemicals) Ltd v Thompson and Voth v Manildra Flour Mills Pty Ltd and concluded that the locus delicti was China. His Honour held that he was therefore bound to apply Chinese law following the lex loci delicti rule adopted in Zhang.

Turning to the expert evidence on Chinese law, McKechnie J found that the applicable Chinese legislation was the General Principles of Civil Law of the People’s Republic of China (hereinafter ‘the General Principles’), adopted at the Fourth Conference of the Third National People’s Congress on 12 April 1986. Article 146 of the General Principles provides:

---
8 Id at [18].
9 Id at [14].
10 Id at [20]–[21], [72]–[80].
11 Id at [41].
12 Ibid.
13 Id at [47].
14 Id at [72]–[94].
16 (1990) 171 CLR 538.
17 Trial decision at [122].
18 Ibid.
With regard to compensation for damages resulting from an infringement of rights, the law of the place in which the infringement occurred shall be applied. If both parties are nationals of the same country or domiciled in the same country, the law of their own country or of their place of domicile may also be applied …

(emphasis added)

The expert evidence before the Court was that Article 146 was not relevant to the proceedings because it was a conflict of laws provision. However, the expert was unable to clearly explain why this would mandate its non-application in the present proceedings. McKechnie J therefore held that in order to ‘do justice and fairness’, his Honour would apply the Article. The effect was that McKechnie J resolved the dispute according to Australian tort law principles and eventually held OPC liable in negligence for Mrs Neilson’s injuries.

McKechnie J’s decision to apply Article 146 proceeded on the assumption that the provision was part of the lex loci delicti — the system of law that he was bound to apply following Zhang. In one sense, this assumption is obviously correct: Article 146 is a provision of Chinese law and Chinese law was the lex loci delicti. However, a reference to the law of a foreign country is usually interpreted more narrowly to mean the law which a court of that country would apply in a case involving no foreign element. In other words, a reference to the law of a foreign country is often treated as referring only to that country’s substantive or ‘internal’ laws, excluding its choice of law rules. This is because of the problem of renvoi, which arises where there is a conflict of choice of law rules between the law of the forum and the foreign system of law to which the forum court is referred. Were we to apply the foreign state’s choice of law rules, those rules may require the application of Australian law, at which point Australia’s own choice of law rules would once more require the application of foreign law. Thus, taking the Neilson case as an example, Australian law requires the application of the lex loci delicti, but Chinese law requires (or at least permits) that the law of the nationality be

19 Id at [127]. It appears that McKechnie J may have erred in his citation of the General Principles of Civil Law of the People’s Republic of China. The General Principles were adopted by the Fourth Session of the Sixth National People’s Congress on 12 April 1986, as opposed to McKechnie J’s finding that that legislation was adopted by the Third National People’s Congress: Kui-Hua Wang and Danuta Mendleson, ‘An Overview of Liability and Compensation for Personal Injury in China Under the General Principles of Civil Law’ (1996) 4 TLJ 137 at 146.
20 Trial decision at [200].
22 Trial decision at [202].
23 Id at [203]–[204].
25 Ibid.
26 See generally, id at 65–80; Nygh & Davies, above n4 at 289–302.
applied. The cycle does not stop there, however, because if the ‘law of the nationality’ incorporates Australian choice of law rules, we are once again sent back to Chinese law and the cycle continues ad infinitum. McKechnie J did not advert to this problem in Neilson, choosing instead to take comfort in the generous homeward gesturing of Chinese law, without recognising that, upon returning to Australian law, his Honour would arguably be sent straight back to the law of China.

Before moving further, it is convenient to say something about the problem involved in using the phrase ‘Australian law’ in the field of tort. In the trial decision, McKechnie J stated:

I propose to apply Article 146. It gives me a right to choose to apply the law of Australia because both parties are nationals of Australia.

The phrase ‘the law of Australia’ requires some further explanation as it applies in relation to tort. While Australia has a single common law, the operation of statute has the effect that the law of tort is governed by a different regime in each state and territory.27 Cases such as Breavington v Godleman,28 Anderson v Eric Anderson Radio and TV Pty Ltd29 and Pfeiffer v Rogerson30 all stand as evidence of the significant differences that exist between Australian jurisdictions in matters of tort. The extent of these differences has increased more recently with several states passing tort reform legislation.31 No system of law can therefore answer the description ‘the law of Australia’ in relation to tort. As such, there is an incongruence between the law to which Article 146 refers (‘the law of the nationality’) and the law which governs the relevant subject matter (tort) in Australia. Article 146 therefore appears to refer the question to a system of law that can offer no answer.

Chinese law is apparently not without a solution to this problem. According to the American scholar Mo Zhang, Chinese law provides that where there is no unitary national law relating to a particular subject matter in the foreign country:

[T]he applicable law shall be determined according to the conflict of law rules of the country. If the foreign country has no such rules, the law of the place that has the closest relationship with the civil relations involved shall be applied directly.32

27 Pfeiffer, above n2 at 514–515; Compare Re Benko [1968] SASR 243.
29 (1965) 114 CLR 20.
30 Pfeiffer, above n2.
31 See Civil Liability Act 2002 (NSW); Civil Liability Act 2003 (Qld); Civil Liability Act 2002 (WA); Civil Liability Act 2002 (Tas); Wrongs (liability and Damages for Personal Injury) Amendment Act 2002 (SA); Civil Law (Wrongs) Act 2002 (ACT); Personal Injuries (Liabilities and Damages) Act 2003 (NT).
In the field of succession, the common law has solved similar problems by adopting whatever solution the relevant foreign legal system would apply when the national law to which it is referred lacks a unitary system of law with respect to that subject matter.\(^{33}\) Thus, consistent with the common law approach and the Chinese rule described by Mo Zhang, it would seem that the applicable law in \textit{Nielsen} was that with which the ‘civil relations involved’ had the ‘closest relationship’. It can therefore be assumed that the applicable law was properly identified as either the law of Western Australia or the law of Victoria. Neither body of law can be described as ‘the law of Australia’ either in general terms or as it applies to the field of tort.

4. \textit{The Appeal Decision}

Appeal proceedings were commenced in the Full Court of Western Australia. The appellant, Mercantile Mutual Insurance (Australia) Ltd (hereinafter ‘Mercantile’), was OPC’s third party liability insurer. There were three named respondents to the appeal being Mr and Mrs Neilson, OPC and OPC’s insurance broker.

Mercantile pressed three grounds of appeal. First, it argued that the doctrine of renvoi does not apply in tort and that McKechnie J therefore erred in applying Article 146. Secondly, it contended that Mrs Neilson’s claim was statute-barred by operation of a Chinese limitation law. Finally, it sought to establish that Mrs Neilson’s injuries arose out of her course of employment with OPC (the result of which would have been that Mercantile was not liable to indemnify OPC in respect of Mrs Neilson’s injuries).\(^{34}\) The Full Court allowed the appeal on the first two grounds, while dismissing the third ground. For present purposes, only the reasoning in relation to the first ground is relevant.

McLure J, with whom Johnson J and Wallwork AJ agreed,\(^{35}\) commenced her consideration of the renvoi issue by providing an overview of the three possible solutions that a legal system can adopt when faced with a renvoi problem: ‘rejecting’ the renvoi (‘the internal law solution’), the ‘single renvoi solution’ and the ‘double renvoi’ solution.\(^{36}\)

The first of these solutions, rejecting the renvoi, is the most straightforward approach. This solution simply requires a court to ignore (or reject) the operation of the foreign legal system’s choice of law rules and instead apply the ‘internal’ or substantive law of that foreign state.\(^{37}\) This is the approach which the common law probably adopts in matters of contract, largely because it is believed that party

---


\(^{34}\) It was for this reason that the third respondent, OPC’s insurance broker, had been joined as a party. In the event that Mercantile established that Mrs Neilson’s injuries were sustained in the course of her employment, OPC claimed against the third respondent: Appeal decision at [7].

\(^{35}\) Id at [75], [76].

\(^{36}\) Id at [29]–[32].

\(^{37}\) Id at [29]; See also Sir Peter North & J J Fawcett, \textit{Cheshire and North’s Private International Law} (13\textsuperscript{th} ed, 1999) (hereinafter ‘Cheshire and North’) at 53; Nygh & Davies, above n4 at 290; \textit{Dicey and Morris}, above n24 at 66.
expectations will be defeated if the substantive law of the proper law of the contract is not applied. In the context of Neilson, the internal law solution would require an Australian court to disregard the operation of Article 146 and instead apply only the substantive provisions of the General Principles.

The second solution, and the one which seems to have been adopted by McKechnie J at first instance, is the ‘single renvoi’ approach (also known as ‘accepting the renvoi’). This is the approach adopted by most civil law systems. The single renvoi solution requires that, when faced with a renvoi problem, a court will only follow the first reference back and then apply the internal law of the country to which that first reference refers. The operation of the single renvoi approach is aptly demonstrated by McKechnie J’s approach at first instance in Neilson. The rule from Zhang referred the Court to Chinese law; however, Article 146 of the General Principles referred the matter back to Australian law. Rather than following Zhang back to Chinese law once more (thereby initiating a seemingly perpetual cycle of references between legal systems), McKechnie J simply ‘accepted’ the reference back from Chinese law and applied the ‘internal law’ of Australia.

The third solution, and the one more commonly adopted in Australian law, is the so called ‘double renvoi’ approach or ‘foreign court theory’. This approach has been applied by common law courts in relation to the formal and intrinsic validity of wills, the intestate succession to moveable property and the recognition of foreign legitimations by subsequent marriage. It may also have application in relation to questions concerning the formal validity of a marriage.

---
39 McLure J, writing in the Full Court, expressed some uncertainty as to whether McKechnie J’s approach was the ‘single renvoi’ solution or the ‘double renvoi’ solution (described below): Appeal decision at [32]. It could be argued that McKechnie J’s approach was a form of double renvoi solution whereby his Honour followed the reference to Chinese law, recognised that Chinese law required the application of Australian internal law, and simply applied the latter. Such an approach would essentially conflate the inquiry as to the relevant Chinese choice of law rule with the inquiry as to the Chinese renvoi rule. As the relevant Chinese choice of law rule bears some similarity to a proper law of the tort approach, McKechnie J may have concluded that the rule incorporated a rejection of renvoi. That said, this account of his Honour’s reasoning seems unsupported by the text of the judgment as McKechnie J made no mention of the renvoi difficulty and did not seek to identify the nature or policy basis of the Chinese choice of law rule. In this context, the single renvoi solution seems to accord far better with the trial judge’s approach.
40 Dicey and Morris, above n24 at 67.
41 Nygh & Davies, above n4 at 291.
42 Compare Appeal decision at [31].
43 In the Estate of Fuld, deceased (No 3) [1968] P 675.
45 Simmons v Simmons (1917) 17 SR (NSW) 419.
46 Re Askew [1930] 2 Ch 259.
and the capacity of individuals to marry. The double renvoi approach requires a forum court to place itself in the position of the foreign court by adopting whatever renvoi solution the foreign court would apply when faced with the renvoi problem. In other words, the double renvoi solution requires the forum court to resolve the renvoi difficulty in the same manner as the courts of the relevant foreign legal system would. For example, if McKechnie J had applied the foreign court theory in Neilson, his Honour would have investigated what a Chinese court would do when, in applying Australian law under Article 146, it came across the rule from Zhang requiring the application of Chinese law.

Of these three possible solutions to the renvoi problem, the Full Court determined that Australian law adopts the internal law solution in the field of tort. This rejection of renvoi proceeded in two steps: first, McLure J surveyed the precedent and commentary relating to the issue of renvoi and held that the internal law solution was the only approach supported by the authorities. In support of this conclusion, McLure J cited the Scottish case of M’Elroy v M’Allister, and two United States cases, Haumschild v Continental Casualty Co and Pfau v Trent Aluminium Co. Her Honour further noted that the United Kingdom Parliament has recently legislated to the same effect and that the bulk of academic literature also supported the rejection of renvoi in its application to tort.

McLure J summarised the state of the authorities as follows:

[T]here is no binding (or any other) authority that renvoi is applicable in torts cases; there is non-binding authority to the contrary; there is widespread academic opinion that renvoi is not applicable to tort; finally, there is much academic

---

48 Nygh & Davies, above n4 at 291–292; See generally Dicey and Morris, above n24 at 67; Cheshire and North, above n37 at 55–56.
49 Appeal decision at [35]–[40].
50 Id at [35].
51 (1949) SLT 139; Morse cites the English case of Church of Scientology of California v Commissioner of Police for the Metropolis (1976) 120 Sol. J. 690 as support for this proposition although he notes that the rejection of renvoi only appears in the full transcript of the case: C G J Morse, Torts in Private International Law (1978) at 268–269, fn12.
52 95 NW 2d 814 (Wis, 1959).
53 55 NJ 511 (NJ 1970); See also Restatement (Second) of Conflict of Laws, §145 (1971).
54 See Private International Law (Miscellaneous Provisions) Act 1995 (UK) s9(5); See also Worker’s Compensation Act 1987 (NSW), s150E as inserted by the Workers Compensation Legislation Amendment Act 2002 (NSW), Sched 1.
55 Appeal decision at [36]–[39]; See for example Dicey and Morris, above n24 at 73, 1517; Cheshire and North, above n37 at 627; Morse, above n51 at 268–269; Michael Tilbury, Gary Davis & Brian Opeskin, Conflict of Laws in Australia (2002) at 1012. The opposite view, namely that the doctrine of renvoi should not be rejected in tort, does have its advocates. See for example, Erwin N Griswold, ‘Renvoi Revisited’ (1938) 51 Hars LR 1165 at 1205–1207; Inglis, above n33 at 499–503. Adrian Briggs, ‘In Praise and Defence of Renvoi’ (1998) 47 ICLQ 877–884; Ed Rimmel, ‘The Place of Renvoi in Transnational Litigation – A Pragmatic Approach to an Impractical Doctrine’ (1998) 19 Holdsworth LR 55 at 85–89.
criticism of the renvoi doctrine in general.\textsuperscript{56}

The Full Court thus concluded that, in light of the authorities, the only supportable solution to renvoi in tort was to reject the doctrine entirely.

The second step in the Full Court’s rejection of renvoi was to analyse whether the operation of renvoi was incompatible with the reasoning of the High Court in\textit{Zhang} and\textit{Pfeiffer}. To support this contention, McLure J set out the three principal reasons for the High Court’s adoption of the lex loci delicti rule in\textit{Pfeiffer}:

(a) subject to the possible difficulty in locating the tort, liability is fixed by reference to geography and is certain;
(b) the lex loci delicti recognises and gives effect to the predominate concern of local legislatures with acts, matters and things within its territory;
(c) so far as the subject matter permits, the lex loci delicti gives effect to the reasonable expectation of parties.\textsuperscript{57}

Taken together, and in light of the High Court’s rejection of any ‘flexible exception’ to the lex loci delicti rule, McLure J concluded that these three considerations suggest that\textit{Pfeiffer} requires that the lex loci delicti be used to determine the substantive rights of the parties:

The implication in the reasoning of the majority in\textit{Pfeiffer}, particularly in their reasoning relating to certainty and territority, is that the Australian choice of law rule identifies or defines the law applicable to determine substantive rights (the lex causae) not the jurisdiction (or law area) which is to determine the relevant applicable law.\textsuperscript{58}

In other words, the Full Court concluded that the reasoning in\textit{Pfeiffer} suggests that the lex loci delicti rule is a rule for identifying the system of the law that will substantively determine the rights and liabilities of the parties, not merely a choice of law device that could lead to the application of some other system of law.

This conclusion was also said to be supported by the reasoning in\textit{Zhang}. According to the Full Court, the High Court’s overriding concern in\textit{Zhang} was to ensure that the chosen choice of law rule would produce ‘certainty and predictability’.

[The requirement for certainty and predictability is not limited to avoiding or minimising forum shopping. It is also aimed at avoiding uncertainty as to the law to be applied in resolving the dispute. It is the need for certainty as to the lex causae that resulted in the High Court refusing to countenance a flexible exception to the universal rule of the lex loci delicti notwithstanding it recognised that there may be situations in which the lex loci delicti would not be in accordance with the reasonable expectation of the parties.]\textsuperscript{59}

\textsuperscript{56} Appeal decision at [40].
\textsuperscript{57} Id at [42].
\textsuperscript{58} Id at [44].
\textsuperscript{59} Id at [47].
In the opinion of the Full Court, the *raison d’etre* of *Zhang* (and *Pfeiffer*) was to ensure that Australia’s choice of law rules were predictable and certain. As the passage quoted above makes clear, ensuring ‘certainty and predictability’ was seen to be a stand-alone objective supporting the lex loci delicti rule; it had intrinsic value, independent of considerations of forum shopping and party expectations. The promotion of ‘certainty and predictability’ was therefore seen to be an end in itself.

Applying this reading of *Zhang*, the Full Court reasoned that the application of renvoi in tort was inconsistent with the lex loci delicti rule. The Court concluded:

The application of the double renvoi doctrine to international torts would not promote certainty and predictability. It would require identification of Australia’s choice of law rules, the foreign country’s choice of law rules and its attitude to renvoi, from which a conclusion can then be reached as to the domestic law of which country applies. This exercise has the potential to be an “extraordinarily complex, unwieldy, phantasmagorical journey to make”.

Renvoi was thus said to be incompatible with the central policy concern expressed in *Pfeiffer* and *Zhang*. For this reason, the Full Court allowed the appeal on the renvoi ground and endorsed the internal law solution to renvoi in tort cases.

5. The Under-Whelming Force of Precedent

As noted above, the Full Court’s decision to reject renvoi in tort was supported by a seemingly overwhelming body of precedent in other jurisdictions and the authority of most academic commentators. Nevertheless, it is respectfully submitted that Full Court’s treatment of the existing authorities was subject to three significant flaws.

First, the Full Court went too far in concluding that there is ‘no binding (or other authority) that renvoi is applicable in torts cases’. In fact, the little Australian case law that exists on the subject suggests that renvoi may have application in tort cases, at least in respect of intranational torts. In *Breavington v Godleman* (*Breavington*), Mason CJ, Wilson, Deane and Gaudron JJ all appeared to endorse the application of the lex loci delicti in intranational torts cases, although they differed in the terms they used to express the applicable law. *Breavington* involved a claim for damages in negligence arising out of a motor vehicle accident in the Northern Territory. The action was commenced in the Supreme Court of Victoria and the issue was which system of law governed the claim. While Mason CJ simply endorsed the application of the ‘lex loci delicti’, Toohey J described the applicable law as ‘the law that would have been applied had

---

60 Appeal decision at [48].
61 Tilbury, Davis & Opeskin, above n55 at 1005.
62 Appeal decision at [39].
63 (1987) 169 CLR 41.
64 Id at 79 (Mason CJ); 98 (Wilson & Gaudron JJ); 138–9 (Deane J).
the appellant’s claim been brought in the Supreme Court of the Northern Territory.\textsuperscript{65} Deane J adopted a similar formulation.\textsuperscript{66} Wilson and Gaudron JJ described the applicable law as ‘the substantive law that would be applied if the matter were adjudicated in a court exercising the judicial power of the State in which the events occurred.’\textsuperscript{67} These statements suggest that Wilson, Gaudron, Deane and Toohey JJ endorsed an approach akin to the double renvoi solution whereby the court was required to place itself in the shoes of a foreign court and apply the system of law that the foreign court would apply. Presumably, this would involve a consideration of the foreign law’s choice of law rules as the foreign court would almost certainly decide the case according to those rules. While it would be going too far to suggest that the judgments in \emph{Breavington} involved any actual endorsement of the application of renvoi in tort, it is nevertheless clear that a majority of the Court approved of an approach that would presumably involve its application.

Secondly, the Full Court failed to recognise the simplistic and flawed reasoning upon which much of the existing case law rejecting renvoi in tort is based. The rejection of the doctrine of renvoi in tort in English law is thought to stem from the judgment of Lord Russell in \emph{M’Elroy v M’Allister}\textsuperscript{68} in which his Lordship casually observed:

\begin{quote}
\[\text{In referring to the lex loci delicti to ascertain by what rules the rights and liabilities of the parties to the action are regulated this Court refers to the internal domestic laws of that locus and not its private international law.}\]
\end{quote}

This was a curious dictum to add in a case in which no renvoi issue actually arose. The statement is explained, however, by the next step in Lord Russell’s reasoning. Relying on the principle stated in the passage quoted above, his Honour went on to hold that, as he was not bound to apply the foreign legal system’s rules of private international law (in this case English law), he was not required to apply English precedent relating to the characterisation of a limitation law as either procedural or substantive.\textsuperscript{70} While this conclusion is certainly correct, the reasoning on which it relies is flawed. As more recent cases have made clear, Lord Russell was not required to apply English precedent on the characterisation question because all such questions of characterisation are to be determined according to the lex fori.\textsuperscript{71}

In other words, Lord Russell did not need to say anything about renvoi to reach the conclusion he did on characterisation. In this context, it is difficult to see why Lord Russell’s confused dictum has been elevated to the status of high principle by textbook authors.

\textsuperscript{65} Id at 167.
\textsuperscript{66} Id at 138–139.
\textsuperscript{67} Id at 98.
\textsuperscript{68} (1949) SLT 139.
\textsuperscript{69} Id at 145.
\textsuperscript{70} Id at 146.
Similarly baseless reasoning is evident in some of the authority from the United States. In *Haumschild v Continental Casualty Co*, Currie J reasoned that renvoi must be rejected in tort because ‘it is likely to result in the court pursuing a course equivalent to a never ending cycle’. This reasoning is hardly convincing; renvoi is no more likely to produce a ‘never ending cycle’ in the law of tort than it is in the law of succession, yet it is rejected in the former context but embraced in the latter. Clearly, if a distinction is to be drawn between the application of renvoi in the law of succession and its treatment in tort, that distinction must be supported by some rational policy. Cases like *Haumschild* and *M’Elroy* demonstrate that the rejection of renvoi in tort has largely crept into the law on an unprincipled basis, having its genesis in the unreasoned dicta of judges, and finding general and unquestioning acceptance by commentators.

Thirdly, the Full Court did not come to terms with the fact that the utility of the doctrine of renvoi in tort will ultimately depend on the nature of the relevant choice of law rule, and foreign precedent rejecting renvoi will only be relevant to the extent that the relevant foreign legal systems adopt the same choice of law principles. For example, where the relevant choice of law rule requires a proper law of the tort approach, it would make little sense to apply renvoi because that may defeat the policy implicit in the rule that a foreign tort claim should be determined according to the system of law with which it has the most significant relationship. Thus, in the United States case of *Pfau v Trent Aluminum Company*, Proctor J rejected the application of renvoi in tort because, ‘[i]t was the purpose of the forum’s choice of law rule, the governmental interest analysis, was to ensure that the law of the state with the most significant interest in the case governed the claim.’ As such, the application of renvoi would defeat the policy underlying this choice of law rule because it would potentially produce arbitrary results by referring the court to the substantive law of a less interested state. Significantly, this same reasoning cannot be applied where a state uses a simple lex loci delicti rule, because the locus delicti might itself be a disinterested state. In fact, in a jurisdiction that uses a lex loci delicti approach, embracing renvoi may actually be a means of promoting the policy outcomes which the governmental interest analysis seeks to protect by rejecting renvoi. This reflects Otto Kahn-Freund’s observation that there is a stronger basis for rejecting renvoi in tort where

---

72 7 Wis 2d 130 (1959).
73 Id at 141–143.
76 Id at 526.
77 See *Mellk v Sarahson* 49 NJ 266, 229 A.2d 625 (1967).
80 Ibid.
a state has adopted a ‘proper law of the tort’ approach to choice of law, than there is in the context of a lex loci delicti rule. Thus, much of the foreign precedent rejecting renvoi is simply inapposite in the context of Australian choice of law rules.

6. Analysing Renvoi in the Context of Zhang

For this reason it was, with respect, entirely appropriate for the Full Court to consider whether the application of the doctrine of renvoi would promote or hinder the policy considerations that the High Court relied upon in Zhang and Pfeiffer in formulating the lex loci delicti rule. That said, in identifying the various policy bases upon which the decision in Zhang relied, the Full Court seemed to concentrate exclusively on the High Court’s concerns for certainty and predictability, ignoring the other considerations which informed its decision. When these other policy concerns are considered, it is clear that the application of renvoi (through the double renvoi solution) will usually be consistent with, and complementary to, the adoption of the lex loci delicti rule in Zhang.

A. Comity and Fidelity to the Law of the Foreign State

One of the principal reasons for adopting the lex loci delicti rule in Zhang was the High Court’s view that the rule was an expression of comity towards other states and gave due recognition to the competency of foreign countries to regulate activities within their own territory. The joint judgment of Gleeson CJ, Gaudron, McHugh, Gummow and Hayne JJ emphasised the merits of this approach by quoting La Forest J in the Canadian case of Tolofson v Jensen:

The underlying postulate of public international law is that generally each state has jurisdiction to make and apply law within its territorial limit. Absent a breach of some overriding norm, other states as a matter of “comity” will ordinarily respect such actions and are hesitant to interfere with what another state chooses to do within those limits.

In a separate judgment, Kirby J likewise argued that the lex loci delicti rule was necessary to respect the competence of states to control and regulate wrongs within their borders:

The law of tort, although now chiefly compensatory in purpose, has additional objectives of establishing standards of reasonable civic conduct, promoting prevention of wrongs and distributing costs amongst the community concerned. A choice of law rule that permits a plaintiff to pick and choose, according to the forum it selects, the law that would be applied, would derogate from the effective control of a given law area over those aspects of its law.

83 Zhang, above n1 at 516.
84 Id at 538.
In light of these statements, it is clear that one of the High Court’s reasons for adopting the lex loci delicti approach in *Zhang* was that it maintains a fidelity to the law and policies of the foreign state.

Turning to the issue of renvoi, it is clear that giving effect to the choice of law rules of a foreign state is consistent with this policy. If the reason for applying the lex loci delicti is a sense of comity towards foreign countries and recognition of their competence to regulate affairs within their own territory, why would we ignore a foreign state’s decision to have a legal controversy decided by the laws of some other country?85 This is certainly no less a regulatory decision than would be a decision to determine the matter according to the state’s own laws.

Indeed, a state may have legitimate reasons for resolving such disputes by reference to foreign law. For example, foreigners who have sustained tortious injury in China have often been paid compensatory damages well above those available under the Chinese law of civil liability because the Chinese Government is concerned that paying compensation at local rates would deter foreign investment and tourism.86 In light of this policy, it would not be surprising if the Chinese government selected a choice of law rule that applied foreign law in tort cases involving foreigners, thereby ensuring that the law of civil liability does not discourage international trade. If an Australian court, in deciding a case like *Neilson*, were to ignore the Chinese choice of law rules, the effect would be to defeat this regulatory policy and compromise the ability of China to regulate the activities of foreigners within its territory. Thus, it follows from *Zhang* that renvoi should be applied in tort because the rejection of a foreign state’s choice of law rules may compromise the regulatory autonomy of that state over matters occurring within its territory.

**B. Uniformity and a Protection Against Forum Shopping**

Another reason for the High Court’s adoption of the lex loci delicti approach in *Zhang* was that it promised greater uniformity between jurisdictions. The joint judgment emphasised that applying the lex loci delicti provided a ‘forum neutral connecting factor’ and hence promised ‘more even handed justice to both parties’.87 This reflected the Court’s observation in *Pfeiffer* that:

> From the perspective of the victim (the plaintiff) application of the lex loci delicti can be said to make compensation depend upon the accident of where the tort was committed, whereas, if the lex fori is applied, the plaintiff can resort to whatever forum will give the greatest compensation.88

---


88 *Pfeiffer*, above n2 at 539.
This is the dreaded scourge of ‘forum shopping’, which arises as a result of the lack of uniformity of internal laws, choice of law rules and procedural rules between states. The effect of these differences is that the legal result in a given case may ultimately depend on the forum in which it is litigated. 89 It is generally thought that forum shopping is an anathema to the principles upon which the law of conflicts is based because ‘the purpose of a choice of laws doctrine is to assure that a case will be treated the same regardless of the fortuitous circumstances which often determine the forum.’ 90 By adopting the lex loci delicti rule, the High Court believed that it was protecting against forum shopping by ensuring that a plaintiff who chooses to litigate a foreign tort claim in an Australian court, rather than in the state in which the wrong occurred, cannot thereby escape the application of the laws of that state.

This assumption only holds if the forum court also applies the choice of law rules of the foreign state. This can be demonstrated by re-examining the facts in Neilson. At trial, McKechnie J found that Article 106 of the General Principles provided for civil liability where one person’s conduct results in harm to another. 91 In this way, the Chinese law provided a basis for liability that, with respect to Mrs Neilson’s case, was broadly similar to that which exists under the common law of negligence. 92 That said, there were significant differences between Chinese law and the Australian common law in relation to the availability of the various heads of damage. While Article 119 of the General Principles provided for the payment of compensation for medical expenses, loss of income and loss of earning capacity, no damages are payable under Article 119 (or any other provision of the General Principles) in respect to pain and suffering. 93 As damages for pain and suffering are available under the Australian common law, any compensation paid in accordance with Chinese law was likely to be significantly lower than that which would have been available had common law principles applied. 94

This distinction between Chinese law and Australian law in relation to the available heads of damage becomes far more significant if Australian law rejects renvoi. Assume that Australian law adopts the internal solution to renvoi in tort such that the Supreme Court of Western Australia looks only to the internal law of China to resolve Mrs Neilson’s tort claim. Assume also, for ease of explanation, that Chinese law also rejects the renvoi. The result would be that, were she to litigate in Western Australia, Mrs Neilson would be entitled to medical expenses and compensation for loss of income, but could not recover damages for pain and

90 Lauritzen v Larsen 345 US 571, 591 (1953) (Jackson J).
91 Trial decision at [144].
92 Ibid.
93 Id at [157]–[159].
94 Id at [160]. The monetary value of this distinction is unclear from the trial decision in Neilson as the parties had earlier agreed as to the quantum of damages payable should OPC be found liable under Australian law.
However, had Mrs Neilson litigated her claim in China, the Chinese court could apply Article 146 of the General Principles and resolve the matter according to the Australian common law. This would mean that Mrs Neilson could recover damages for pain and suffering. In this way, the choice of forum greatly affects Mrs Neilson’s entitlement to compensation. As Adrian Briggs notes, the application of renvoi provides a defence against this type of forum shopping because it ensures that a forum court decides a matter exactly as the foreign court would do.\(^9\) If the Supreme Court of Western Australia had to decide Mrs Neilson’s claim by exactly the same principles that a Chinese court would adopt, the effect would be to substantially neutralise the effect of forum selection on the legal consequences of the case. In this context, the application of renvoi is critical to the policy of uniformity expressed by the High Court in *Zhang* and *Pfeiffer*.

That said, it is important not to overstate the extent to which applying a foreign state’s choice of law rules will promote uniformity and prevent forum shopping. Even if we were to adopt renvoi in tort, there remain several practical and legal obstacles to ensuring that foreign tort claims are resolved in a forum-neutral manner. As forum courts are bound to apply forum public policy and procedural law, total fidelity to the laws of the foreign country is simply not possible.\(^9\) Renvoi is therefore clearly not a complete solution to the problem of forum shopping or the lack of international uniformity in relation to substantive law. Notwithstanding this qualification, renvoi is properly regarded as an important tool by which a court can achieve greater inter-jurisdictional uniformity and discourage forum shopping, thereby promoting one of the underlying policies of the decision in *Zhang*.

**C. Party Expectations**

Moreover, the application of renvoi will contribute to the attainment of a third policy objective behind the lex loci delicti rule; that of meeting reasonable party expectations. In *Zhang*, Kirby J claimed that the lex loci delicti rule has ‘sure foundations in human psychology’\(^9\) because a person will ‘ordinarily assume that he or she is governed by the law of the law area in which the event, critical to legal liability, happens.’\(^9\) A similar view was expressed by the joint judgment in *Pfeiffer*.\(^10\) As Kahn-Freund has explained, party expectations play a role in choice

---

95 It should be noted that it is well-established that the availability of particular heads of damages is a question of substance, not procedure, and will hence be governed by the lex causae: *Stevens v Head* (1993) 176 CLR 433 at 459–460. In *Zhang*, the Court left undecided the question of whether matters going to the quantification of damages should now also be considered to be substantive; *Zhang*, above n1 at 520. Nothing the Court said in *Zhang* called into question the principle that matters relating to the available heads of damage are properly considered to be substantive.

96 Briggs, above n55 at 879–882.

97 Bell, above n89 at 26–36, 46–47; Rimmel, above n55 at 60–61.


99 Ibid (Kirby J).

100 *Pfeiffer*, above n2 at 540 (Gleeson CJ, Gaudron, McHugh, Gummow & Hayne JJ).
of law policy because potential tortfeasors should be able to calculate the extent to which their conduct exposes them to the risk of legal liability.\textsuperscript{101} In Kahn-Freund’s words, ‘They should be able to feel safe in Rome if they do as the Romans do.’\textsuperscript{102}

At first glance, this consideration may be thought to tell against the application of renvoi in tort because the operation of the doctrine may result in the application of some system of law other than the lex loci delicti. After all, how can party expectations be satisfied if, when in Rome, one does as the Romans do, only to have the law of Gaul apply? However, once it is appreciated that the concern with party expectations is ultimately for ensuring that potential tortfeasors can calculate their liability (and that potential tort victims can determine the extent of their legal protection), it is clear that the operation of renvoi is consistent with this objective because it will usually ensure that the full extent of a tortfeasor’s liability is referable to a single legal system. This point is aptly demonstrated from the facts in \textit{Neilson}. Assuming that Australian law applies the double renvoi solution, OPC would be able to calculate the full extent of its liability according to the law of China as it would be applied by the Chinese courts. By contrast, were we to reject the renvoi, OPC would have to calculate its potential liability not simply according to Chinese law as it applies in China, but also in terms of Chinese law as it would be applied by Australian courts, and Chinese law as it might be applied in numerous other jurisdictions. This latter scenario would entail significant expense for Australian companies seeking to do business overseas as they would be forced to consider their exposure to tortious liability in numerous jurisdictions. Party expectations are thus better served by applying renvoi in tort as the doctrine can alleviate much of the uncertainty involved in multi-jurisdictional operations.

\textbf{D. ‘Certainty and Predictability’}

The foregoing analysis sheds some light on what the High Court meant in \textit{Zhang} when it stressed the importance of ‘certainty’ in choice of law:

\begin{quote}
The selection of the lex loci delicti as the source of substantive law meets one of the objectives of any choice of law rule, the promotion of certainty in the law. Uncertainty as to the choice of the lex causae engenders doubt as to liability and impedes settlement.\textsuperscript{103}
\end{quote}

Certainty is valued because it promotes the rapid settlement of claims and removes difficult legal obstacles that impede the resolution of disputes. This is consistent with the other policy goals outlined above such as the minimisation of forum shopping and the satisfaction of party expectations. Thus, contrary to the interpretation of \textit{Zhang} adopted by the Full Court, the High Court never suggested that certainty and predictability were ends in themselves. Their value arises from the fact that certainty and predictability facilitate international transactions by minimising the legal confusion that can arise from multi-jurisdictional liability.

\textsuperscript{102} Ibid.
\textsuperscript{103} \textit{Zhang}, above n1 at 517 (Gleeson CJ, Gaudron, McHugh, Gummow & Hayne JJ).
In this context, the Full Court’s conclusion that the application of the doctrine of renvoi is destructive of certainty and predictability cannot be maintained. In *Neilson*, McLure J reasoned:

The High Court in *Zhang* has deliberately selected a rigid choice of law rule in tort to promote certainty and predictability. It would be inconsistent with the reasoning and result in *Zhang* to superimpose a *renvoi* doctrine the purpose and effect of which is to soften or avoid the rigidity of choice of law rules.¹⁰⁴

This cannot, with respect, be correct. The High Court in *Zhang* endorsed an approach to choice of law that would minimise uncertainty as to liability and remove legal impediments to international transactions. Provided that the application of renvoi facilitates these objectives, the doctrine’s application will be consistent with the rule adopted in *Pfeiffer* and *Zhang*. In this context, questions of whether the rule is rigid or not are somewhat besides the point. The lex loci delicti rule is simply a mechanism for implementing a particular set of policies. Ultimately, renvoi should be applied if its application is consistent with those policies and rejected if it is not. The fact that the mechanism may be described as ‘rigid’ offers nothing. For example, the rigid rejection of renvoi will preserve discrepancies in the way in which the same cause of action is decided between jurisdictions; it will be productive of uncertainty as to liability and will thereby reduce the chance of settlement. By contrast, the application of renvoi in tort should minimise these differences thereby ensuring greater fidelity to the reasoning in *Pfeiffer* and *Zhang*. The fact that the former approach may be described as ‘rigid’, and the latter as ‘flexible’, can have no bearing on which approach is adopted. Only the latter is consistent with the policies which underlie *Pfeiffer* and *Zhang* and therefore only the latter should be applied.

### 7. Dealing with Evidentiary Uncertainty

One final comment needs to be made with respect to questions of certainty. As the Full Court noted, a degree of uncertainty is inherent in applying the double renvoi solution because some legal systems lack a clearly articulated attitude to renvoi. The difficulties involved in the double renvoi solution are aptly captured in the judgment of Wynn-Parry J in *Re Duke of Wellington*:¹⁰⁵

> [I]t would be difficult to imagine a harder task than that which faces me, namely, of expounding for the first time either in this country or in Spain the relevant law of Spain as it would be expounded by the Supreme Court of Spain, which up to the present time has made no pronouncement on the subject, and having to base that exposition on evidence which satisfies me that on this subject there exists a profound cleavage of legal opinion in Spain, and two conflicting decisions of courts of inferior jurisdiction.¹⁰⁶

---

¹⁰⁴ Appeal decision at [49].
¹⁰⁵ [1947] Ch 506.
¹⁰⁶ Id at 515 (Wynn-Parry J).
This reflects Elliot Cheatham’s observation that, ‘[I]t is difficult to ascertain and apply foreign local law; it is more difficult to ascertain and apply foreign conflict of laws rules; it is yet more difficult to ascertain and apply the foreign rule on the acceptance and rejection of renvoi.’

Given that such difficulties are inherent in the double renvoi solution, would not the application of renvoi in tort simply create the sort of uncertainty Zhang was intended to avoid? Regardless of the policy considerations that may support it, would not a rule that cannot be applied with certainty also undermine international transactions and disappoint party expectations?

While the logical force of this objection is unimpeachable, its premise, that a foreign country’s attitude to renvoi is near-impossible to prove, is increasingly without foundation. Over fifty years have passed since Wynn-Parry J had to contend with Spain’s silence on the renvoi question; nearly seventy-five years have passed since Maugham J famously observed that the application of renvoi was undesirable because it ultimately depends on the ‘doubtful and conflicting evidence of foreign experts’. During this time, two significant developments have occurred; first, access to sources of foreign law have improved significantly, not least with the innovation of the internet and online legal materials. Secondly, and more significantly, this period has been marked by the adoption in many countries of specific and clearly formulated attitudes to renvoi.

The extent of this latter development was captured by J Georges Sauveplanne in his 1988 survey of renvoi laws around the world. Sauveplanne’s survey reveals that countries as diverse as France, the Netherlands, Turkey, Brazil, Thailand, Japan and Egypt all have discernible attitudes to renvoi. Even Spain and Italy, two countries that have troubled the common law courts in the past, have now made clear their attitudes to the doctrine. This process has also been assisted by the increased codification of private international law rules in states such as Austria and Portugal. The upshot of these developments is that states are now both more aware of the renvoi difficulty and more vocal in relation to their attitude to its operation. Thus, whatever force the evidentiary uncertainty argument previously enjoyed, its ongoing rhetorical force is, if not yet spent, certainly in decline. It is not nearly as difficult as it once was to determine a foreign state’s attitude to renvoi; the effect of which is that it cannot be assumed that the application of the double renvoi solution will create any great uncertainty as to the applicable law.

---

107 Cheatham, above n85 at 338.
108 Re Askew [1930] 2 Ch 259 at 278.
110 Id at 15–33.
112 Sauveplanne, above n109 at 26.
113 Id at 25–33.
8. Conclusion

*Neilson* in many ways represents the paradigm case in which renvoi should not be rejected in tort. First, the governing law of the locus delicti expressed a clear preference that cases involving foreign nationals be resolved according to law of the nationality. Secondly, the reasonable expectations of the parties may well have been that the operation of Article 146, together with their close contacts with Australia, would ensure that Australian law would govern any claim between them. Thirdly, the rejection of renvoi meant that the extent of the defendant’s liability differed depending upon the jurisdiction in which the action was litigated. Thus, by rejecting the renvoi, the Full Court defeated each of the significant policy considerations which the High Court relied upon in *Zhang* for adopting the lex loci delicti rule.

While it is clearly simpler and more convenient to reject the renvoi altogether, to do so is simply inconsistent with Australia’s choice of law rule in tort. The High Court’s concern for certainty and predictability in *Pfeiffer* and *Zhang* was not motivated by a desire to ensure that cases were easily dealt with, or evidentially simple; rather, the Court endorsed the lex loci delicti approach because the certainty it afforded would facilitate international transactions and remove impediments to settlement. The application of renvoi will not defeat these objectives, and will usually be consistent with them. For this reason there is no basis in policy or precedent for rejecting renvoi.
Abstract

This short article evaluates Part 3 of the Civil Liability Act 2002 (NSW). The author compares the legislative treatment of plaintiffs seeking to recover compensation for mental harm in NSW (formerly known as ‘nervous shock’), resulting from the NSW government’s extensive tort law reform in 2002, with that of the common law. In this context, the concepts of psychiatric injury, proximity, sudden shock and normal fortitude are analysed. It is argued that the legislative provisions are regrettable and an unfortunate intrusion into the common law of negligently inflicted psychiatric injury. The author concludes that the effect of the new legislative provisions on mental harm is to impede the logical development of the common law in this area, which the High Court of Australia signalled would flow from its decision in Tame v New South Wales; Annetts v Australian Stations Pty Ltd.

* BComm LLB (Hons I). I would like to thank Ross Anderson for his time and patience in the preparation of this comment. All opinions expressed and any errors are mine.
1. Introduction

As part of an attempt to reform negligence law in New South Wales, Part 3 of the Civil Liability Act 2002 (‘the Act’) \(^1\) represents a significant narrowing of the scope of the common law action for pure mental harm. \(^2\) The common law principles governing recovery for negligently inflicted psychiatric injury, particularly those respecting the duty of care, were restated in Tame v New South Wales; Annetts v Australian Stations Pty Ltd \(^3\) prior to the introduction of the legislation. Part 3 contains broader coverage of the area than the High Court decision in Tame and Annetts. \(^4\) However, given that Tame and Annetts \(^5\) was an attempt to address the more recurrent and pressing uncertainties in this area of the law, the following evaluation will only focus on those sections of Part 3 that respond to the common law principles as recently (re)determined. \(^6\)

First, this comment will briefly outline the crux of the decision in Tame and Annetts. Part 3 of the comment will then evaluate the Act by comparing the common law position on the need to establish a recognisable psychiatric injury, and on the relevance of geographical, temporal and relational proximity, sudden shock, and normal fortitude with the referable legislative provisions. In Part 4, the author concludes, arguing that the developments in the common law have been

\(^1\) This Act, as amended by the Civil Liability Amendment (Personal Responsibility) Act 2002 (NSW), was an effort to confine the operation of the common law principles of negligence in response to ‘widespread public concern about large increases in the cost of … public liability insurance, the impact of this on the community and government’, and the perception that courts were awarding some claimants ‘unjustified and/or excessive damages’: John Ball, ‘Uniform Tort Law: “A Fairer System for a Fairer Result”, Recommendations for a Balanced System That Treats Everyone Equally’, Law Society of New South Wales (2002) at 8. See also Reg Graycar, ‘Public Liability: A Plea for Facts’ (2002) 54 Plaintiff 37 at 37. It was felt that ‘[t]he insurance crisis served to highlight just how far the law has drifted away from the concept of personal responsibility’, and the amending Act was ‘intend[ed] to filter out unmeritorious claims and return tort law to a system that protects people who have been harmed by the bona fide negligence of others and who therefore deserve fair compensation’: New South Wales, Legislative Council, Parliamentary Debates (Hansard), No 52, 19 November 2002 <http://parliament.nsw.gov.au/prod/lec/chansar52.nsf/> (accessed 14 March 2003). The provisions in Part 3 largely adopt the recommendations on mental harm in the report of the Review of the Law of Negligence Panel, which was convened to inquire into tort law and to make suggestions for a principled approach to its reform: Jocelyn Kellam, ‘Negligence Review Has Broad Ramifications’ (2002) 13 Australian Product Liability Reporter 57 at 59. For commentary on the review basis, its assumption that existing tort law principles are contributing to unaffordable and unsustainable public liability insurance, and its approach to tort law reform, see Kellam generally. See also Australian Plaintiff Lawyers Association, ‘Tort Reform and Consumer Protection’ (2002) 54 Plaintiff 46; Law Council of Australia, ‘Second Submission by the Law Council of Australia to the Negligence Review Panel on the Review of the Law of Negligence’, paper submitted to the Review of the Law of Negligence Panel, 2 September 2002 at 2–5.

\(^2\) By pure mental harm is meant psychiatric injury that is not consequent on personal injury of any other kind: see definitions in s27 of the Act.

\(^3\) (2002) 191 ALR 449 (hereinafter ‘Tame and Annetts’).

\(^4\) Ibid.

\(^5\) Ibid.

\(^6\) Other provisions of the Act that will not be discussed include the Part’s excluded application from workplace injuries but not from motor vehicle accident injuries: s5A, and proportionate reduction of damages based on contributory negligence of the immediate victim: s30(3).
desirable in many respects, and that in winding back these advances and forestalling future developments in the common law, the NSW legislature has taken a regrettable step backwards.

2. The Position at Common Law

Just prior to the introduction of the Act, in *Tame and Annetts*, the High Court of Australia clarified the common law position on negligently inflicted psychiatric illness. The High Court retained the need for a plaintiff to demonstrate a recognisable psychiatric injury. For the purposes of cases in which the fault of the defendant immediately harms a person other than the plaintiff, the High Court declined to close the class of eligible claimants according to the directness of their perception of the event in question or its aftermath, and the closeness of their ties with the person imperilled. These circumstances, together with sudden shock, were rejected as preconditions to recovery, and were deemed potentially relevant considerations in the overall question of whether a duty of care was owed only. A majority also held that there was no need to establish reasonable foreseeability of injury to a person of normal fortitude.

3. Evaluating the Civil Liability Act 2002 Part 3

A. Recognisable or Recognised Psychiatric Injury

Section 31 of the Act mirrors the common law in stipulating that only recognisable psychiatric illness is compensable. From early on, courts have permitted recovery for recognisable psychiatric illness but not for mere emotional distress or suffering. The distinction is predicated on policy considerations, particularly

---

7 Above n3. This is the first time since *Jaensch v Coffey* (1984) 155 CLR 549, decided almost 20 years ago, that the High Court has revisited the principles governing the duty of care in psychiatric injury cases. While a significant decision in itself in allowing recovery in the circumstances of the case, *Jaensch v Coffey* left the law uncertain in many respects. The High Court was pressed and attempted to resolve these in *Tame and Annetts*.

8 *Tame and Annetts*, id at 460 (Gaudron J); 495–496 (Gummow & Kirby JJ); 522 (Hayne J) and 542 (Callinan J).

9 Id at 456 (Gleeson CJ); 461–462 (Gaudron J); 494, 505 (Gummow & Kirby JJ) and 517 (Hayne J). Callinan J dissented at 451.

10 The idea of restricting claimants according to family relations was not considered an option by the High Court: see id at 460 (Gleeson CJ) and 461 (Gaudron J).

11 Id at 456 (Gleeson CJ); 465 (Gaudron J); 494, 500, 502 (Gummow & Kirby JJ) and 519–520 (Hayne J). Callinan dissented at 540, 541.

12 Id at 455–456 (Gleeson CJ); 464 (Gaudron J) and 494, 497, 498 (Gummow & Kirby JJ).

13 Section 31 uses the term ‘recognised psychiatric illness’. There does not seem to be any significance in the use of ‘recognised’ as opposed to ‘recognisable’.

14 See for example, *Coultas v Victorian Railways Commissioners* (1886) 12 VLR 895 at 896 (Williams J); *Dulieu v White* [1901] 2 KB 669 at 672–3 (Kennedy J); *Chester v Vaverley Corporation* (1939) 62 CLR 1 at 7–8 (Latham J); *Behrens v Bertram Mills Circus Ltd* [1957] 2 QB 1 at 28 (Devlin J); *Hinz v Berry* [1970] 2 QB 40 at 42–43 (Lord Denning MR); *Mount Isa Mines Ltd v Pusey* (1970) 125 CLR 383 at 394 (Windeyer J); *McLoughlin v O’Brien* [1983] 1 AC 410 at 431 (Lord Bridge); *Jaensch v Coffey*, above n7 at 559–60 (Brennan J) and 587 (Deane J).
fear of opening the floodgates,\textsuperscript{15} given, inter alia, the ubiquity of emotional reactions to a given stimulus and a perception that psychiatric injury is less verifiable than physical injury.\textsuperscript{16}

However, it is still possible to address this concern, assuming it is valid,\textsuperscript{17} without requiring proof of recognisable psychiatric injury. The problem with ‘recognisable psychiatric illness’ is that it is elusive. As Butler argues, despite its medically aligned terminology, it is questionable whether psychiatry and courts can properly distinguish between mere emotional distress and recognisable psychiatric injury.\textsuperscript{18} This makes the use of the standard to determine the viability of a claim potentially unfair and highly unsatisfactory. It is suggested therefore, that a preferable approach is to question whether the plaintiff has been rendered unable to restore psychiatric homeostatic equilibrium, and if so, has this adversely affected the plaintiff’s normal enjoyment of life.\textsuperscript{19} The first limb accommodates the law’s policy refusal to compensate mere grief and sorrow,\textsuperscript{20} and its practice of


\textsuperscript{16} Butler, ‘Identifying the Compensable Damage in “Nervous Shock” Cases’, id at 73–74.

\textsuperscript{17} Mullany & Handford argue that the validity of policy concerns is overstated: above n15 at 43–44.

\textsuperscript{18} Butler ‘Identifying the Compensable Damage in “Nervous Shock” Cases’, above n15 at 78–9. There will be some cases where it is clear that the plaintiff has suffered a recognisable psychiatric injury rather than just emotional distress and vice versa, but there will also be borderline cases where the harm sustained cannot be objectively categorised. See also Hayne J’s discussion in Tame and Annetts, above n3 on psychiatric illness and emotional distress at 522–523, and contrast with Gummow & Kirby JJ’s discussion at 496. For other criticisms, see Nicholas Mullany, ‘Distress, Disorder and Duty of Care: The New Zealand Front’ (2001) 117 \textit{Law Quarterly Review} 182 at 186 and Harvey Teff, ‘Mental Suffering Revisited in New Zealand’ (2001) 9 \textit{Tort Law Review} 109 at 113.

\textsuperscript{19} Butler, id at 82–84.

\textsuperscript{20} Id at 84.
admitting expert testimony. The second limb resonates with the concept of 'normal enjoyment of life' relevant in the assessment of damages, and is a fairer criterion in a multicultural society where the notion of mere emotional distress is less definable.

Arguably then, it would have been preferable had the legislature reformed the common law and removed this evidently, theoretically and practically unsound barrier to action, especially considering now that the risks that are to be deemed foreseeable have narrowed.

B. Geographical and Temporal Proximity: Direct Perception at the Scene of the Accident

In cases where psychiatric injury is sustained as a result of death, injury or peril to someone other than the plaintiff, as well as having to prove a recognisable psychiatric illness, the Act permits recovery only if the plaintiff witnessed the incident at the scene, or alternatively, is a close relative of the immediate victim. Thus the Act reintroduces the notion of geographical and temporal proximity rejected as irrational and arbitrary by the High Court in Tame and Annetts.

As the High Court reasoned, presence at the scene must be rejected as a precondition to the founding of duty for its want of logicality. It is difficult to see why for example, other than for floodgates fears, a plaintiff who is present at the scene and sees a close friend being hit and seriously injured by a motor vehicle is able to recover, whereas a plaintiff who is also a close friend but only able to

21 Ibid.
22 Id at 84–85.
23 Id at 85.
24 Section 5B defines reasonable foreseeable risks as risks that are 'not insignificant'. This is intended to narrow the range of risks that would otherwise be held to be reasonably foreseeable under the common law test of 'not far-fetched or fanciful': see New South Wales, Legislative Council, Parliamentary Debates (Hansard), No 52, 19 November 2002, above n1.
25 The victim will be subsequently referred to as 'the immediate victim'.
26 Section 30.
27 This had origins in Hambrook v Stokes Bros [1925] 1 KB 141, and was subsequently applied by the High Court in Chester v Waverley Corporation, above n14.
29 Tame and Annetts, id at 456 (Gleeson CJ); 461–462 (Gaudron J) and 494–495, 504–505 (Gummow & Kirby JJ). In the words of Gummow & Kirby JJ at 504 with whom Gleeson CJ at 456 and Gaudron J at 461–462 agreed:

A rule that renders liability in negligence for psychiatric harm conditional on the geographic or temporal distance of the plaintiff from the distressing phenomenon, or the means by which the plaintiff acquires knowledge of that phenomenon, is apt to produce arbitrary outcomes and to exclude meritorious claims.

Gaudron J at 462 also noted that requiring geographical and temporal proximity has the effect of limiting the categories of possible claimants in disconformity with the neighbourhood principle recognised in Donoghue v Stevenson [1932] AC 562.
attend the hospital a few hours later cannot.\textsuperscript{30} Clearly, ‘because the supposed rule lacks any principled foundation, it mandates differential treatment of plaintiffs in substantially the same position.’\textsuperscript{31} Hence the courts were soon compelled to ameliorate the irrationality of the rule, developing the ‘aftermath’ doctrine to allow recovery in cases in which it was foreseeable that a plaintiff not present at the scene of the accident would arrive shortly afterwards.\textsuperscript{32} This was then extended to include attendance at the hospital: \textit{McLoughlin v O’Brian},\textsuperscript{33} and further again to encompass ‘the period of immediate post-accident treatment’: \textit{Jaensch v Coffey}.\textsuperscript{34}

Thus floodgates fears or other policy considerations have not restrained the courts from attenuating the rigour of the presence rule and ultimately rejecting its utility as a precondition to recovery. Nonetheless, intent on curtailing tortious liability, the legislature has felt it necessary to reintroduce this restriction for plaintiffs who are not relatives, and in its pre-\textit{McLoughlin v O’Brian}\textsuperscript{35} form. Insofar as being present and witnessing the accident may be relevant to whether or not psychiatric injury was foreseeable, it is submitted that presence can only appropriately be a relevant consideration. To go further than this – s32(2)\textsuperscript{36} – becomes indefensible.

Providing further support for this is the fact that to require presence is to require perception with unaided senses. To witness something as it happens is to see it first hand. Over the years, cases have prohibited recovery because the relevant event was not perceived of through unaided senses.\textsuperscript{37} The High Court in \textit{Tame and Annetts}\textsuperscript{38}

\begin{flushleft}

\textsuperscript{31} \textit{Tame and Annetts}, above n3 at 505. Continuing at 505, Gummow & Kirby JJ said that to transform what may be a relevant consideration to a prerequisite for duty in all cases ‘carries the risk of attribution of disproportionate significance to what may be no more than inconsequential circumstances.’

\textsuperscript{32} See \textit{Boardman v Sanderson} [1964] 1 WLR 1317; \textit{Storm v Greeves} [1965] Tas SR 252; \textit{Mount Isa Mines Ltd v Pusey}, above n14; \textit{Chadwick v British Railways Board} [1967] 1 WLR 912; \textit{Benson v Lee} [1972] VR 879. This development was noted by Gummow and Kirby JJ in \textit{Tame and Annetts}, id at 495.

\textsuperscript{33} Above n14, see particularly at 422 (Wilberforce LJ).

\textsuperscript{34} Above n7 at 607–8 (Deane J).

\textsuperscript{35} Above n14.

\textsuperscript{36} Section 32(2)(b) provides that whether the plaintiff witnessed at the scene a person being killed, injured or put in peril is relevant to the inquiry of foreseeability.

\textsuperscript{37} See for example, \textit{Hambrook v Stokes Bros}, above n27 at 152 (Banks LJ); 159 (Atkin LJ) and 165 (Sargant LJ); \textit{Chester v Waverley Corporation}, above n14 at 43 (Evatt J); \textit{Bourhill v Young} [1943] AC 92 at 103 (MacMillan LJ); \textit{King v Phillips} [1953] 1 QB 429 at 441 (Denning LJ); \textit{Mount Isa Mines Ltd v Pusey}, above n14 at 407 (Winedey J); \textit{Benson v Lee}, above n32 at 422–423 (Wilberforce LJ); \textit{Jaensch v Coffey}, above n7 at 567 (Brennan J) and 612 (Dawson J); \textit{Alcock v Chief Constable of South Yorkshire Police}, above n14 at 398 (Keith LJ); 400–401 (Ackner LJ); 416–418 (Oliver LJ) and 423 (Jauncey LJ); \textit{Annetts v Australian Stations Pty Ltd}, above n30 at 61, 63 (Ipp J, Malcolm CJ and Pidgeon J agreeing); \textit{Gifford v Strang Patrick Stevedoring} (2001) 51 NSWLR 606 at 616–617, 622 (Hodgson JA, Handley JA and Ipp AJA agreeing). See also Mullany & Handford, above n15 at 153–166.
\end{flushleft}
has held that this rule, apparently based on floodgates fears,\(^3^9\) cannot be defended as an elemental part of the action. As medically accepted, learning of a traumatic event, for instance by third party communication, is not necessarily less affronting and lasting on the mind than seeing something happen with one’s own eyes at the scene.\(^4^0\) For example, it makes little sense to deny relief to a godparent who learns of the death of his or her godchild from a car collision by telephone but not to an acquaintance who happened to be a passenger in the same car and survived. The rule is also ‘disjoined from the realities of modern communication.’\(^4^1\) Arguably, a rule so devoid of rationality and arbitrary is indefensible on policy grounds.\(^4^2\) Arguably, it makes a mockery of the legal system.

C. Relational Proximity: Close Ties

As noted, s30 only discards geographical and temporal proximity if plaintiffs are ‘close members of the family of the [immediate] victim’ as defined.\(^4^3\) It would seem that the legislature takes the view that defined relatives of the immediate victim should be allowed to recover damages as a matter of policy, or that psychiatric illness may be foreseeable to these plaintiffs regardless of proximity factors.\(^4^4\) This resembles the common law position insofar as it recognises the foreseeability of close relatives being mentally afflicted by somehow learning of a traumatic event involving a member of family,\(^1^5\) which is just a reflection of reality.

---

38 Above n3 at 456 (Gleeson CJ); 461–462 (Gaudron J) and 494–495, 504–505 (Gummow & Kirby JJ).
39 Note the reasoning and decision in Alcock v Chief Constable of South Yorkshire Police, above n30.
40 McLoughlin v O’Brian, above n14 at 442 (Bridge LJ); Jaensch v Coffey, above n7 at 555 (Gibbs CJ); 606–609 (Deane J); Coates v Government Insurance Office of New South Wales (1995) 36 NSWLR 1 at 8–11 (Kirby P), 5, 7 (Gleeson CJ) and 22–23 (Clarke JA); and more generally, the outcome in Alcock v Chief Constable of South Yorkshire Police, ibid. See also Mullany & Handford, above n15 at 166–170. Some decisions at first instance have even abandoned the necessity of the principle: see for example, Petrie v Dowling, above n28; Reeve v Brisbane City Council, above n28; Quayle v State of New South Wales (1995) Australian Torts Reports 81–367. In South Africa the South African Supreme Court of Appeal has rejected the unaided senses rule: Barnard v Santam Bpk 1999 (1) SA 202; Nicholas Mullany, ‘Personal Perception of Trauma and Sudden Shock – South Africa Simplifies Matters’ (2000) 116 Law Quarterly Review 29.
41 Tame and Annetts, above n3 at 504 (Gummow & Kirby JJ). See also Coates v Government Insurance Office of NSW, id at 11 (Kirby P).
42 Note that there is actually no evidence that permitting recovery for psychiatric illness induced by learning of a traumatic event other than directly would proliferate claims: see Thomas J (in dissent) at 203 in van Soest v Residual Health Management Unit, above n15.
43 Section 30(5) defines ‘close member of the family’ as (a) a parent or other person with parental responsibility for the immediate victim; or (b) the spouse or de facto spouse of the immediate victim; or (c) a child or stepchild of the immediate victim or other person for whom the victim has parental responsibility; or (d) a brother, sister, half-brother or half-sister, or stepbrother or stepsister of the victim.
44 Compare this to the view of the English Law Commission that where there is a close tie of love and affection between the plaintiff and immediate victim the requirement of direct perception should be abandoned, above n15 at 88.
45 McLoughlin v O’Brian, above n14 at 422 (Wilberforce LJ); Jaensch v Coffey, above n7 at 605–606 (Deane J) and 569–570 (Brennan J); note also 555 (Gibbs CJ) relating relationship with proximity; Alcock v Chief Constable of South Yorkshire Police, above n30 at 337 (Hidden J); 397 (Keith LJ); 415–416 (Oliver LJ) and 422 (Jauncey LJ).
However, the High Court specifically left open the definition of ‘close members of family’. Thus the common law appreciates that the closeness of certain relations, such as that of parent and child and husband and wife, varies, and accordingly, their response to traumatic events involving their relatives varies. Thus the common law also recognises that some plaintiffs, although not parent, sibling or spouse, may be just as close or even closer to the person injured, killed or imperilled than a parent, sibling or spouse. For example, a child may in fact be closer to a friend than a stepbrother or stepsister.

The legislation also gives regard to the fact that there inevitably is variation in the closeness of bonds in a given family relationship through s32. However, because membership of only certain specified relationships entitles a plaintiff to claim, a close relationship with the immediate victim that would make the infliction of psychiatric injury reasonably foreseeable is necessarily non-compensable if it falls outside the defined class in s30(5).

It is not hard to imagine that s30(2) will work injustice to some plaintiffs, such as fiancées, who were not present at the scene of the accident, but nonetheless, should have been contemplated as likely to suffer mental harm. Therefore, to move away from the flexibility of the common law position and categorically define who may claim, as the legislation does, is regrettable.

**D. Sudden Shock**

In respect of sudden shock, however, the legislation admits of more flexibility. As noted, the High Court rejected the sudden shock rule as essential, seeing it only as relevant to the foreseeability inquiry. Using the increase in diagnoses of depressive illness and the need to reach a decision based on sometimes conflicting medical evidence, courts have sought to justify the need for shock to stem the flow of claims.

46 Tame and Annetts, above n3 at 460 (Gleeson CJ) and 461 (Gaudron J). As Gaudron J stated at 461: Much of the reasoning in Jaensch v Coffey pointed to the illogicality of excluding claims by those in a close and personal relationship with the person killed or injured, but the actual decision serves to signify, in the words of Brennan J in that case, that “the categories of claimants [who suffer pure psychiatric injury] are not closed”.

47 For further examples, see the discussion of the judgments at the three levels in Alcock v Chief Constable of South Yorkshire Police in Mullany & Handford, above n15 at 119–126.

48 Section 32(2)(c) provides that the nature of the relationship between the plaintiff and any person killed, injured or put in peril is relevant to the question whether or not psychiatric illness is foreseeable.

49 Compare the fiancée plaintiff in Alcock v Chief Constable of South Yorkshire Police, above n30.

50 Section 32(2)(a) provides that ‘whether or not the mental harm was suffered as the result of a sudden shock’ is a factor to consider in the question whether the harm was foreseeable. ‘Sudden shock’ means shock in the sense described by Brennan J in Jaensch v Coffey, above n7 at 567: ‘the sudden sensory perception – that is, by seeing, hearing or touching – of a person, thing or event, which is so distressing that the perception of the phenomenon affronts or insults the plaintiff’s mind and causes a recognisable psychiatric illness.’ The need to establish shock denies claims for gradual onset of, or cumulative events causing, psychiatric disorders: see for example, Pratt v Pratt [1975] VR 378.
It is argued that rationality and plausibility require the rejection of sudden shock as a precondition to recovery. For one thing, it has been pointed out that the floodgates concern is misplaced. As Teff states, ‘no liability arises unless factors inducing psychiatric illness are demonstrably bound up with … the accident’. 54

Secondly, the notion of a single isolated event causing psychiatric illness is medically unsound. 55 Thirdly, faced with the rule, witnesses may ‘distort opinions on what they may feel to be legitimate claims out of deference to [the law]’. 56 Furthermore, the shock requirement sits incongruously with the common law approach to mental suffering consequent upon property damage, which does not require shock. 57 It is also conceptually irreconcilable with the ‘eggshell skull’ rule, which enables a plaintiff to recover for the full extent of his or her illness if the kind of damage sustained was foreseeable, even though the actual gravity or extent of the damage was not foreseeable. 58

In embracing the common law’s rejection of shock as a precondition to liability, the legislature has given recognition to the artificiality and problematic application of the rule. However, arguably it would have been better had the legislature gone further and declined to specifically provide for its relevance in

51 Tame and Annetts, above n3 at 456 (Gleeson CJ); 465 (Gaudron J); 494, 500, 502 (Gummow & Kirby JJ) and 519–520 (Hayne J). Prior to the High Court pronouncement in Tame and Annetts rejecting sudden shock as a precondition to the founding of a duty, courts denied recovery where psychiatric injury was caused other than by ‘shock’ in the sense described by Brennan J in Jaensch v Coffey, ibid. See for example, Campbelltown City Council v McKay (1988) 15 NSWLR 501; Anderson v Smith (1990) 101 FLR 34; Spence v Percy [1992] 2 Qd R 299; Reeve v Brisbane City Council, above n28; Chiaverini v Hockey (1993) Australian Torts Reports 62–254; Pham v Lawson, above n28; Davis v Scott (1998) 71 SASR 361. This has met with resistance from some quarters: see for example, Campbelltown City Council v McKay at 503 (Kirby P); Mullany & Handford, above n15 at 200–201; Harvey Teff, ‘The Requirement of “Sudden Shock” in Liability for Negligently Inflicted Psychiatric Damage’ (1996) 4 Tort Law Review 44 at 55–56; Des Butler, ‘A “Kind of Damage”: Removing the “Shock” From “Nervous Shock”’ (1997) 5 Torts Law Journal 255 at 269–275.

52 Morgan v Tame (2000) 49 NSWLR 21 at 49 (Mason P).

53 See Brennan J in Jaensch v Coffey, above n7 at 567: ‘If mere knowledge of a phenomenon sufficed, the bearers of sad tidings … might be held liable’, and Lord Oliver in Alcock v Chief Constable of South Yorkshire Police, above n30 at 416 who expressed the concern that were it not for sudden shock the law would have to permit recovery for psychiatric illness induced by the prolonged aftercare of primary victims and for pathological grief following bereavement. See also Teff, above n51 at 45–46, 47.

54 Teff, id at 55.

55 Campbelltown City Council v Mackay, id at 503 (Kirby P); Frost v Chief Constable of South Yorkshire Police [1998] QB 254 at 271 (Henry LJ); Butler, above n51 at 270–271. The Tame case provides an example of how psychiatric disorders can be sustained without sudden shock: see Tame and Annetts, above n3 at 500 (Gummow & Kirby JJ). In South Africa, the rule has been rejected by the Court of Appeal: Barnard v Santam Bpk, above n40, and in Singapore, the High Court has also rejected rule: Pang Koi Fa v Lim Djoe Phing [1993] 3 SLR 317.

56 Campbelltown City Council v Mackay, id at 503–504.

57 Butler, above n51 at 271–272.

s32(2) as this suggests psychiatric illness is generally less foreseeable absent proof of sudden shock.

Nonetheless, perhaps the most regrettable feature of s32, and indeed, the entire Part 3, is the provision that no duty of care is owed unless a person of normal fortitude might have suffered psychiatric illness in the circumstances.59

E. Normal Fortitude

By circumscribing recovery to those plaintiffs whose illness was foreseeable to a person of normal fortitude, save those instances where the defendant knew or ought to have known of the susceptibilities of the plaintiff,60 the legislation has abrogated the application of the common law. In Tame and Annetts,61 the majority did away with normal fortitude as an essential element in the test of foreseeability.62 Arguably, this is the preferable formulation.

The minority in Tame and Annetts63 expressed the view that normal fortitude ought normally be retained as a precondition because it forms an inextricable part of the notion of reasonableness in reasonable foreseeability.64 According to McHugh J, normal fortitude ‘strikes a fair balance between victim and tortfeasor’, precluding liability for unforeseeable responses and respecting ‘the right of citizens … not to have their freedom of action and communication unreasonably burdened.’65

The problem with this reasoning is that it is premised on the assumption that there exists a normal standard of susceptibility to psychiatric illness. In fact, as medically recognised,66 ‘[t]he community is not formed of normal citizens, with all those who are less susceptible or more susceptible to stress to be regarded as extraordinary.’67 The discussion concerning the difficulty of differentiating recognisable psychiatric illness and mere emotional grief also applies here. Hence

59 Section 32(1) provides that the defendant does not owe a duty of care to the plaintiff to take care not to cause the plaintiff mental harm unless the defendant ought to have foreseen that a person of normal fortitude might, in the circumstances of the case, suffer a recognised psychiatric illness if reasonable care were not taken.

60 Section 32(4).

61 Above n3 at 455–456 (Gleeson CJ); 464 (Gaudron J) and 494, 497, 498 (Gummow & Kirby JJ).

62 While Jaensch v Coffey, above n7 cannot be taken as expressly requiring a plaintiff to prove that psychiatric injury was foreseeable to a person of normal fortitude: see at 556 (Gibbs CJ); 568 (Brennan J); 557 (Murphy J); 609–610 (Deane J) and 613 (Dawson J), lower courts have generally interpreted the case as requiring that psychiatric injury be foreseeable to a person of normal fortitude: see for example, Stergiou v Stergiou (1987) 4 MVR 435; Miller v Royal Derwent Hospital Board of Management (1992) Australian Torts Reports 81–175; Wodrow v Commonwealth (1993) Australian Torts Reports 81–260.

63 Above n3

64 Id at 475–476 (McHugh J); 519, 524 (Hayne J) and 534 (Callinan J).

65 Id at 476.


67 Chadwick v British Railways Board, above n32 at 992 (Waller J). See also Mount Isa Mines Ltd v Pusey, above n14 at 405–406 (Windel Shar J).
application of the rule necessarily involves value judgments being made by judges as to the normalcy of the plaintiff’s reaction.68 This makes the outcome of cases dependent on ‘ordinary reaction’ predisposed to arbitrary outcomes.69 The normality of the Annetts’ reaction to their son’s disappearance and death, as perceived by the Court of Appeal70 rejected as incorrect by the High Court, bears testimony to this proposition, as does the enactment of the *Law Reform (Miscellaneous Provisions) Act 1944* (NSW) in response to *Chester v Waverley Corporation*.71

Accordingly, to bar recovery unless injury was foreseeable to a person of normal fortitude is medically unsustainable, artificial and productive of anomalies. This makes the legislative reintroduction of this rule as indispensable to the test of foreseeability highly problematic.

It is exacerbated by the fact it applies to all plaintiffs (subject to s32(4)), thereby representing a significant reversal of the common law position. Judgments in *Tame and Annetts*72 indicate that where there is a pre-existing relationship between plaintiff and defendant, normal fortitude is irrelevant to foreseeability.73 The Act, however, mandates the application of normal fortitude regardless of whether there exists a pre-existing relationship between the parties.74 The inversion of the common law is thus pervasive. In this regard, it should also be noted that analysis of the minority reasoning in *Tame and Annetts*75 reveals that, were it not for the undemanding test of foreseeability – not far-fetched or fanciful – all judges would apparently have concurred as to the relevance of normal fortitude.76

4. **Conclusion**

Evidently therefore, motivated by concerns that the tort system is geared too heavily in favour of plaintiffs, Part 3 of the Act now restricts plaintiffs in NSW from the benefit of establishing their cause of action free from some illogical and artificial policy limitations. While it is welcome that the legislation has left the relevance of sudden shock to reflect the common law, since the rule seems indefensible medically, conceptually or on policy grounds, regrettably the legislation retains the need for recognisable psychiatric injury, a term that cannot

---

68 Butler, above n66 at 128–129.
69 Ibid.
70 *Annetts v Australian Stations Pty Ltd*, above n30 at 55–56 (Ipp J).
71 Above n14. In this case the High Court held that a mother of normal susceptibility witnessing the body of her dead son being extracted from a water-filled trench would not suffer psychiatric illness.
72 Above n3.
73 Id at 466, 471, 481–483 (McHugh J); 508 (Gummow & Kirby JJ) and 534 (Callinan J); but see 521 (Hayne J).
74 Section 32(1), (2)(d).
75 Above n3.
76 As noted above, the legislation has provided in s5B that only risks ‘not insignificant’ are foreseeable.
be defined with sufficient precision and fairly applied. The Act also reinstates the notion of direct perception and unaided senses, which as has been shown, is devoid of rational bases so as to be unsupportable on considerations of policy. Moreover, while duly recognising the foreseeability of close relatives suffering psychiatric illness on learning of a traumatic event involving family, the legislation goes an undesirable step further to define which members of family can recover. The effect of this is to arbitrarily preclude some genuine claims.

Probably most lamentable of all, is the legislation’s returning the concept of normal fortitude to the foreseeability test. It is highly regrettable because it affects virtually all claims for psychiatric injury, and the concept’s elusiveness means its application predisposes plaintiffs to subjectivity and potentially arbitrary outcomes. The required application of normal fortitude highlights and compounds the legislature’s arguably grievous attempt to curtail the positive and continuing development of the common law towards a more rational and principled approach to the law of negligently inflicted psychiatric injury.