articles

After Objectivity: Expert Evidence and Procedural Reform
Gary Edmond
131

Burdens and Standards in Civil Litigation
CR Williams
165

Mental Illness, Rationality, and Criminal Responsibility (Tropes of Insanity and Related Defences)
Steven T Yannoulidis
189

Blue Asbestos and Golden Eggs: Evaluating Bankruptcy and Class Actions as Just Responses to Mass Tort Liability
Peta Spender
223

case note

Western Australia v Ward & Ors
Kate Stoeckel
255
After Objectivity: Expert Evidence and Procedural Reform

GARY EDMOND*

Now that the “disinterested” are praised so widely, one has, perhaps not without some danger, to become conscious of what it is these people are really interested in.

— Friedrich Nietzsche, 1886

Abstract

Is objective expert evidence possible or even desirable? Recent reforms to civil procedure in England, New South Wales and the Federal Court of Australia all seem to be predicated upon the desirability of obtaining (more) objective expert evidence. This article examines several of the reforms, including: (a) the creation of a paramount duty to the court; (b) pre-trial expert conferences; and (c) the use of court-appointed experts. Contending that the model(s) of expertise underpinning each of the reforms is highly simplistic, it endeavours to anticipate some of the consequences of grafting these procedural reforms onto our adversarial system. The article discusses how the reforms threaten to surreptitiously: raise admissibility standards; contribute to the transformation of the judicial role; privilege repeat litigants; and increase the vulnerability of the judiciary to exogenous criticism. The article also examines the concept ‘junk science’ and questions the empirical grounding of assertions about a crisis associated with expert opinion evidence. That is, in addition to questioning the efficacy of the reforms, it challenges the empirical basis apparently motivating them. Finally, by treating the reforms and their justifications ironically, the article illustrates how objectivity is not some essential quality inhering in knowledge or practice but primarily a representational accomplishment.

* Faculty of Law, University of New South Wales, Sydney 2052, Australia and Visitor, Law Program, Research School of Social Sciences, The Australian National University. e-mail: g.edmond@unsw.edu.au.
1. **The Limits of Objectivity**

... impartiality is of paramount importance.¹

... it is essential that there be a duty of objectivity, ...²

“How then is objective expert assistance to be obtained? The answer has to be: by modification of the adversarial system.”³

“A hired gun” philosophy will become a thing of the past.⁴

---


The new regimes for the admission of expert evidence in England and Wales, the Federal Court of Australia and the NSW Supreme Court are all premised upon the possibility and desirability of obtaining (more) objective expert opinion evidence. This article aims to consider whether such aspirations are realistic, whether the new procedures are appropriate and whether they will produce the anticipated results.

In their analysis of fact-finding and the assessment of expert evidence, advocates for procedural reform typically conjure simplistic images of objectivity. In conjunction with concerns about the operation of the civil justice system — usually expressed in terms of an evidentiary (or liability) crisis — apparent derogations from these images are then employed to warrant a series of reforms. In the following discussion it is my intention to adopt an alternative course. Instead of recommending procedures designed to produce demonstrably objective expert evidence, it will be my contention that ascriptions such as impartial, neutral or objective are not productive ways to think about expertise and expert knowledge, especially in relation to legal procedural reform.

Initially, then, it is my intention to problematise judicial images of expert objectivity — like those expressed in the epigraph set out above. Discrediting implausible models of objectivity is important, because controvertible images of expertise underlie and are actually embedded in the procedures adopted during the recent spate of

5 My comments are restricted to the treatment of expert evidence and, at least in this article, do not extend to other dimensions of recent procedural reforms such as case management or the use of ADR. In practice, the reforms may encourage increased interest in the reasonableness of expert opinion and recourse to (images of) professional standards, rather than just abstract images of expertise.

6 I accept that historically, at least from early modern times, legal commentators and reformers have sought degrees of objectivity (or impartiality) from experts. This article, therefore, is critical of not only the current reforms but also earlier practices and the models of science and expertise historically mobilised by judges. The recent reforms, however, may be particularly vulnerable to critique because of their explicit confidence in the value and possibility of obtaining objective expertise. Many of the earlier practices allowed, at least in theory, for more complex and perhaps more localised treatments of expertise. These earlier approaches — centred around the existence of a field, recognition of competence, the standing of particular individuals and levels of acceptance — reflected social relations and trust between professional bodies which are, perhaps of necessity, being managed somewhat differently today. For some discussion of relations between knowledge, authority and trust in particular historical settings, consider: Steven Shapin, *A Social History of Truth: Civility and Science in Seventeenth-Century England* (1994) at 3–125 and Barbara Shapiro, *Probability and Certainty in Seventeenth-Century England* (1983) at 163–193.

7 During the last few decades in the academic study of the history, philosophy, sociology and anthropology of science there have been very few attempts to locate or defend objectivity — either as an analytical conceit or adequate description of scientific practice. Consequently, it would seem to be the responsibility of proponents to prove the possibility and value of (obtaining) objective knowledge rather than for me to repeat, *ad nauseum*, the theoretical and practical problems which are almost universally acknowledged within the specialist fields. Of interest, in the Australian Law Reform Commission, *Review of the Federal Civil Justice System*, Discussion Paper 62 (1999) at 13.68, a highly qualified image of ‘partisanship’ is offered: ‘Where an expert witness is briefed and remunerated by one side, it is *often assumed* that the expert is *likely to exhibit a tendency* to give evidence which favours that side.’ (italics added)
reforms. Alternative, and perhaps more sophisticated, approaches to expertise might encourage different types of procedural innovation or enable the civil justice system to be conceptualised in ways that do not invariably lead to perceptions of a crisis.

Representations of objectivity vary considerably. Notwithstanding this variation, in descriptions of science, scientific practice, and explanations for the historical success of science, objectivity, as some kind of essential quality, regularly assumes a prominent position. Typically, objectivity is equated with qualities such as independence, impartiality and neutrality. Good science, so this story goes, derives its authority from being evidence-based, efficacious, communal, critical and driven by a powerful method. These characteristics, which are often seen as dimensions of scientific (or mechanical) objectivity, purportedly function to liberate science from a range of contaminants, such as subjectivity, personal interests, partisanship, fraud, speculation, bias, gratuitous assumptions and so forth. The appeal of such images of objectivity and method — to regulators, governments and judges involved in practical decision making around complex subjects — is fairly easy to comprehend.

We should not mistake these myths about science for the far more complicated circumstances which have historically shaped and continue to shape scientific practice and the production of knowledge. The images of objectivity discussed in the previous paragraph are highly idealised versions which displace a range of real-world influences and concerns profoundly influencing the contemporary sciences. Once we recognise, as we inevitably must, that the modern sciences are exposed to: social and economic pressures; institutional politics; diverse funding arrangements; shifting hierarchies and reward structures; ethical considerations;
competition — whether financial, personal or disciplinary; a range of techniques, instruments and methods; different levels of relevance and potential application; complex relations with other professions, through activities such as patenting; sensitivities to public concerns, especially around risk; changing public perceptions and levels of trust, and differing employment opportunities, then appeals to some extra-social image of objectivity become untenable and expert knowledge becomes more complex and inescapably political. All of the factors shaping the practices and production of knowledge raise very serious implications for understanding the contemporary sciences as well as the utility of simplistic images of objectivity and method.

Exceptionally, (a crude image of) objectivity might provide some assistance in making an evaluation about some particular expertise: such as where an expert tries to testify in an area in which they obviously lack skill, training or experience or where they possess a clear personal or pecuniary interest. In the normal course of affairs, however, such conspicuous displays of bias or impropriety may be infrequent and of limited utility in actually determining the value of the evidence. Even if such cases were common in legal settings, they could be effectively managed through existing rules and traditions of practice. More commonly, and more problematically, when experts disagree, recourse to objectivity, method(s), particular norms or the ((proper) value of the) evidence will not automatically command general assent or resolve a dispute. This is not to suggest that claims about objectivity, method(s), practices and interpretations will not form part of any given dispute. Often they will be actively contested. Rather, it is intended to suggest that for the purposes of decision making and the resolution of disputes, in the majority of cases appeals to objectivity — or idealised images of methods and attitudes — will be of limited practical value.

---


13 In this example, concerned with attributions of interest, images of fairness are integrated into assessments of value.

14 Even if, for good reasons, we typically use interests (or agency) to interpret knowledge claims, this does not mean that interested (implicitly tainted) knowledge is automatically unreliable. Interestingly, several historians, philosophers and sociologists of science have stressed the importance, to successful scientific practice, of qualities such as: perseverance, obduracy, secrecy, passion, commitment, competitiveness and faith. See, for example: Ian Mitroff, The Subjective Side of Science (1974); Michael Mulkay, ‘Interpretation and the Use of Rules: The Case of the Norms of Science’ in Thomas Gieryn (ed), Science and Social Structure: A Festschrift for RK Merton (1980) at 111–125. For a recent critical overview of ‘agency’ see: Barry Barnes, Understanding Agency: Social Theory and Responsible Action (2000).
These observations find considerable support in research from the history and sociology of science. In an analysis of objectivity and scientific controversy, the historian of medicine Randall Albury observed:

matters of disagreement between scientific experts are not typically conflicts between objectivity on one side and bias on the other, but conflicts involving two rival conceptions of objectivity — that is, two different ways of assigning relevance to the available data and of interpreting their meaning … The question of objectivity, then, as it relates to the problem of conflicting advice from scientific experts on matters of social importance, is not properly a question of deciding in the abstract which expert is more objective. It is a concrete question of which expert’s version of objectivity is to be preferred.\(^{16}\)

Controversy, as Albury explains, rarely involves dichotomised contests between identifiably objective and biased positions. (Although protagonists may characterise controversies in this way.)\(^{17}\) In the absence of an agreed Archimedean point, what is considered objective depends on the stance, commitments and assumptions of the observer (or adjudicator).\(^{18}\)

On the basis of his study of replication and scientific controversy, the sociologist Harry Collins suggests that what will be counted as the correct practice, method(s), interpretation(s), (competent use of) equipment and to some extent attitudes, which are often equated with objectivity, or objective performance, are rarely resolved before a controversy has been ‘closed’. During the controversy — that is, before closure — these areas of practice and commitment are frequently contested domains. It is only after the controversy has been settled or closed, however temporarily, that we know what counts as objective scientific practice and, inseparably, what nature is like.\(^{19}\) Not coincidentally, the approaches adopted by the side that is eventually seen to be correct are usually presented — and accepted — as objective.\(^{20}\) The winners are (or become) those who use the right methods, and implicitly, possess the correct (or most accurate) interpretation of


\(^{18}\) For an introductory discussion, see: Alan Chalmers, *What is this Thing Called Science?* (1982) at 1–37.

similarly, in legal settings judges usually accept evidentiary packages and the images of objectivity associated with them. where judges decide between one or more competing expert opinions, images of objectivity — also extending to images of proper practices, method(s), assumptions, equipment, interpretation(s), attitudes and even demeanour — are invariably and inextricably linked to the evidence (and expert) conceived as most reliable. conversely, the evidence which is not preferred is characterised as deficient: not objective or insufficiently certain or reliable to meet the requisite burden of proof. scientific controversies, both inside and outside the courtroom, are frequently resolved in a manner where one side is portrayed as basically correct (and this is presented as good science) and one side incorrect (and this is presented as bad science). in the vast majority of cases, appeals to objectivity, or greater objectivity, will offer little assistance in distinguishing between interpretations, how evidence should be understood and represented, what will count as correct practice, appropriate methods and so forth. rather, issues of objectivity and professional propriety depend on where a judge stands on particular issues, the weight of the evidence, the legal implications, concerns about consistency, ideas of justice and the ability to produce a plausible judgment. where judges endeavour to use objectivity in rationalising a decision they need to manage, usually through emphasis and omission, the social constituents of expert practice. ascriptions of objectivity, therefore, should be understood as the outcome of decision making. they concern the stance, and/or the way that practices and opinions conform (and are enrolled and linked) to expectations, other values and practices which provide the specific local content to particular images of objectivity.

from this perspective, images of objectivity and methodological propriety are integral to actual controversy rather than entities existing outside controversies waiting to be enlisted to assist resolution. if these resources were independently available, they would be used prescriptively to guide expert practice and we would have far fewer expert disagreements. objectivity, therefore, may not be a particularly fertile way of thinking about expert evidence or resolving conflicting

---

20 this is how non-specialists tend to understand the resolution of expert controversy, even if specialists often have quite different impressions of the stability of particular closures, see: harry collins & robert evans, ‘the third wave of science studies: studies of expertise and experience’ (2002) 32 social studies of science 235 at 242–251.
22 this asymmetrical approach, sometimes described as ‘sociology of error’, is discussed in: gary edmond, ‘constructing miscarriages of justice: misunderstanding scientific evidence in high profile criminal appeals’ (2002) 22(1) oxford journal of legal studies 53. see, also, the description of flip-flop thinking developed by: harry collins & trevor pinch, the golem: what everyone should know about science (1993; 2nd ed, 1998) at 142 and the golem at large: what you should know about technology (2002) at 155.
24 edmond, above n17.
25 rule following and the intransitivity of relations of sameness can be problems in the sciences as well as in law. see mary hesse, the structure of scientific inference (1974).
expert opinions. Each side in a controversy will describe its own practices as objective and methodologically rigorous. What gets counted as objectivity for rationalisation and accounting purposes, however, is an artefact rather than a cause of the particular resolution. Objectivity is a way of representing practice and behaviour, not of going about it. It is a quality associated with expertise rather than a quality inhering in it. Commitment to objectivity, as most judges are no doubt acutely aware, is hardly a prescription for going on in the vast majority of professional activities. For both judges and scientists, trying to be objective reveals very little about what to do in specific situations; especially when confronted with choices or discretions. In consequence, embracing objectivity as a guide to practice and decision making is likely to produce inconsistent results, require considerable rhetorical finesse and be vulnerable to a range of alternative interpretations.

Limitations to the value of objectivity, especially as a means of making useful practical distinctions in areas of technical controversy, have led some particularly strident critics to wonder about its strategic use and to question its ideological value:

The impossibility of objectivity and the consequent irrelevance of notions of bias (based as they are upon an assumption that non-bias is possible) should be clear..., but that should not blind us to the ideological role that the concept of “objectivity” plays.

While this approach to expertise and objectivity warrants serious consideration, the arguments in this article are not entirely dependent upon the adoption of such radical formulations.

In sum, these observations lead to two basic conclusions. The first, which has found widespread support in recent historiography of science, is that the success of the modern sciences is not a consequence of some special objectivity. Instead, a range of more complex social, political and institutional relations, ideological commitments, shared interests, enrolments and alliances are routinely used to

27 John Fiske, Television Culture (1989) at 288. Fiske is an influential cultural studies scholar rather than a sociologist of science or knowledge. Nevertheless, he expresses concerns, shared by many academics in the social sciences and humanities, about the value and use of objectivity more generally.
28 To some extent, in this article, I am playing devil’s advocate. In part, because there has been such limited critical discussion of the reforms and also because proponents of procedural reform have been so inattentive to legal scholarship and research emanating from other fields. Even if critics do not agree with these approaches to objectivity, that does not relieve them from accounting for the discussion in Sections 1 and 3. There is little in Section 3 which necessitates the adoption of more radical epistemic framing. Significantly, even the work of philosophers with (more) realist sensitivities resonates with many of the issues raised there. See for example: Schwartz, above n15; Haack, above n9.
explain the rise of the modern sciences. The second observation, more germane to my purposes, is that objectivity is not a useful way of practically — as opposed to rhetorically — classifying or distinguishing expert knowledges. This entails quite serious implications for reforms designed ostensibly to produce more objective knowledge in legal settings.

2. The New Rules and Procedures

Under the reforms in New South Wales, the new rules will encourage an economy in the use of experts, and a less adversarial culture …

Initially, it is my intention to describe a few of the reforms to procedure in the New South Wales Supreme Court. (These are substantially similar to those previously adopted by the Federal Court of Australia.) This article will restrict itself to an examination of three: (a) the general duty to the court; (b) joint expert conferences; and, (c) court-appointed experts. For the purposes of the following discussion it is important to consider both the model(s) of science (or expertise) underlying the reforms in combination with how the new procedures might actually impact upon practices identified as egregious. To varying degrees, each of the procedures considered below presupposes the possibility and utility of making use of a simplistic image of objectivity and scientific practice.

A. General Duty to the Court: The Code (Schedule K)

The centrepiece of the new procedural framework is the emphasis placed on the expert’s duty to the court. In an adversarial system, like ours, this emphasis is characterised as a significant conceptual shift. That shift is captured in the ‘Expert Witness Code of Conduct’:

---


31 Abadee, ‘New Millennium’ above n4 at 1. The tension between the potential value of expert opinion and the practical need to ration the number of experts is rarely developed.

32 I have deliberately excluded hot-tubbing from this discussion. In part, because as a forensic technique hot-tubbing retains the use of multiple (party-based) experts, is integrated into the trial process, exists on the public record and leaves experts open to cross-examination. Notwithstanding some conceptual limitations, as a procedural reform hot-tubbing is probably a more politically salient and pragmatic response to expert disagreement which deserves further study and discussion. See Justice Heerey, ‘Expert Evidence: The Australian Experience’ (Jan–Feb 2002) Bar Review 166 at 170 and Justice Heerey, ‘Expert Evidence in Intellectual Property Cases’ (1998) 9 AIPJ 92. Recent examples include: Australian Competition and Consumer Commission v Boral Ltd [1999] FCA 1318; Coonawarra Penola Wine Industry Association Inc & Others and Geographical Indications Committee [2001] AATA 844.
General Duty to the Court

2. An expert witness has an overriding duty to assist the Court impartially on matters relevant to the expert’s area of expertise.

3. An expert witness’s paramount duty is to the Court and not to the person retaining the expert.

4. An expert witness is not an advocate for a party. (Italics added)

The Code imposes upon the expert witness an ‘overriding duty to assist the Court impartially’ (2). The expert is not to adopt the role of the advocate (4) but, instead, owes a paramount duty to the court (3). Notwithstanding that expert witnesses have, since early modern times, owed a duty to assist the court impartially, the imposition of an explicit obligation is described as a central and innovative component in the new procedural regime.

The aim of this reform is to reduce partiality (and increase objectivity) by explaining the proper orientation of the expert witness and by inaugurating a regime where the expert might be subject to sanctions — such as having evidence excluded or being charged with professional misconduct — for non-compliance.

B. Joint Conferences

The NSW Supreme Court Rules provide for expert conferences or convocations of experts beyond the courtroom under the following conditions:

Conference between experts

36.13CA. (1) The Court may, on application by a party or of its own motion, direct expert witnesses to:

(a) confer and may specify the matters on which they are to confer;
(b) endeavour to reach agreement on outstanding matters; and
(c) provide the Court with a joint report specifying matters agreed and matters not agreed and the reasons for any non-agreement.

(2) An expert so directed may apply to the Court for further directions.

(3) The Court may direct that such conference be held with or without the attendance of the legal representatives of the parties affected, or with or without the attendance of legal representatives at the option of the parties respectively.

---


(4) The content of the conference between the expert witnesses shall not be referred to at the hearing or trial unless the parties affected agree.

(5) The parties may agree, at any time, to be bound by agreement on any specified matter. In that event, the joint report may be tendered at the trial as evidence of the matter agreed. Otherwise, the joint report may be used or tendered at the trial only in accordance with the rules of evidence and the practices of the Court.

(6) Where, pursuant to this rule, expert witnesses have conferred and have provided a joint report agreeing on any matter, a party affected may not, without leave of the Court, adduce expert evidence inconsistent with the matter agreed.  

Expert meetings are designed to bring experts together to confer in order to reach agreement or clearly specify the extent of (dis)agreement and reasons for it. Both require reasons (1)(c). Such conferences may be held with or without legal representation (3).  

The parties may agree to be bound by mutually accepted ‘matters of fact’ (5), but are bound to the extent that the experts actually agree (6). Where the experts concur in a written report, a party requires the leave of the court to adduce additional expert evidence (6).

Expert conferences are premised on the notion that released from traditional adversarial legal processes, experts will reach consensus or narrow the extent of their disagreement. Sometimes commentators suggest that disagreements between experts in legal settings are largely apparent, or epiphenomenal. In this vein Justice Wood provides a fairly typical justification for expert conferences: ‘when experts need to justify their opinions to fellow experts, extreme views are usually moderated, bias or adherence to junk science being quickly apparent and abandoned …’.  

This particular procedure presupposes that legal processes cause or accentuate expert disagreement, and that parties should be bound on matters where experts agree. There is, in addition, an untested assumption which implies that where expert witnesses agree the consensus is both reliable and objective.


36 Wood, ‘Expert Witnesses’, above n2 at 3, indicates that he would prefer lawyers to attend. But compare the use of an expert as a referee chairperson, in Triden Properties Ltd v Capita Financial Group Ltd (1993) 30 NSWLR 403 at 408–409, where lawyers were deliberately excluded. In a much earlier case, Cement Linings Ltd v Rocla Ltd (1939) SR (NSW) 491 at 494–495, Justice Nicholas did not even disclose the expert’s report.

C. Court-Appointed Experts

The Supreme Court Rules also enable the court to appoint an expert of its own volition or on the application of one or more parties.\(^\text{38}\)

**Court Appointed Expert and Assistance to the Court**

**Selection and appointment**

39.1. (1) Where a question for an expert witness arises in any proceedings the Court may, at any stage of proceedings, on application by a party or of its own motion, after hearing any party affected who wishes to be heard:

(a) appoint an expert … to inquire into and report upon the question;

(b) authorise the expert to inquire into and report upon any facts relevant to the inquiry and report on the question;

(c) direct the expert to make a further or supplemental report or inquiry and report; and

(d) give such instructions (including provision concerning any examination, inspection, experiment or test) as the Court thinks fit relating to any inquiry or report of the expert.

(2) The Court may appoint as the expert a person selected by the parties affected or a person selected by the Court or selected in a manner directed by the Court.

\(...\)

39.5. Any party affected may cross-examine the expert and the expert shall attend court for examination or cross-examination if so requested on reasonable notice by the registrar or by a party affected.

39.6. Where an expert has been appointed pursuant to this Part in relation to a question arising in the proceedings, the Court may limit the number of other experts whose evidence may be adduced on that question.\(^\text{39}\)

Part 39 enables the court to appoint an expert of its own choosing (1.(1)) to inquire into and report on a question or issue relevant to the proceedings (1.(1)(a), (b), (c)


& (d)). It also empowers a judge to limit the number of experts called on an issue (6). The rules provide little indication of how objective expert witnesses should be identified, selected or the conditions of their tenure (1.2)).

Once again, some of these powers pre-existed the reforms. The power to appoint an expert was, arguably, available through the inherent jurisdiction of Anglo-American courts. Historically, however, judges in adversarial jurisdictions have been reluctant to appoint experts unilaterally.

The remainder of this article endeavours to examine some of the assumptions apparently motivating the procedural reforms as well as some of their possible effects.

3. The Politics of Procedure

Before embarking on a somewhat critical analysis of the new procedures it is my intention to consider two factors which seem to have influenced their character and implementation. Ubiquitous in discussions of expert evidence and expert procedure, the two factors are: first, the manner in which expert evidence is routinely characterised as a social (or legal) problem; and, second, the prevalence of a phenomenon characterised as junk science.

First, several commentators and judges in Australia, and elsewhere, have portrayed the use of expert opinion evidence as a serious social problem. The empirical evidence for such a portrait is quite limited, rendering the claim fairly contentious. Each of the NSW judges cited in the epigraph to this article (Section 1) supported their concerns about partisanship and the need for procedural reforms through reference to research sponsored by the AIJA. That research, a quantitative survey of judicial attitudes, is used alongside personal experience as evidence for the existence of a litigation crisis caused by expert partisanship, delays and high costs.

According to the Report of the survey published by the AIJA, when asked 35% of judges intimated that lack of expert objectivity was the most serious problem with expert evidence. Some commentators suggest that the AIJA survey is conservative in its estimate of the problem. Justice Sperling, for
example, thinks that: ‘It is likely, therefore, that the incidence of bias as assessed by surveyed judges in the Freckelton report is an under-estimate.’

For Justice Abadee, the data from the AIJA survey was supplemented with authority from other jurisdictions: ‘Lord Woolf expressed the view that the expert’s role should be that of an independent adviser to the court, “with lack of objectivity” being a serious problem.’

Notwithstanding recourse to such authority there is very little actual evidence — beyond the anecdotal — of a serious or widespread legal problem.

Quite problematically, reference to the AIJA survey is used to equate judicial opinion about the role of experts and perceived problems with expert evidence with how the world actually is. Judicial perspectives — we might say judicial opinion — is privileged, indeed reified. While it may be reasonable to contend that judges have an important perspective on the operation of the civil justice system, it is not (methodologically or politically) appropriate to simply and uncritically substitute judicial perspectives for the real.

Judges are not — in any simple sense


46 Sperling, above n3 at 1. Wood adopts a more conservative approach.

objective commentators. Without intending to suggest that judicial views are homogeneous, judges are keenly sensitive to — as the promotion of procedural reform seems to indicate — managing resources; case loads; the disposition of cases; the reliability of decisions; explaining decisions; the implications of decisions; limiting appeals; their own credibility, and the standing of the profession and legal institutions.

In addition, we should not forget that the vast majority of cases, claims and actions are never litigated. Most are dropped or settled. Routinely, judges are exposed to only the very tip of the litigation iceberg. Consequently, much of the evidence presented in support of a crisis is derived by privileging the perspectives of special groups and straining the data from a methodologically questionable survey. Together they facilitate a form of ontological gerrymandering (or, more colloquially, judicial bootstrapping). Given that representations of ‘the real’ are often actively contested should we accept the institutionally sensitised (subjective) opinions of judges over (or separate from) those of lawyers, different types of expert, parties, peak professional bodies and various publics?

It is perhaps not surprising to find that alternative perspectives on the state of the civil justice system abound. Three examples will serve our purposes. In contrast to the AIJA survey and its dominant readings, the ALRC draft report on Experts presents a more circumspect portrait of the Australian litigation landscape:

Any attempt to assess the advantages and disadvantages of the present use of expert evidence and expert witnesses is made difficult by the lack of empirical information on the nature and extent of the problems that may exist. (Italics added.)

50 Even if the AIJA survey is more circumspect, in practice it tends to be read — even by some of its authors — in a way that simply adopts (or endorses) the judicial perspectives. See, for example, Australian Law Reform Commission, *Review of the Federal Civil Justice System*, above n7 at 13.21, 13.22 and 13.72. Such tendencies are also evident in US surveys of judges, see: Sophia Gatowski et al, ‘Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a post-Daubert World’ (2001) 25 Law & Human Behavior 433.
51 Australian Law Reform Commission, ‘Approaches to Reform’ in “Experts”, ch 2, above n37. See also: Australian Law Reform Commission, *Review of the Federal Civil Justice System*, above n7 at 13.2, 13.6. ‘Any attempt to assess the advantages and disadvantages of the present use of expert evidence and expert witnesses is made difficult by a lack of empirical information, including how much time and money is spent on adducing expert evidence.’ (Italics added) at 13.6 of the ALRC Report. Though, we should be sensitive to the observation that recourse to uncertainty is itself a political (and/or rhetorical) move, see: Brian Campbell, ‘Uncertainty as Symbolic Action in Disputes Among Experts’ (1985) 15 Social Studies of Science 429.
In addition, another study, with similar methodological limitations to the AIJA study of judicial attitudes, surveyed forensic scientists about their impressions of judicial understanding of their evidence and its reception in courts and judgments. It is perhaps unremarkable to observe that the attitudes of judges (taken from the AIJA study) are not entirely consistent with the attitudes of a limited sample of state-employed forensic scientists.\(^\text{52}\) It would seem that different social and professional groups maintain different views about expertise in our legal system and the need for reform. We can imagine that other stakeholders, such as plaintiff and defence lawyers, might also interpret the new procedures somewhat differently.\(^\text{53}\) The third example is drawn from the US. There, perhaps surprisingly, several eminent evidence scholars and social scientists have challenged the empirical grounding of assertions about the extent of serious problems associated with the use of expert opinion evidence.\(^\text{54}\) Contests over admissibility standards, the roles played by experts and expertise and the existence of charlatans and junk science are a feature of broader cultural struggles around the shape and operation of our legal system, especially the extent of civil liability.\(^\text{55}\)

Concerns about a litigation crisis are closely linked to anxieties about junk science. During the last decade ‘junk science’ has become a conspicuous feature in judicial and legal discourses on expertise, especially in the US.\(^\text{56}\) Having criticised the concept elsewhere, at length, it is not my intention to belabour those concerns here.\(^\text{57}\) Nevertheless, I should reiterate that however it is defined, ‘junk science’ offers little assistance in attempts to understand the roles played by expertise within the legal system. Predominantly, this is because boundaries


\(^{54}\) See the amicus curiae briefs submitted in Kumho Tire Co. v Carmichael 119 S Ct 1197 (1999). In particular the: Brief of Margaret A Berger, Edward J Inwinkelried, & Stephen A Saltzburg as Amicus Curiae in Support of Respondents and Brief Amici Curiae of Neil Vidmar, Richard O Lempert, Shari Seidman Diamond, Valerie P Hans, Stephan Landsman, Robert Maccoun, Joseph Sanders, Harmon M. Hosch, Saul Kassin, Marc Galanter, Theodore Eisenberg, Stephen Daniels, Edith Greene, Joanne Martin, Steven Penrod, James Richardson, Larry Heuer & Irwin Horowitz in Support of Respondents. See: Marc Galanter, ‘An Oil Strike in Hell: Contemporary Legends about the Civil Justice System’ (1998) 40 Arizona LR 717. The construction of public problems, such as ‘crime waves’, ‘people smuggling’, ‘jury competence’ or a ‘war on terrorism’ can have serious public policy implications. This is not to suggest the absence of a crisis, but to question how to understand and describe the situation. For a useful introduction to the public problem literature, see: Joseph Gusfield, The Culture of Public Problems (1981).

\(^{55}\) Edmond, above n8.

\(^{56}\) See Edmond, above n8.
constructed around purportedly good and bad science are less obvious and often more controversial than routinely acknowledged by proponents of the concept.\textsuperscript{58} Typically, there is an implicit suggestion that ‘junk science’ is as conspicuous to modern sensibilities as astrology, necromancy or phrenology.\textsuperscript{59} However, the limited value of ‘junk science’ as a mechanism for practical demarcation becomes evident when we consider how infrequently these types of evidence enter the courtroom.

Conceptually, ‘junk science’ is more of a polemical tool than an analytical one. Those who use the term may not encounter difficulties identifying what they believe to be ‘junk science’, but they do suffer when it comes to credibly distinguishing it from authentic scientific knowledge on the basis of consistent application of their preferred demarcation criteria.\textsuperscript{60} Often the process of classification (or demarcation) seems to incorporate overt ideological dimensions; predicated upon \textit{a priori} commitments rather than practical or technical distinctions.\textsuperscript{61} Consider, by way of example, the favourable response to the US Supreme Court’s \textit{Daubert} decision as a purported solution to the difficulties posed by psychological syndrome evidence in the work of Freckelton, Odgers and Richardson.\textsuperscript{62} The \textit{Daubert} decision, and particularly a version of the philosophy


\textsuperscript{58} Definitions and boundaries developed to describe and demarcate particular sciences are always local, temporal and contingent, see: Thomas Gieryn, \textit{Cultural Boundaries of Science: Credibility on the Line} (1999); Geoffrey Bowker & Susan Star, \textit{Sorting Things Out: Classification and its Consequences} (1999); Harriet Ritvo, \textit{The Platypus and the Mermaid and other Figments of the Classifying Imagination} (1997).

\textsuperscript{59} \textit{US v Downing} 753 F 2d 1224 at 1239 (3rd Cir 1985); \textit{Kumho} (1999), above n54 at 1175. See also Huber, \textit{Galileo’s Revenge}, above n56. For a more scholarly treatment, consider: Roger Cooter, \textit{The Cultural Meaning of Popular Science: Phrenology and the Organisation of Consent in Nineteenth-Century Britain} (1984).


\textsuperscript{61} For an influential study of attempts to distinguish the study of parapsychology from the realm of legitimate science, see: Harry Collins & Trevor Pinch, \textit{Frames of Meaning: The Social Construction of Extraordinary Science} (1982).
of Karl Popper, was mobilised because of the apparent serviceability of falsification as a means of demarcating science from non-science in addition to his well-known animosity toward much psychological and social theory. But we should not be deceived by apparently easy solutions. There is no simple identifiable essence distinguishing all so-called good science from so-called bad science. Rather than the distractions caused by pejorative labels and allegations we ought to be concerned with attempts to locate expertise that is sufficiently reliable for the purposes of legal decision making. Use of the term ‘reliable’ only begs the question — how reliable? Unfortunately, this is one of the elisions in the US Daubert and Kumho decisions as well as in our own procedural developments.

Finally, a fairly obvious limitation with the concept of ‘junk science’ concerns the inconsistency, or incoherence, in the way it is deployed. If, for example, ‘junk science’ is easily identified and a pressing problem in contemporary litigation, why then do judges experience such difficulty managing it? What prevents judges from excluding such evidence as clearly inadequate under the existing rules of evidence (Evidence Act 1995: ss55, 56, 79, 135 and 136) or affording it little weight in their assessments? Why are additional reforms required? Could it be that difficulties associated with expertise are not resolved by appeals to simplistic models of science or naïve images of objectivity? And, if ‘junk science’ is not always so easy to identify, might concerns about its prevalence be misguided or illusory? Notwithstanding its rhetorical value, the idea of ‘junk science’ is predicated upon junk philosophy and sociology of science.

At this stage, having questioned whether the reforms are actually necessary, we move to consider some of their limitations and possible effects.

First of all, the imposition of an explicit duty to the court seems unlikely to make much difference to the practice and culture of expert witnessing. Such a duty may produce limited changes such as encouraging experts to divulge slightly more information, structure their reports differently or incline them toward circumspection, but given that experts have always been under a duty to the court...
and have always testified under an oath or affirmation — to tell the ‘whole truth’ — it does not follow that clarifying the duty will produce more objective evidence (or less ‘junk science’). In addition, the imposition of a paramount duty to the court will not operate in a social vacuum. Most expert witnesses already owe a variety of professional, ethical, pecuniary and personal allegiances. In the absence of substantive sanctions (more below), expecting the imposition of an additional obligation to have manifest effects on expert practice might seem a little optimistic.

A useful illustration of some of the tensions flowing from the legal attribution of roles and duties arose in the Hindmarsh Island Bridge litigation. An anthropologist sued by developers alleging negligence in her professional performance held, simultaneously, (at least) a duty to the court, a duty to her profession, a duty to her employer, a duty to the university and their insurer, an ethical duty to the local Aboriginal peoples and was concerned to defend her reputation, dignity and assets. Confronted with conflicting expert opinions about the adequacy of the anthropologist’s performance, Justice von Doussa escaped this morass through the (re-)construction of her contract and instructions. The

67 There is some scepticism about the value of the Woolf Reforms in England: Christopher Taylor, ‘Distilling the Spirit of Woolf’ (2001) 151 New Law Journal 1035; Zuckerman (ed), Civil Justice in Crisis, above n1. Interestingly, even preliminary assessments of the effects of reform on expert evidence are ambivalent, see: Lord Chancellor’s Division, Emerging Findings: An Early Evaluation of the Civil Justice Reforms (London: HMSO 2001). In their study of court-appointed experts in US federal courts, Cecil & Willging found that, when they were used, fewer cases than (they) expected actually settled: Cecil & Willging, Court-Appointed Experts, above n41 at 13.

68 This is not to suggest that the different duties are necessarily irreconcilable. In their ‘Role and Duties of an Expert Witness in Litigation’ (1998) <http://www.mica.org.au/pdf/xwitness.pdf> (15 March 2003) the Australian Council of Professions lists three duties for members acting as expert witnesses: (1) a primary duty to the court; (2) a secondary duty to the body of knowledge and understanding from which the expertise is drawn; and (3) a third duty to the party. Many professional societies, such as the Australian Anthropology Society and Australian and New Zealand Forensic Science Society have their own charter, expectations and ethical precepts. In the UK the Civil Procedure Working Party has produced a draft ‘Code of Guidance on Expert Evidence’ which distinguishes between experts who provide advice and experts who provide reports, see: Christopher Taylor, ‘Expert Evidence — A Review of the Proposed New Code’ (2001) 151 New Law Journal 1052.

69 ALRC, Review of the Federal Civil Justice System, Discussion Paper 62 (1999) 13.18. The ALRC discussion paper summarises some of the problems (it suggests are) most frequently associated with expert evidence. (Given that in the same Review, at 13.6, the empirical basis for concern is questioned, the foundation for such assertions are somewhat perplexing). It is not clear how the procedural reforms will address concerns such as: not hearing the most expert opinions; corrupt experts; how lay judges should assess evidence; the effects of demeanour, and the confidence of the expert. Elsewhere in the discussion paper, at 13.94, the metaphors and cautious tone suggest ambivalence: ‘Experts, reminded of their duty to the Court, may be more confident to resist any suggestions from lawyers to tailor reports to secure a particular legal outcome. Lawyers, in turn, may become less likely to suggest such tactical play to their experts. This is at least the plan.’ (italics added)

anthropologist’s contractual engagement with the Aboriginal Legal Rights
Movement (ALRM), in relation to a controversial application under the Heritage
Protection Act (Cth), entitled her to act as an advocate. Some may find such an
approach unconvincing. Presumably, many of the advocates of reform would find
it most unsatisfactory. But the example does raise the question about expert
allegiances and priorities. Many potential expert witnesses will hold pre-litigation
and litigation-sensitive contractual relations with employers who are likely to be
sued. Pharmaceutical and insurance companies are obvious examples.

Expert conferences, the second of the reforms considered in Section 2, are
predicated upon the belief that litigation accentuates disagreement. Implicitly,
experts released from substantive legal and procedural constraints will be able to
broker agreement (or limit the extent of disagreement). This assumes that much
expert disagreement is the result of communication problems, distortion caused by
legal practices or the restricted orientations encouraged by the partisan
commitments of clients. We should not forget, however, that there is considerable
scientific controversy beyond legal and regulatory fora. In addition, those few
cases that eventually get to trial — the ‘tip of the iceberg’ that are not settled or
abandoned — often involve disputes in areas presented as uncertain or
controversial. Can we expect expert meetings to expose or satisfactorily resolve
what cross-examination, an oath, or duty, and extra-legal processes often
cannot?

The contention, promoted by Justice Wood and others, that when experts meet
to negotiate in their own terms they moderate their views, abandon ‘junk science’
and need to justify their positions to fellow experts, is a highly romantic view of
expert controversy. Indeed, several proposed procedural solutions to complex
forms of expert disagreements, such as science courts and expert panels, have
encountered mixed success.

In practice, sometimes experts could not even agree on the terms of their disagreement. In the case of the proposed science court, those
from the more mainstream (or orthodox) position would not engage in debate with
those holding alternative (minority) positions.

Entering into debate was perceived to cede (extra) credibility to minority opponents. The example illustrates

---

71 Consider: Carol Henderson Garcia, ‘Expert Witness Malpractice: A Solution to the Problem of
72 Chapman v Luminis [2001] FCA 1106 at paras 89–90, 101, 109, 120, 198, 278–300, 311, 313,
400.
73 See: John Abraham, ‘Distributing the Benefit of the Doubt: Scientists, Regulators, and Drug
Safety’ (1994) 19 Science, Technology & Human Values 493; Evelleen Richards, Vitamin C and
74 For a theoretical discussion of expert disagreement in a public inquiry, see: Brian Wynne,
Rationality and Ritual: the Windscale Inquiry and Nuclear Decisions in Britain (1982); and for
legal contexts, see: Roger Smith & Brian Wynne (eds), Expert Evidence: Interpreting Science
75 For some introductory literature, see: Tristram Engelhardt & Arthur Caplan (eds), Scientific
Controversies: Case Studies in the Resolution and Closure of Disputes in Science and
Technology (1987); David Mercer, Understanding Scientific/technical Controversy (1996)
Occpaper–1.html> (15 March 2003); Collins & Pinch, The Golem and The Golem at Large,
above n22.
how institutional arrangements predicated upon simplistic models of scientific knowledge, practice and professional normative order may be untenable or fail to operate in the ways anticipated. These kind of issues loom large and operate in addition to the more publicly contentious selection of experts, choice of appropriate fields and the relevance of any prior decision, agreement, publications or report for subsequent disputes or litigation.

Expert conferences, which have the potential to bind the parties, invest repeat-players (such as insurance companies) with a range of procedural advantages, which might appear superficially innocuous. Pre-trial expert conferences, ostensibly aimed to reduce the amount of expert disagreement and possibly litigation, transform the choice of expert into a more important strategic decision. Those lawyers and plaintiffs least familiar with the system may effectively dispose of their case through insufficient attention to expert selection. Such choices would be less significant if there really were recognisably objective experts. However, in their absence, parties may cede considerable advantages and even compromise their action by failing to anticipate what an expert might say or agree to. Repeat players — those with litigation experience or stables of experts — can avoid this situation, without contravening the new procedures, by

76 Prominent American socio-legal scholar Michael Saks summarises the various experiments as follows: ‘Court appointment of non-party experts is one of the most commonly recommended reforms. It also has been a resounding failure everywhere it has been tried.’ Michael Saks, ‘The Phantom of the Courthouse’ (1995) 35 Jurimetrics 233 at 240. The judges and proponents of reform seem to have paid little attention to a range of reforms attempted in other (adversarial) jurisdictions. Two conspicuous oversights are the New York Medical Expert Testimony Project and a current initiative in the US to produce panels of high quality experts under the auspices of peak scientific organisations such as the American Association for the Advancement of Science (AAAS). Consider: Impartial Medical Testimony (1956); Francis Van Dusen, ‘A United States District Judge’s View of the Impartial Medical Expert System’ (1962) 32 FRD 498; William Wick & Eric Kightling, ‘Impartial Medical Testimony Under the Federal Civil Rules’ (1967) 34 Insurance Counsel Journal 115; Lowell Myers, ‘The Battle of the Experts’: A New Approach to an Old Problem in Medical Testimony’ (1965) 44 Nebraska LR 539; Robert Winikoff, ‘Medical-Legal Screening Panels as an Alternative Approach to Medical Malpractice Claims’ (1972) 15 William and Mary LR 695 and ‘Court-Appointed Scientific Experts: A Demonstration Project of the AAAS’ at <http://www.aaas.org/spp/case/case.htm> (15 March 2003).


familiarity with particular experts, shared values, tacit understandings and continuing professional relations. All this before we even consider whether the new procedures might send many of the negotiations between parties, lawyers and experts ‘underground’.

Further, expert conferences and the use of court-appointed experts may actually alter the balance between respective parties in other ways. Experts participating in conferences may reach agreement around the least controversial aspects of their knowledges. Judges seeking to create robust and socially credible resolution to legal disputes may, understandably, tend to prefer the evidence of more credible and established scientists and disciplines. ‘Credible’, ‘established’ and ‘least controversial’ possess a range of social resonances: for example, with experts who write authoritative textbooks or are based in prestigious universities or research institutions. Given that plaintiffs bear the burden of proof any tendency to reduce disagreement and perhaps favour (what are presented as) more established or orthodox opinions may tend to disadvantage them (relative to defendants). Further, the new emphasis on the experts’ duty to the court may itself encourage circumspection and greater use of uncertainty — again potentially disadvantageous to plaintiffs. In these ways the reforms may actually transform (and possibly raise) the traditional entry requirements for expert evidence (and burdens of proof) making it harder for some groups to access substantive tort doctrine. Commitment to particular manifestations of objectivity may come at the price of making the types of expert preferred by courts more mainstream and

---


83 Enhanced discovery might also impact to the detriment of plaintiffs. If parties were required to disclose all of their communications and reports, much more would depend on the initial selection of the expert and prior knowledge of particular experts. Consider: Anthony Champagne, Daniel Shulman & Elizabeth Whitaker, ‘An Empirical Examination of the Use of Expert Witnesses in American Courts’ (1991) 31 Jurimetrics 375 at 376. For an informative and influential account of relations between law and actual social behaviour see: Stewart Macaulay, ‘Non-Contractual Relations in Business: A Preliminary Study’ (1963) 28 American Sociological Review 55. Macaulay’s article explained how doctrinal contract law exerted limited influence on the majority of business relations. Instead, business practice tended to be shaped by levels of trust, reciprocity and tolerance contingent upon continuing business relations, shared values, professional credibility, and the high cost of litigating. See also: Robert Ellickson, ‘Of Coase and Cattle: Dispute Resolution Among Neighbors in Shasta County’ (1986) 38 Stan LR 623. Similar types of informal or tacit practices will presumably emerge around the new expert evidence procedures.

84 Studies of experts appearing in the media suggest that informal relations, availability, visibility and institutional prestige are important factors in the selection of experts. Interestingly, these studies also suggest that personal eminence or affiliations with prestigious schools are often more important than specific expertise, see, for example: Dorothy Nelkin, Selling Science: How the Press Covers Science and Technology (1987). Cecil & Willing report that judges often selected experts on the basis of ‘informal networks of friends and acquaintances’ which raises concerns about ‘the extent to which such networks can be relied upon to provide skilled and neutral experts’: Cecil & Willing, Court- Appointed Experts, above n41 at 31–32.

more orthodox. These preferences may alter the traditional rights and duties between plaintiffs and defendants, manufacturers and consumers. Such a substantial policy shift might be desirable but it deserves public discussion and should not be insinuated through procedural reform predicated upon non-problematised and purportedly apolitical images of objectivity. Given that ‘objectivity’ is almost as useless as ‘junk science’ as a means of demarcating or assessing expertise there needs to be further discussion about the types of expertise that should be used in the resolution of disputes, the requisite standards of reliability and the implications of uncertainty.

In conjunction with Lord Woolf, Justices Abadee and Sperling emphasised that it is essential for experts to maintain their objectivity during litigation. Notwithstanding the intuitive appeal of ‘objectivity’ as some kind of intrinsic attribute manifested in a person(s) or their knowledge, behaviour and methods, as indicated, objectivity is better understood as a representational accomplishment shaped by the ability of those supporting or defending a panel, expert, opinion or report to manage or sustain that appearance. The appearance of objectivity is easiest to manage in the absence of controversy. In legal settings the apparent objectivity of opinions or experts will be most resilient and persuasive where the stakes in the litigation are small, where the appearance of independence and credibility are strong, consensus in related fields high and compromising interests difficult to establish. This is precisely the type of situation where settlement is likely. However, if the stakes are high, the social and financial implications considerable or where experts (or judges) can be represented to have conflicting interests and affiliations then the authority to comment, derived from an expert’s imputed objectivity, can be compromised. This can impact upon the ability of a legal institution to reach a socially plausible closure, as well as maintain its own claim to political (or epistemic) credibility. As they stand, the procedures will offer least assistance in the circumstances where they would be most valuable.

The new procedures may raise problems of social legitimacy and introduce new difficulties as court-appointed experts and experts involved in conferences become tarnished by the way their evidence and opinions are perceived and used by judges, plaintiffs and defendants, particularly if the same experts are used over and over (especially by the same judge). The credibility of judicial decision making may become more intimately related to the authority and legitimacy of specific experts and their opinions. This will extend beyond the conventional adversarial expectation that the judge will decide between the experts produced by the parties. In the actual selection of one or more experts, judges may have to make

86 Reform and proposals for reform (as well as critique) as Leubsdorf observes, are always part of a broader vision of social justice: J Leubsdorf, ‘The Myth of Civil Procedure Reform’ in Zuckerman (ed), Civil Justice in Crisis, above n1, 53 at 67.
87 In a sense, the use of procedures like court-appointed experts is intended to convert all disagreements to this form.
89 There have already been cases where judicial intervention, by way of critical comment, has been successfully appealed. See, for example, Vakauta v Kelly (1989) 167 CLR 568.
potentially controversial choices between competing fields and possibly impose
some kind of (potentially dispositive) epistemic hierarchy.\footnote{In his Report and throughout the proceedings of the Royal Commission into the Chamberlain convictions, Justice Morling was very cautious about the way he treated those experts which he invited to assist the Commission. Typically they were represented as independent or neutral. In the Report those characterisations provide an important persuasive resource. See: Hon Justice T Morling, Report of the Commissioner: Royal Commission of Inquiry into Chamberlain Conviction (Darwin, Government Printer: 1987); Gary Edmond, ‘Law, Science and Narrative: Helping the “Facts” to Speak for Themselves’ (1999) 23 Southern Illinois University LR 555 at 574–583.} Elsewhere I have discussed the inconsistent judicial use (or more accurately representation) of concepts such as: communities, fields, protocols, methods and norms (even those outlined in the much vaunted Ikarian Reefer).\footnote{One example is the way epidemiological evidence was used in the Bendectin (Debendox in Australia and Britain) litigation, see: Edmond & Mercer, ‘Litigation Life’, above n85, another is the use of historical sources in the Yorta Yorta native title litigation: Members of the Yorta Yorta Aboriginal Community v Victoria [1998] FCA 1606, see: Alexander Reilly, ‘The Ghost of Truganini: Use of Historical Evidence as Proof of Native Title’ (2000) 28 Fed LR 453.} When there are few experts, or perhaps only one, these, often complex, strategic decisions will become even more important.\footnote{Edmond, ‘Judicial Representations of Scientific Evidence’, above n21 at 222–224. See: National Justice Compania Naviera SA v Prudential Assurance Co Ltd [1993] 2 Lloyd’s Law Reports 68 at 81–82, on appeal National Justice Compania Naviera SA v Prudential Assurance Co Ltd [1995] 1 Lloyd’s Law Reports (CA) 455 at 496.} It may be that a court-appointed expert, especially a very eminent one, will be permitted more latitude to offer opinions at the boundary of their legal competence. Where courts are implicated in the creation and management of expert ‘credibility’ and ‘objectivity’ they will become more vulnerable to a range of exogenous groups and institutions which might challenge their attempts to procure and maintain legitimate closure.\footnote{Further, if judges are implicated in expert selection, then accuracy and reliability may become more important dimensions of participants’ perspectives of fairness and justice. If the parties have their ability to select experts restricted then they might be far more attentive to, and potentially critical of, judicial performance. The selection and appointment of experts by judges is not a trivial modification to our adversarial procedures. See also n94.} In addition, judges will have to make decisions about what kind of actions might be sufficient to impugn any (judicially) invested objectivity.

In the US there have been criticisms about the use of so-called independent experts in expert panels investigating the role of silicone gel breast implants in a range of illnesses.\footnote{In the US Federal Court, there have been attempts to sue prosecutors for their alleged negligent selection (also described as ‘expert shopping’) and use of experts. See, for example: Mowbray v Cameron County, Texas 274 F 3d 269 (2001). See also: Robert Park, Voodoo Science: The Road from Foolishness to Fraud (2000) at 162–171.} Of interest, expert panels in different jurisdictions, ostensibly concerned with similar issues, were configured utilising different types of

---

\footnote{Different approaches to the use of experts were adopted in the following cases: In re Silicone Gel Breast Implant Prods Liab Litig (MDL 926) 887 F Supp 1455 (1995); Hall v Baxter Healthcare Corp 947 F Supp 1387 (1996). For an overview, see: Karen Resinger, ‘Court-Appointed Expert Panels: A Comparison of Two Models’ (1998) 32 Indiana LR 225. Even in Australia, court-appointed experts have been challenged on various grounds, see for example: Trade Practices Commission v Arnotts Ltd (1989) 89 ALR 131.}
expertise and incorporated into proceedings in a variety of ways. In those cases, where the litigants had access to considerable resources and the requisite motivation, panel members and judges were subjected to a range of (ultimately unsuccessful) challenges based on alleged financial and cognitive biases. If we conceptualise objectivity as a judicial construct then the use of joint or court-appointed experts may elide a range of socio-political practices which are occasionally susceptible to strategic exposure. All of these concerns operate in addition to the more regular concerns about additional costs, loss of party control, selection processes, difficulties in integrating court-appointed expert evidence into the adversarial trial as well as the authority and influence a court-appointed expert might exert upon a judge (or jury).

Notwithstanding these observations, some of the procedures designed to reduce the cost and delay associated with litigating may achieve that outcome by encouraging judges — and experts, yes experts — to make preliminary decisions which become dispositive. Through the reforms judges are encouraged to adopt more active interest in narrowing the extent of expert disagreement and deciding whether any disagreement — as opposed to agreement — is in fact genuine. Only ‘genuine’ scientific disagreement will be relevant to the legal proceedings. These types of decision were already being made by judges, but the new procedures and the ideology underpinning them offer new institutional pressures (and means) to reduce expert disagreement and determine whether ‘full, red blooded adversarialism’ should be permitted. According to Lord Woolf, adversarialism is not suited to all cases:

There are in all areas some large, complex and strongly contested cases where the full adversarial system, including oral cross-examination of opposing experts on particular issues is the best way of producing a just result. That will apply particularly to issues on which there are several tenable schools of thought, or where the boundaries of knowledge are being extended. It does not, however, apply to all cases.

A reduction in the range of admissible expert evidence will presumably influence outcomes. The judicial location of ‘boundaries of knowledge’, ‘tenable schools’

---


98 Woolf, Access to Justice, above n1 at 13.6.

99 Id at 13.19.
and ‘strongly contested cases’ may effectively dispose of cases. These decisions are complex socio-legal choices. In the US, even though the Daubert judgment purported to inaugurate a new regime which encapsulated the liberalisation motivating the enactment of the Federal Rules of Evidence (1975), the post-Daubert landscape has been generally detrimental to plaintiffs’ experts. Federal judges, acting in the capacity of vigilant gatekeepers, have insisted upon higher standards of ‘reliability’ and frequently privileged defendants’ experts — especially those routinely employed (‘contracted’ in Justice von Doussa’s terminology) in industry.100

Reforms involving the imposition of a duty, requiring expert conferences and even the appointment of an expert by the judge are unlikely to produce more objective knowledge for two main reasons. The first results from recognising that ascriptions of objectivity are better understood as representational achievements — often post factum — rather than some kind of intrinsic essence. This means that concepts such as objectivity (or neutrality, independence, impartiality, ‘junk science’ and so on) will be difficult to operationalise as a means of demarcating expertise. Judicial representations of objectivity are not surrogates for truth, reliability or accuracy.101 The other difficulty is that these definitional problems mean that few sanctions can be effectively mobilised against bad experts (or ‘junk scientists’).102 Without a sophisticated model of objectivity or even expertise, on what grounds are judges to apply sanctions against experts who breach their obligation to the court or who are unable to achieve consensus around their opinions? How should judges determine whether reluctance to agree or narrow the grounds of disagreement at an expert conference constitutes legitimate

100 See for example Kumho, above n54. Claims about the modern liberalisation of admissibility standards should not be accepted uncritically. A range of liberal, perhaps excessively liberal (by our own standards) admissibility decisions have been watered-down in recent decades. Some of the early permissive Australian decisions include: Adelaide Stevedoring Co v Forst (1940) 64 CLR 538; Commissioner for Government Transport v Adamick (1961) 106 CLR 292; Clark v Ryan (1960) 103 CLR 486; Weal v Bottom (1966) 40 ALJR 436; and more recently: R v Gilmore [1977] 2 NSWLR 935; Murphy v The Queen (1989) 167 CLR 94. Against these relatively, and some extremely, inclusive approaches we might contrast a few more restrictive judgments such as: Hocking v Bell (1945) 71 CLR 430 and Blackstock v Foster [1958] SR (NSW) 341. Compare Alan Abadee, ‘Professional Negligence Litigation’, above n4 at 2: ‘The modern approach to expert evidence is less exclusionary than it has been in the past.’ The majority Daubert judgment, above n1, written by Justice Blackmun, was (also) characterised as a more liberal approach to the rules of evidence even though it — and the cases that followed, particularly General Electric Company v Joiner 78 F3d 524 (1997) and Kumho, above n54 — have not had that effect. We should also be attentive to differences in treatment of plaintiffs’ expert evidence and the expert evidence adduced by the state in criminal prosecutions. For some discussion of these issues see: Gary Edmond & David Mercer, ‘Conjectures and Exhumations: Citations of History, Philosophy and Sociology of Science in US Federal Courts’ (2002) 14 Law & Literature 309; Edmond, ‘Legal Engineering’, above n8.

101 The tensions between process and accuracy have long been noted: John Rawls, A Theory of Justice (1971) at 85–86. See also: Daniel Farber, Beyond all Reason: The Radical Assault on Truth in American Law (1997); Norman Levitt, Prometheus Bedevilled: Science and the Contradictions of Contemporary Culture (1999) at 211–230.

102 Sanctions also operate to punish parties for the performance of an expert.
professional differences or obduracy driven by a party’s desire for a trial? What in
the process divulges this? When is adherence to a particular ‘school of thought’
partisan and under what circumstances might it be reasonable or objective? What
can judges do when experts hold firm opinions about areas characterised as
uncertain or disagree about the extent or significance of certitude in a field? Could
experts be punished — by contempt of court proceedings, for attempting to pervert
the course of justice, or even perjury — for steadfastly holding an opinion, or even
changing their mind in relation to other opinions or fresh evidence? When should
fresh evidence or assumptions excuse or require such shifts?103

If judges were to attempt to discipline expert witnesses a range of novel and
perhaps unanticipated consequences may emerge. Some of these might create or
exacerbate serious inter-professional tensions. An instructive hypothetical
example might focus on the performance of a medical specialist in professional
negligence litigation. For the time being judges may be able to determine what
constitutes negligence in relation to disclosure and medical practice. It does not
follow, however, that professions will accept the evidentiary findings in judgments
or the concerns of judges in relation to professional misconduct in the presentation
of evidence. Were they to be established, it is quite possible that respectably
constituted medical boards of inquiry would exonerate, or even vindicate, the
professional performance of their members in their provision of evidence: even
performances subjected to severe judicial criticism. Medical concern over the
extension of legal intervention — exemplified in cases like Rogers v Whitaker —
may be avenged if (or when) medical panels are asked to review (legally)
controversial expert witness performances.104 Professionally independent, are the
medical colleges likely to accept the findings or simplifications of a judge in
relation to controversial opinion evidence? What happens to a case or the
credibility of a judge (qua fact-finder) who criticises an expert performance that is
later exonerated or, worse, endorsed by a peak professional body? In some cases,
attempts to maintain the primacy of legal institutional perspectives might appear
not only controversial but even antiquarian and possibly irrational.105 Judicial
recourse to justice may not be convincing in competition with professional claims
to truth.

Removing expertise from the dynamics of the adversarial trial and burdening
judges with additional roles such as selecting neutral experts and monitoring a
range of duties transforms judging into a more explicitly epistemological activity.
The more technocratic and evidence-based courts endeavour to become, the more
vulnerable they will be to forms of exogenous criticism. Where there are

103 Compare judicial uses of ‘fresh evidence’ in Edmond, ‘Constructing Miscarriages of Justice’,
above n22 at 80–82.
104 Rogers v Whitaker (1992) 175 CLR 479. Rogers overturned the earlier Bolam approach to
medical negligence which privileged medical opinion: Bolam v Friern Hospital [1957] 1 WLR
582; Sidaway v Bethlehem Royal Hospital [1985] AC 871.
105 This seems to be a feature of AMA approaches to medical negligence and recent calls for
statutory protections and self-regulation. See the speech by Dr Karen Phelps to Australian
SGB3MB> (15 March 2003).
potentially career-destroying judicial assessments of expert performance, we should not expect that the paramount duty to the court will produce acquiescence from experts or that judicial representations will remain unchallenged. Experts, not coincidentally, like lawyers (and judges) will endeavour to actively and creatively reinterpret and appropriate the proposed reforms. Recall the solution to the Hindmarsh Island Bridge litigation. There, Justice von Doussa circumvented the expert’s paramount duty to the court and the need to choose between the experts through recourse to contract and differentiating between advocacy and testimony. Alternatively, professional groups may endeavour to publicly censure judges and judicial performances on the basis that they, and not judges, are exclusively entitled to interpret and apply specialised expert knowledge.

These types of criticism may be hard for judges to withstand. Ultimately, the imposition of a paramount duty offers few possible sanctions. It is, in effect, a sheep in Woolf’s clothing.

Both judges and scientists routinely draw upon images of objectivity to describe their practices. There may, however, be differences in the way each conceptualise and use these images in particular situations. Even though the procedural reforms imply that by replacing the obligation to parties with a paramount duty to the court experts will, in effect, be (more) ‘objective’ this does not guarantee that (groups of) scientists and judges will maintain similar interests or share understandings or the implications of objectivity in specific instances. Groups of scientists may hold professional interests and values which are not consistent with those held by judges or the parties. For example, without alleging conspiracy, should we accept that a pre-trial conference of obstetricians will be disinterested (in the legal sense) in the professional effects and insurance implications of professional negligence litigation or in identifying the appropriate standards of obstetric practice? Alternatively, at what stage and in what ways should legally significant and ostensibly competing concerns such as justice, cost, access, efficiency, fairness, social effects, truth and delay impact upon judicial concerns about ‘objectivity’?

Lord Woolf indicated that a ‘full, “red-blooded” adversarial approach is appropriate only if questions of cost and time are put aside.’ How do these socio-legal factors affect scientists’ understanding of legal


107 The case and some of the expert opinions in this case were developed prior to the procedural reforms, but the judgment discusses the evidence in terms of the Federal Court’s new procedures.


proceedings, judicial impressions of objectivity and the way both use and represent scientific expertise? Technical legal justifications, alleged expert incompetence, and claims about fresh evidence did not prevent sustained criticism of the judiciary as a result of a number of alleged miscarriage of justice cases here and in the UK in previous decades. Is there any reason to believe that scientists (or the public) will accept *sui generis* or technical judicial resolutions in the face of controversial scientific opinions?112

Having identified a number of potential limitations to the procedural reforms (if they are actually used) we should not be surprised by their ability to produce apparently legitimate and robust legal closures using fewer experts. Changing the rules of procedure may simultaneously transform our understanding of justice and procedural fairness as effectively as any proper characterisation of the sciences, scientific practice or valuation of the evidence. However, scientific evidence and judicial representations will still be open to a range of (often competing) interpretations.113 Some of the structural means of attacking evidence available in more traditional adversarial jurisdictions may be precluded, but this is a classic example of how the epistemological robustness of the evidence will be closely linked to the social legitimacy of the institutions endorsing it.114 The new procedural rules are designed to restrict costly and protracted expert disagreement. If judges actively implement them they may have this effect. But it will not be a result of producing ‘objective’ scientific evidence or ensuring ‘proper’ conduct, although this is almost certainly how it will be portrayed. Instead it will be the consequence of excluding or displacing some types of practices, conflicts and claims from the courtroom.

A more informative question is: Will the reforms work in the ways anticipated? If we mean will the new rules produce objective knowledge — the preferred goal, despite the absence of clear elaboration — then the likelihood is low. If, however, working means that the rules will reduce the number of experts appearing in courts, reduce the length of some trials, or contain the extent of disagreement, then to some extent they probably will.115 Of course, these results will not necessarily be

---

112 The implications of various innocence projects are currently being worked-out in Australia, UK and the US.
113 Gary Edmond, ‘Science in Court: Negotiating the Meaning of a “Scientific” Experiment During a Murder Trial and Some Limits to Legal Deconstruction for the Public Understanding of Law and Science’ (1998) 20 *Syd LR* 361 at 391–400.
115 There is also the possibility that the reforms will have some of the desired effects for the wrong reasons. Sometimes desired outcomes are achieved where causal attributions are flawed. For example, early natural magicians treated people with blood disorders (which we would understand as anaemia) with iron compounds because they linked blood to the planet Mars, which was associated with war (and iron) through mythical and semiotic connections. Brian Easley, *Witch-Hunting, Magic and the New Philosophy: An Introduction to the Debates of the Scientific Revolution 1450–1750* (1980).
causally linked to the production of more reliable or socially credible opinions. Instead, a range of other, potentially more compelling sociological insights are available to explain changes to practice. For example, expert agreement in pre-trial meetings may not only be attributable to clarification of the issues. The rules are silent on subjects such as expert–expert intimidation, the relations between theory and practice, previous performances, professional relations, competence, the relevance of particular types of expertise, relations within and across fields, personal animosities and interests shared by particular expert sub-groups, such as obstetricians. Further, experts may reach agreement to avoid going to trial and confronting the rigours of cross-examination and inconvenient delays while still recovering substantial fees. The new rules imply that disagreement is caused, or at least accentuated by the legal system. There may be some basis to this claim, but the rules offer few means of resolving expert disagreement other than by foreclosing the possibility or extent of disagreement through encouraging consensus or by reducing the number of expert opinions available. None of the proposals actually deal with the difficulty of resolving what the rules unavoidably recognise as genuine expert disagreement. In an age exhibiting a profusion of often contradictory expert opinions, attempts to reduce the number of experts in legal institutions may appear as an unwarranted form of legal suppression.

The new procedures exemplify aspects of both managerial and technocratic approaches to expert disagreement. Sensitive to a range of institutional and resource pressures, judges are to endeavour to reduce the number and length of trials by limiting the number of experts and narrowing the extent of their disagreement. Where possible judges should encourage experts to resolve factual disagreements, or appoint an expert to resolve the case or constrain the extent of disagreement. If consensus is to be imposed procedurally or bureaucratically, does the court need to concern itself with the quality of expert evidence and the reliability of the consensus, or are agreement and resolution — aspects of judicial managerialism — the dominant motivations? How should the expeditious resolution of disputes be weighed against accuracy? The procedures around the production of expert consensus provide a useful illustration of managerial concerns. Interestingly, and revealingly, the procedures are designed to bind the experts. Not to enhance certainty per se, but for the purpose of ‘enhancing certainty as to how the expert evidence will come out at the trial.’ Here judicial ends motivate the procedural means. From the perspective of many judges the

116 Craig Down, ‘Crying Woolf? Reform of the Adversarial System in Australia’ (1998) 7 JJA 213; Thomson, ‘Life after Woolf’, above n38; Zuckerman & Cranston (eds), Reform of Civil Procedure, above n1; Zuckerman, Civil Justice in Crisis, above n1. Increased judicial managerialism may seem, and in fact be, inevitable, but its implications should not be conceived as system neutral nor its precise manifestations pre-determined. For an early discussion, see: Judith Resnik, ‘Managerial judges’ (1982) 96 Harv LR 376.

117 Joint Conferences of Experts, Supreme Court of New South Wales, Practice Note 121s 2(e) (italics added). Such ‘binding’ may actually contradict some of the norms celebrated elsewhere by judges who wish to encourage experts to modify their views. Where experts modify their views in good faith away from the joint report, there may be institutional and logistical pressures to prioritise the report and perhaps disallow additional expert evidence or cross-examination.
reforms must appear to offer a number of desirable features. They are designed to reduce the number of experts appearing, to reduce the extent of expert disagreement and reduce the volume of litigation. While all of these goals may appeal to fact-finders, none need be equated with reliability or accuracy. Practical and logistical pressures on judges are sometimes in tension with the expectations of litigants and even citizens.\(^{118}\)

Some of the technocratic and managerialist sentiments find endorsement in judicial pronouncements. According to Justice Abadee, these goals may even grant experts a more substantive role in the legal system:

> Joint conferences in my view will increase in number and frequency of use. In some ways conferences of experts are an alternative way of resolving disputes. In the event of joint experts agreeing on a matter there may be little left to litigate. Thus the expert’s role is no longer to be perceived as merely a participatory role in the adversarial system of litigation.\(^{119}\)

Here, ‘little left to litigate’ is implicitly interchangeable with ‘little left to decide’. This equation represents a sleight of hand. Managerialist goals surreptitiously collapse truth into process. Viewed from this perspective it is arguable that the new procedures offer few advances in terms of the accuracy of decision making. Rather, the reforms promise a *bureaucratisation of truth* predicated upon simplistic sociological models of expertise.\(^{120}\)

The recent reforms place the legal system on a road to technocracy.\(^{121}\) It may, in a narrow economic sense, be efficient to have experts meet and negotiate consensus in private, but is this the type of resolution of disputes that is desirable, or that will facilitate enhanced public confidence in and understanding of legal procedures?\(^{122}\) At present, few plaintiffs talk of ‘having their day in an (private) expert conference.’ Further, will it be possible for judges to maintain their

---


120 John Anthony Jolowicz, ‘The Woolf Report and the Adversary System’ (1996) 15 *Civil Justice Quarterly* 198. In contrast, the comparativist scholar and francophile, Jolowicz, suggests that notwithstanding some loss to the adversarial nature of the civil justice system the Woolf reforms to civil procedure in England and Wales will ensure substantive as well as procedural justice. Compare, Cecil & Willging, *Court-Appointed Experts*, above n41 at 94.

authority over the entire proceedings if they are confronted with pre-assembled ‘matters of fact’ which may be dispositive, however (un)fair? Will the rule of fact, or some short-sighted judicial manifestation of it, eventually trump the rule of law?

A. Are (Legal) Academics Irrelevant?

In closing, it might be an informative reflexive exercise — exemplifying the core role of representation in the treatment of expertise — to project a simplistic image of objectivity back onto the proponents of reform to assess (and simultaneously construct) their performance in relation to the expectations raised by the procedural reforms. We might ask: As exponents of legal reform, how do these judges and commentators perform when assessed by the very criteria they seek to apply to others?

It might be argued that certain reformers have become advocates for particular images of expertise (and law). Rather than being objective, descriptions of the legal system suggest the existence of serious public problems — junk science and partisanship — which require specific procedural innovations. Earlier, we saw how both of these concerns are at the very least contentious. However, it might be argued that rather than reflecting a commitment to truth and justice — or an overriding duty to the court — the reforms are consistent with a range of professional interests and managerial (and corporate) aspirations such as expediting trials, employing fewer experts, and reducing the extent of expert disagreement. These goals are intuitively appealing because they operate as a procedural metonymic for truth and reliability. They are, however, merely an attempt to bureaucratise expert knowledge and should be recognised as fairly crude and unwieldy developments. They are not a substitute for reliability and should not be mistaken for one. Returning to the standards required of other experts, can we be satisfied with the answers to questions such as: Are these reforms impartial? Will they assist the court? and, the persistent issue of (relevance): Were they necessary in the first place?

If some judges find it so difficult to engage with relevant literatures, identify and list their own assumptions, and their duty to the court is susceptible to subversive, or at least inconsistent, readings, then can we realistically expect expert witnesses to perform (or have their performances presented) any more satisfactorily?

---


123 Hilgartner, Science on Stage, above n96; Cecil & Willging, Court-Appointed Experts, above n41 at 76.

124 See the collection: Steven Woolgar (ed), Knowledge and Reflexivity (1988).

125 These are set out in Practice Note 121 s2(e), above n117.
In an era when judges seem sensitive and vulnerable to external criticism and are particularly attentive to representations of public confidence, the appropriate response to complex expert evidence might not be to promote procedures which reduce or displace the amount and extent of disagreement, especially when expert disagreement is an unavoidable, and anticipated, feature of contemporary life. The cultural cringe manifested in crying Lord Woolf, in the absence of detailed knowledge about the operation of expertise within our legal system, in combination with extremely limited empirical evidence supporting the prevalence of charlatans and junk science, without an explicit theoretical model of expertise or the nature of objectivity and with an increasingly sceptical public, might facilitate the more rapid disposition of cases but that goal can hardly be equated with legal legitimacy and the dispensation of justice.
Burdens and Standards in Civil Litigation

C R WILLIAMS*

Abstract

This article provides an account of where the burden of proof lies in respect of some of the major issues that commonly arise in civil cases, and an account of the standard of proof that must be met in civil cases. Determining which party carries the burden of proof requires consideration of the substantial nature of the plaintiff’s cause of action and the defences available rather than merely having recourse to the form in which legal rules are stated or causes of action created by statute are expressed. The distinction must be drawn between the legal, evidential and tactical burdens, and regard had to considerations of practical convenience. In expressing the standard of proof the expression ‘balance of probabilities’, as interpreted by Dixon J in Briginshaw v Briginshaw (1938) 60 CLR 336, gives the court a degree of flexibility in applying the civil standard without requiring the development of intermediate standards between the normal civil standard and the criminal.

* Sir John Barry Chair of Law, Monash University.
1. Introduction

The aim of a trial is to reconstruct the facts of a past event so that the tribunal may then apply the law to those facts. This is a process that can be achieved with only a limited degree of success, and in consequence the question ‘how is the finder of fact to proceed in cases of uncertainty?’ assumes considerable significance. The questions ‘who has the burden of proof?’, or ‘who will fail if the tribunal is unable to determine what occurred?’ become important at two points of time in a trial.¹

First, at the close of a party’s case the opposing party may submit that there is ‘no case to answer’, i.e. that the first party has not discharged the burden of proof on a particular issue to the extent of requiring the opponent to respond. If such submission is accepted, then in cases tried by judge alone the issue will not be further considered by the judge, and in cases tried with a jury will be withdrawn from the consideration of the jury. Where the burden of proof lies upon the plaintiff then, if that issue is an essential element of the cause of action pleaded, the case will be dismissed.

Secondly, the burden of proof becomes crucial if at the conclusion of the case the tribunal is unable to determine where the truth lies. In such situations each issue in respect of which the tribunal of fact is undecided is determined against the party who carried the burden of proof on that issue.

Belief is a matter of degree, and the tribunal requires a standard by which belief must be measured before being accepted as a basis for decision. The standard of proof is the degree or level to which the party carrying a burden of proof must meet that burden.

The present article is intended to provide an account of where the burden of proof lies in respect of some of the major issues which commonly arise in civil cases, an analysis of the factors which lead to the placing of the burden on one party rather than the other, and an account of the standard of proof which must be met in civil cases.

2. Terminology and the Shifting Burden of Proof

In both criminal and civil cases the phrase ‘burden of proof’ is commonly said to be used in two quite distinct senses. In one sense it means ‘The peculiar duty of him who has the risk of any given proposition on which the parties are at issue — who will lose the case if he does not make this proposition out, when all has been said and done.’ ² The burden of proof in this sense has variously been termed ‘the

¹ The burden of proof also determines which party has the right to begin calling evidence.
² James B Thayer, A Preliminary Treatise on Evidence at the Common Law (1898) at 355.
legal burden, ‘the risk of non-persuasion’, ‘the fixed burden of proof’, ‘the probative burden’, and ‘the burden of persuasion’. The phrase ‘the legal burden’ is the most commonly used, and is the one that will be adopted in this article.

Thayer defines the second sense in which the phrase ‘burden of proof’ is used as ‘the duty of going forward in argument or in producing evidence; whether at the beginning of a case or at any later moment throughout the trial or the discussion’. Cross refers to this burden as:

[T]he obligation to show, if called upon to do so, that there is sufficient evidence to raise an issue as to the existence or non-existence of a fact in issue, due regard being had to the standard of proof demanded of the party under such obligation.

There is a significant difference in the formulations adopted by Thayer and by Cross, and this will be considered presently. This burden has been referred to as ‘the evidential burden’, ‘the duty of producing evidence satisfactory to the judge’, and ‘the burden of proof in the sense of introducing evidence’. The phrase ‘the evidential burden’ is the one that will be adopted in this article.

These two burdens are, of course, quite distinct and arise at different points in a trial. Where the question is whether a no case submission should be accepted, or whether a particular issue should be left to the jury or considered by the judge in a case tried by judge alone, it is the evidential burden which is in issue. Where the question is what should be done if, at the end of the day, the tribunal is unsure where the truth lies, it is the legal burden that is in issue. Wigmore summarises the distinction between the two burdens as follows:

The risk of non-persuasion operates when the case has come into the hands of the jury, while the duty of producing evidence implies a liability to a ruling by the judge disposing of the issue without leaving the question open to the jury’s deliberations.

---

5 Nigel Bridge, ‘Presumptions and Burdens’ (1949) 12 MLR 273.
6 *DPP v Morgan* [1976] AC 182 at 209 (Lord Hailsham); at 217 (Lord Simon).
8 Thayer, above n2.
9 Heydon, *Cross on Evidence*, above n3 at 199.
11 Wigmore, above n4 at §2487; Cleary, above n7 at 783.
12 *Purkess v Crittenden* (1965) 114 CLR 164 at 168 (Barwick CJ, Kitto & Taylor JJ).
13 Wigmore, above n4 at 299.
Considerable confusion has arisen from a failure to distinguish adequately between these two burdens. Thus, appeals have been allowed because the trial judge directed the jury that the legal burden rested on a party whereas in fact only the evidential burden rested upon that party.\textsuperscript{14} On occasion appeal courts themselves discuss questions of burden of proof without making it clear whether they are talking about the legal burden or the evidential burden.\textsuperscript{15}

A more subtle source of confusion, however, lies in the fact that the expression ‘evidential burden’ is commonly used to refer to two notions that are in fact quite distinct. In the first sense the evidential burden means the burden of adducing evidence on an issue on pain of having the trial judge determine that issue in favour of the opponent. This is the sense in which the expression ‘evidential burden’ is used in the passage quoted from Cross, above.

The second sense in which the expression ‘evidential burden’ is used includes the burden resting upon a party who appears to be at risk of losing on a given issue at a particular point in the trial. The party is under an evidential burden in the sense that if the party does not produce evidence or further evidence he or she runs the risk of ultimately losing on that issue. The passage from Thayer immediately preceding that from Cross incorporates this meaning into the concept of burden of proof.

Both judges\textsuperscript{16} and academic commentators\textsuperscript{17} commonly treat the evidential burden as incorporating both these notions. Yet they are clearly distinct. The former involves a question of law, while the latter involves merely a tactical evaluation of who is winning at a particular point in time. Among commentators who in fact distinguish between these two concepts, the expression ‘evidential burden’ is commonly reserved for the former notion and the expression ‘tactical burden’\textsuperscript{18} or ‘provisional burden’\textsuperscript{19} is used to refer to the latter notion. The expression ‘tactical burden’ is the one that will be used in this article.

It is commonly said that the burden of proof ‘shifts’ during the course of a trial. Sometimes it is asserted that the evidential burden but not the legal burden may

\textsuperscript{14} Eg, Sutton v Sadler (1857) 3 CB (NS) 87; Purkess v Crittenden (1965) 114 CLR 164; R v Guerin [1967] 1 NSWLR 255; R v Joyce [1970] SASR 184; R v Hinc [1972] Qd R 272.
\textsuperscript{15} Eg Redpath v Redpath and Milligan [1950] 1 All ER 600; Watts v Rake (1960) 108 CLR 158.
\textsuperscript{17} Thayer, above n2; Phipson, above n7 at 64–68.
\textsuperscript{18} Heydon, Cross on Evidence, above n3 at 197; Peter K Waight & Charles Robert Williams, Evidence: Commentary and Materials (6th ed, 2002) at 57.
\textsuperscript{19} Denning, above n3. Note also Poricanin v Australian Consolidated Industries Ltd [1979] 2 NSWLR 419.
Sometimes it is asserted that both burdens may shift. Such assertions involve, it is submitted, an inadequate analysis of the concept of burden of proof.

Any case, whether civil or criminal, involves a finite number of potential issues; a number of elements comprise each cause of action and each criminal offence, and a number of potential defences are available to the defendant or the accused. In respect of each of these issues, rules of law determine upon which party the legal and the evidential burden lies. The only burden that does shift in the course of a trial is the tactical burden. A failure to distinguish between the evidential burden and the tactical burden would appear to be responsible for most of the suggestions that the burden of proof shifts. At any given point in time a party who has the legal burden in respect of a particular issue may appear more or less likely to be able to discharge that burden. If that party appears likely to be able to discharge the legal burden, then the tactical burden shifts to the other party; the other party must produce contradictory evidence or run the risk of losing on that issue. If that other party produces such evidence, then the tactical burden may shift back to the party bearing the legal burden. Such swings of the forensic pendulum as a case progresses involve, however, no shift in either the legal or the evidential burden.

It is sometimes suggested that the operation of presumptions is to shift the evidential and even the legal burden of proof. However, this is not so. When one party establishes facts that give rise to a presumption, he or she merely shifts the tactical burden onto the other party in a particularly strong sense. The other party is faced not merely with the likelihood of losing on that issue unless contradictory evidence is produced, but is faced with the certainty of doing so. The operation of a presumption is that it is sufficient, unless rebutted, to satisfy the requirements of a burden. It is not to shift a burden.

3. **The Incidence of the Burden of Proof in Civil Cases**

**A. The Legal Burden**

(i) **The Legal Burden is Decisive When the Evidence is Evenly Balanced**

If the tribunal of fact is satisfied to the appropriate standard as to the existence or non-existence of the facts representing the issues in the case, then of course no question of burden of proof arises. If the tribunal is not so satisfied in respect of any issue, then that issue must be determined against the party carrying the burden of proof. There is no power in the court to attempt to achieve a measure of rough justice by adopting an intermediate hypothesis.

In *Nesterczuk v Mortimore* a head-on collision occurred between two vehicles being driven by the parties on a straight and level road. The plaintiff sued

---

20 Thayer, above n 2 at 355–370; Wigmore, above n 4 at 300–301; Parkess v Crittenden (1965) 114 CLR 164; R v Morgan [1976] AC 182 at 217.
22 Id at 228–229.
23 *Robins v National Trust Co* [1927] AC 515 at 520.
24 (1965) 115 CLR 140.
the defendant in negligence, and the defendant counter-claimed. Both claimed that
they were on their correct side of the road, and that the accident had occurred when
the other party had suddenly swerved onto his incorrect side. The trial judge
dismissed both the claim and counter-claim.

On appeal to the High Court it was argued that if the trial judge was unable to
decide that the accident was due to the negligence of one party rather than the
other, then the judge was bound to take the view that both were to blame and
apportion the damages recoverable accordingly. The court rejected the argument.
Owen J, with whom Kitto and Windeyer JJ agreed, stated:

The learned trial judge was unable to determine whether the collision had
occurred on the northern or the southern side or in the centre of the road and no
good reason has been shown why, on the meagre evidence, an appellate court
should take a different view. In the circumstances of the case, to say that the
probabilities favour the view that both drivers were to blame rather than one or
the other was wholly responsible would be a mere guess.25

A finding of negligence on the part of both drivers could only, his Honour held, be
justifiable if there was evidence to suggest this was a probable cause of the
accident, as for example if the accident had occurred in the centre of the road.26

(ii) The General Rule

A basic test for determining which party has the burden of proof is contained in the
judgment of Walsh JA in *Currie v Dempsey*.27 His Honour stated:

In my opinion [the legal burden of proof] lies on a plaintiff, if the fact alleged
(whether affirmative or negative in form) is an essential element in his cause of
action, eg if its existence is a condition precedent to his right to maintain the
action. The onus is on the defendant, if the allegation is not a denial of an essential
ingredient in the cause of action, but is one which, if established, will constitute a
good defence, that is, an “avoidance” of the claim which, prima facie, the plaintiff
has.28

The difficulty with such a test is that most matters can be classified either as an
‘essential element’ in the establishment of a cause of action, or as a matter
‘avoiding’ the prima facie claim of the plaintiff. For example, a plaintiff sues a
defendant for repayment of a loan, and the defendant claims either (a) that the
money was given by way of gift, or (b) that the loan has been repaid. It may be
argued that the elements of the plaintiff’s claim are (i) the plaintiff gave the
defendant money, (ii) that it was given by way of loan, and (iii) that the money has
not been repaid. If this is correct, the burden of proof is on the plaintiff. With equal

25 Id at 155. See also *Holloway v McFeeters* (1956) 94 CLR 470; *Maher-Smith v Gaw* [1969] VR 371.
26 As in *Bray v Palmer* [1953] 1 WLR 1455; *Baker v Market Harborough Industrial Co-Operative Society Ltd* [1953] 1 WLR 1472.
28 Id at 539.
logic however, it may be argued that items (ii) and (iii) operate to avoid liability, in which case the burden of disproving these items is on the defendant. In fact, it is established in Australia that the burden of proof is on the plaintiff in respect of item (ii) and on the defendant in respect of item (iii). In Joaquin v Hall the Victorian Supreme Court held that the burden of proving that money was given by way of loan and not gift is on the plaintiff. In Young v Queensland Trustees Ltd the High Court held that 'the law has always been that it lies upon a defendant to make out a defence of payment by way of discharge'.

Decisions such as these cannot be the result purely of logic. This point was made vividly by Professor Julius Stone in the following passage:

What is the difference in logic between a quality of a class as contained in the definition of the class, and a quality of a class as contained in an exception to the class? The answer appears to be — none at all. Every qualification of a class can equally be stated without any change of meaning, as an exception to a class not so qualified. Thus, the proposition “All animals have four legs except gorillas”, and the proposition “All animals which are not gorillas have four legs”, are, so far as their meanings are concerned, identical….

What then determines where the burden of proof lies? Often the question will turn on the form in which a legal rule is traditionally stated. If an issue is commonly listed among the constituent elements of a cause of action, the burden of proof will be said to be on the plaintiff. If the issue is commonly referred to as a factor leading to the avoidance of liability, the burden of proof will be on the defendant. Thus, in torts law the burden of proving negligence rests on the plaintiff, while the burden of proving contributory negligence is on the defendant. In marine insurance, the burden of proving that the ship was lost by perils of the sea rests on the shipowners. However, legal rules usually assume particular forms for reasons of clarity of exposition quite unrelated to questions of burden of proof. In one case, Sir Wilfred Greene MR remarked:

---

30 (1956) 99 CLR 560.
31 Id at 562.
33 For this reason the rules relating to burden of proof are sometimes said to be the province of substantive law rather than the law of evidence, see Heydon, Cross on Evidence, above n2 at 196. The Evidence Act 1995 (Cth) and Evidence Act 1995 (NSW) do not deal with the question of determining which party carries the burden of proof in civil cases.
35 Lopes v Taylor (1970) 44 ALJR 412.
36 Rhesa Shipping Co SA v Edmunds [1985] 1 WLR 948.
Nothing can be more dangerous when a question of burden of proof has to be considered, than to pick out the language of learned judges used in cases where neither they nor anybody else was thinking of the question of burden of proof.\(^{37}\)

Where the form in which a legal rule is commonly stated is not decisive, the basic starting point is that the burden of proof should normally be on the plaintiff as to all issues. In the absence of reasons to the contrary, the party who invoked the judicial process and compelled the defendant’s involvement in that process, should run the risk of having decided against her or him any issue as to which the tribunal of fact is, at the end of the day, undecided.

In some cases a key factor, either re-enforcing the policy of placing the burden of proof on the plaintiff in cases of uncertainty or displacing that policy, may be the consideration that it is possible to predict that, if no evidence were given on a particular issue, the correct result would be reached in most cases by deciding in favour of one party rather than the other. If this is so then, of course, the percentage of correct decisions will be maximised by placing the burden of proof on the opponent.\(^{38}\)

Illustrative is the decision of the House of Lords in *Joseph Constantine Steamship Line Ltd v Imperial Smelting Corporation Ltd*.\(^{39}\) The defendants chartered a ship to the plaintiffs for a voyage from Port Pirie to Europe. While anchored at Port Pirie, an explosion caused such damage to the ship that it could not perform the charter party. The plaintiffs claimed damages, and the defendants raised the defence of frustration. The plaintiffs claimed that the defendants were entitled to rely on frustration only if they established that the explosion occurred without any fault on their part. The defendants claimed that once the frustrating event is established, the burden of proof is on the plaintiffs to establish such default on the part of the defendants as would deprive them of their right to rely upon it. The trial judge held that the burden was upon the plaintiffs to show the frustration occurred through negligence on the part of the defendants. The Court of Appeal reversed this decision. On appeal the House of Lords affirmed the decision of the trial judge. The decision of the House of Lords would appear correct. In the majority of cases in which the performance of a contract is frustrated, that frustration is, presumably, unlikely to be the result of negligence or other fault on the part of the defendant. If this is so, then placing the burden of proof on the plaintiff will maximise the percentage of correct decisions that will be reached.\(^{40}\)

The House of Lords may have been influenced by similar considerations in *Henderson v Henry E Jenkins & Sons*.\(^{41}\) A lorry owned by the defendants was descending a hill when the brakes failed. The lorry struck and killed a pedestrian.

\(^{37}\) *Imperial Smelting Corporation Ltd v Joseph Constantine Steamship Line Ltd* [1940] 2 KB 430 at 433 (Sir Wilfred Greene MR). Although the decision of the Court of Appeal was reversed on appeal to the House of Lords (*Joseph Constantine Steamship Line Ltd v Imperial Smelting Corporation Ltd* [1942] AC 154), nothing in their Lordships’ judgments reflected upon the accuracy of this dictum.

\(^{38}\) For elaboration of the significance of this form of reasoning, see Stone, above n32.

\(^{39}\) [1942] AC 154.
The failure was due to the sudden escape of brake fluid from a hole in a pipe in the hydraulic braking system resulting from corrosion of that pipe. The widow of the deceased claimed damages, alleging that the defendants had been negligent in failing to keep the braking system in efficient repair. The defendants pleaded that the accident had been caused by a latent defect in the braking system. The defence of latent defect applies only where all reasonable care has been taken. The defendants established that they had carried out routine maintenance, but failed to show that there were no special circumstances in the past use of the vehicle to indicate that the lorry might have been subjected to a corrosive agent resulting in the corrosion of the pipe. The House of Lords held, by a majority of three to two, that the burden of proving that all reasonable care had been taken rested on the defendant and that the plaintiff was therefore entitled to succeed.\textsuperscript{42} As a matter of policy this decision also would seem correct. Since in most cases where injury is caused as a result of mechanical failure, the explanation is likely to be failure to keep the items in a sufficient state of good repair rather than the presence of a latent defect not discoverable by taking all reasonable steps, it follows that the percentage of correct results will be maximised by placing the burden of proof on the defendant rather than the plaintiff.

(iii) Agreement Between the Parties

In cases of disputes arising under agreements between the parties, the form of the agreement may affect the onus of proof. In \textit{Stateliner Pty Ltd v Legal and General Assurance Society Ltd}\textsuperscript{43} a policy of insurance on a passenger bus provided that the policy did not cover damage caused while the vehicle was being used "in an unsafe or unroadworthy condition, unless such condition could not reasonably have been detected by the Insured".\textsuperscript{44} The Supreme Court of South Australia held that this agreement obliged the insurer to prove the bus had been used while in an unsafe or unroadworthy condition, but placed the obligation of establishing that such defective condition could not reasonably have been detected by the insured on the insured.\textsuperscript{45}

(iv) Measure of Damages in Tort Cases

The issue of burden of proof in relation to damages in tort cases is of particular significance because of the difficulties involved in measuring damages. The plaintiff must, of course, prove the extent of her or his damages.\textsuperscript{46} Where the

\begin{itemize}
\item Similar reasoning explains the rule that the burden of proof is on the bailee to show that bailed goods were lost without her or his default since no doubt in most cases fault on the part of the bailee is the correct explanation for the loss of such goods (\textit{Tozer Kemsley and Millbourn (Australasia) Pty Ltd v Collier’s Interstate Transport Service Ltd} (1956) 94 CLR 384; \textit{Fankhauser v Mark Dykes Pty Ltd} [1960] VR 376).
\item \textsuperscript{1970} AC 282.
\item \textsuperscript{1981} 29 SASR 16.
\item \textit{Stateliner Pty Ltd v Legal and General Assurance Society Ltd} (1981) 29 SASR 16.
\item \textsuperscript{1967} VR 161 at 164.
\end{itemize}
plaintiff alleges a partial loss of earning capacity, the burden is upon the plaintiff to prove the extent of that partial loss.\textsuperscript{47} Even where the plaintiff establishes that a certain figure represents the likely extent of her or his loss, the burden remains on the plaintiff to disprove any offsetting gains. In \textit{Stewart v Dillingham Constructions Pty Ltd}\textsuperscript{48} an action was brought on behalf of infant plaintiffs in respect of the death of their father. The mother gave evidence on behalf of the plaintiffs, in the course of which she testified as to her intended re-marriage and the prospective adoption of the plaintiffs by her future husband. The trial judge directed the jury that the burden of proof in relation to the benefits the plaintiffs might obtain through their mother’s proposed marriage and their prospective adoption rested on the defendant. On appeal by the defendant, the Supreme Court of Victoria held that the direction of the judge was incorrect. The court held that it is for the plaintiff to prove her or his net loss, and that this includes the legal burden of disproving any offsetting gains.

Differing views have been adopted by the High Court in relation to burden of proof on the issue of whether a plaintiff’s damages ought to be reduced because a pre-existing condition would ultimately have led the plaintiff to her or his present state. In \textit{Watts v Rake}\textsuperscript{49} the High Court held that the burden of proof rested on the defendant. In \textit{Purkess v Crittenden}\textsuperscript{50} however, the court interpreted \textit{Watts v Rake} as placing the evidential burden only on the defendant and held the legal burden to rest upon the plaintiff.\textsuperscript{51} It would seem that on principle the decision in the earlier case is to be preferred. Cases in which the injuries resulting from a tort constitute no more than a hastening of the inevitable are comparatively rare. This being so, the method by which the percentage of correct decisions would be maximised would be to place the legal burden upon the defendant.

The one issue in respect of which it has been established that the legal burden of proof is upon the defendants is mitigation of damages. In \textit{Munce v Vinidex Tubemakers Pty Ltd}\textsuperscript{52} the defendants alleged that an operation would alleviate the plaintiff’s condition and that it was unreasonable of him not to submit to the operation. On appeal the Supreme Court of New South Wales upheld the direction of the trial judge that the burden of proof on this issue rested on the defendant. Glass JA stated:

\begin{quote}
There is authority of long standing which establishes an exception to the principle that the plaintiff bears the onus of proving all matters relating to damages. The exception relates to any disputed question that is truly a matter of mitigation of damages. In relation to questions properly so classified, the defendant is subject to both burdens. He must not only introduce evidence that the plaintiff has failed
\end{quote}

\textsuperscript{49} (1960) 108 CLR 158.
\textsuperscript{50} (1965) 114 CLR 164.
\textsuperscript{51} See also \textit{Pastras v Commonwealth} [1967] VR 161; \textit{Apollo Shower Screens Pty Ltd v Building and Construction Industry Long Service Payments Corp} (1985) 1 NSWLR 561.
\textsuperscript{52} [1974] 2 NSWLR 235.
It may be difficult to determine whether an issue is properly classified as one relating to the extent of the plaintiff’s damages or as a question of mitigation of damages. Thus, if the plaintiff fails to obtain employment following her or his accident, and alleges that this failure was the result of her or his incapacity to work as a consequence of the accident, then the burden of proof will be upon the plaintiff since this relates to the measure of damages. If, however, the reason assigned is that the plaintiff was unable to find suitable employment, this will be classified as being a question of mitigation of damages and the burden will be on the defendant to show that the plaintiff could have found such employment.

B. The Evidential Burden

The general rule in civil cases is that the party who has the legal burden also has the evidential burden. In a negligence action, for example, the plaintiff has the legal burden of ultimately persuading the tribunal of fact that the defendant owed the plaintiff a duty of care, that the defendant was in breach of that duty, and that the breach caused damage to the plaintiff. If the plaintiff does not discharge this legal burden, then the plaintiff’s claim will fail. As the plaintiff carries the legal burden in respect of these matters, he or she also carries the evidential burden, and the plaintiff’s case will be withdrawn from the jury unless he or she discharges this burden.

Several exceptions to this general rule are to be found in the area of tort law. As stated above, while the legal burden on the issue of whether a pre-existing condition would in time have led to the plaintiff’s present state rests upon the plaintiff, the evidential burden is upon the defendant. It is also established that, while the legal burden of disproving the existence of offsetting gains is upon the plaintiff, the evidential burden of showing the likelihood of such gains rests upon the defendant.

In the tort of deceit the legal burden of establishing all the elements of the tort rests upon the plaintiff. However, where it is shown that a representation that was false and fraudulent was made, the defendant carries the evidential burden of showing the representation was not acted upon. In Gould v Vaggelas, the appellants claimed they had been induced to enter into a contract by the deceit of

---

53 Id at 239. See also Watts v Rake (1960) 108 CLR 158 at 159 (Dixon CJ); Buczynski v McDonald (1971) 1 SASR 569; Plenty v Argus [1975] WAR 155; Craig v Garfit-Mottram (1977) ACTR 12; Fazlic v Milingimbi Community Inc (1982) 38 ALR 424 at 427; Metal Fabrications (Vic) Pty Ltd v Kelcey [1986] VR 507. A contrary view was adopted by the Privy Council in Selvanayagam v University of the West Indies [1983] 1 All ER 824.


55 Purkess v Crittenden, above n50.


the respondents. The High Court held that the evidential burden of refuting causation was cast on the respondent, notwithstanding that the legal burden in respect of this issue rested with the appellant. Wilson J stated:

Where a plaintiff shows that a defendant has made false statements to him intending thereby to induce him to enter into a contract and those statements are of such a nature as would be likely to provide such inducement and the plaintiff did in fact enter into that contract and thereby suffered damage and nothing more appears, common sense would demand the conclusion that the false representations played at least some part in inducing the plaintiff to enter into the contract. However, it is open to the defendant to obstruct the drawing of that natural inference of fact by showing that there were other relevant circumstances. Examples commonly given of such circumstances are that the plaintiff not only actually knew the true facts but also knew them to be the truth or that the plaintiff either by his words or conduct disavowed any reliance on the fraudulent representations. It is entirely accurate to speak of an onus resting on a defendant to draw attention to the presence of circumstances such as those I have described in order to show that the inference of the fact of inducement which would ordinarily be drawn from the fraudulent making of a false statement calculated to induce a person to enter into a contract followed by entry into that contract should not in all the circumstances be drawn. But it is no more than an evidentiary onus—an obligation to point to the existence of circumstances which tend to rebut the inference which would ordinarily be drawn from the primary facts.58

Exceptions are to be found in other fields also. In the area of wills, while the legal burden of proving testamentary capacity rests on the party seeking to take under the will, the evidential burden of showing lack of competence rests on the party who challenges the will.59 In appeals against assessments by the Commissioner of Taxation, the taxpayer bears the legal burden of proof while the Commissioner bears the evidential burden of raising a particular matter in evidence so as to require the taxpayer to deal with that matter.60

4. The Burden of Proof in Causes of Action Created by Statute

Some statutes distinctly deal with the question of burden of proof, eg s 35(2) of the Bills of Exchange Act 1909 (Cth).61 In most cases, however, statutes creating causes of action do not refer to the question of burden of proof. However, in such cases the statute nonetheless provides an authoritative formulation of the requirements of the cause of action and this formulation may be treated as carrying implications as to which party bears the burden of proof. The courts have thus been required to develop principles designed to aid in determining which party should be treated as bearing the burden of proof in respect of each issue in dispute.62

58 Id at 238.
59 Sutton v Sadler (1857) 3 CB (NS) 87; 140 ER 671.
61 For discussion of this provision, see Heydon, Cross on Evidence, above n3 at 213.
If the issue in dispute constitutes a necessary ingredient of the plaintiff’s cause of action, then the burden will be upon the plaintiff. If, however, it is classified as a matter exempting the defendant from liability in a certain event, then the burden will be upon the defendant. While the form and structure of the particular statute is significant, what is involved is often more than a mere question of construction. It is, rather, an attempt to determine the substantial nature of the cause of action created by the particular statute.

Illustrative is *Darling Island Stevedoring and Lighterage Company Limited v Jacobsen*. The *Workers Compensation Act 1926* (NSW) was expressed to cover cases where the employee ‘received injury without his own default or wilful act on any of the daily or other periodic journeys’ between the employee’s place of abode and place of employment. The plaintiff was injured in the course of such a journey. The High Court held that in spite of the fact that the form in which the clause was cast favoured ‘the view that the words in question express part of the description of the primary or general grounds of liability’, the burden of proving that the injury occurred as a result of the plaintiff’s ‘own default or wilful act’ rested upon the defendant. Rich J stated:

> This seems to me to be but another case in which we should weigh the substantial meaning against the form of its expression. I am clearly of opinion that the substantial meaning of the provision is to lay down a general rule of liability arising from the nature of the journey on which the worker was engaged when he sustained the injury. That rule is simple and all embracing. Having adopted it as a principle of general application, the legislature proceeded to introduce by way of exclusion the case of such an injury so sustained having been caused by a special element, namely the default or wilful act of the worker. It is a special exemption, exception or exclusion from the operation of the rule on a new and additional factor, namely fault or willfulness. These are things which, according to the sense of fairness and justice which inspires the common law, we usually require to be proved against not disproved by the worker, or in the case of his death, his representatives. The substantial nature of the provision is more demonstrate to my mind than the form of the draftsmanship, which was probably determined more by the craving for brevity than the consciousness of the niceties of the rules affecting the burden of proof.

Similarly, considerations of substance governed the decision in *Vines v Djordjevich*. The plaintiff was attempting to cross a street when she was

---


63 (1945) 70 CLR 635.

64 Id at 644 (Dixon J).

65 Held by Rich, Dixon, McTiernan & Williams JJ, Starke J dissenting.

66 (1945) 70 CLR 635, 639. See also *Pye v Metropolitan Coal Company Ltd* (1934) 50 CLR 614; (1936) 55 CLR 138.
knocked down and injured by a motor vehicle. The identity of the driver was not
ascertained. In an action against the nominal defendant brought under the Motor
Car Act 1951 (Vic) the question arose whether the plaintiff carried the burden of
proving that notice of his claim was given to the minister ‘as soon as possible after
he knew that the identity of the motor car could not be established’. The
requirement of notice was contained in a proviso, which favoured the view that the
burden rested upon the defendant. The High Court held, however, that the form of
the legislation was not decisive. The court stated:

In the end, of course, it is a matter of the intention that ought, in the case of a
particular enactment, to be ascribed to the legislature and therefore the manner in
which the legislature has expressed its will must remain of importance. But
whether the form is that of a proviso or of an exception, the intrinsic character of
the provision that the proviso makes and its real effect cannot be put out of
consideration in determining where the burden of proof lies. When an enactment
is stating the grounds of some liability that it is imposing or the conditions giving
rise to some right that it is creating, it is possible that in defining the elements
forming the title to the right or the basis of the liability the provision may rely
upon qualifications, exceptions or provisos and it may employ negative as well as
positive expressions. Yet it may be sufficiently clear that the whole amounts to a
statement of the complete factual situation which must be found to exist before
anybody obtains a right or incurs a liability under the provision. In other words it
may embody the principle which the legislature seeks to apply generally.68

The court held that since the proviso expressed a requirement that had to be
fulfilled by all as a condition precedent to a cause of action being available, the
burden of proof rested on the plaintiff.

In contrast to Vines v Djordjevitch may be placed Nominal Defendant v
Dunstan.69 An admittedly uninsured vehicle injured the plaintiff. In an action
against the nominal defendant under the Motor Vehicles (Third Party Insurance)
Act 1942 (NSW) the question arose as to who had the burden of proof on the issue
of whether the vehicle was exempt from the operation of the Act by regulations
made pursuant to the Act. The High Court held that since such exemptions
constituted ‘special grounds of exculpation’ which are capable of being
established only by proof of ‘additional facts of a special nature’, the burden rested
upon the defendant.70

67 (1955) 91 CLR 512.
68 Id at 519 (Dixon CJ, McTiernan, Webb, Fullagar & Kitto JJ). See also Neptune Oil Co Pty Ltd
v Fowler (1963) 63 SR (NSW) 530, Currie v Dempney [1967] 2 NSW R 532, Lynch v Attwood
69 (1963) 109 CLR 143.
70 Id at 151. See also Romano v Foggo [1974] 2 NSWLR 336; Banque Commerciale SA, en
Liquidation v Akhil Holdings Ltd (1990) 169 CLR 279 at 285; Attorney-General (ACT) v ACT
Minister for Environment (1993) 43 FCR 329; Minister for Immigration and Ethnic Affairs v
Stateliner Pty Ltd v Legal and General Assurance Society Ltd (1981) 29 SASR 16.
In cases where the form or structure of the legislation does not give definite guidance on the question of burden of proof, the courts will have regard to considerations of policy and convenience. The fact that a matter is ‘peculiarly within the knowledge of one party’, or that it will be easier for that party to prove the matter than her or his opponent, may be significant. In *Nimmo v Alexander Cowan and Sons Ltd*\(^1\) the plaintiff brought an action against his employers for breach of statutory duty in respect of injuries sustained when he fell while unloading bales from a railway wagon. He claimed that his place of work had not been kept safe as required by the *Factories Act* 1961 (UK). Section 29(1) of that Act requires employers ‘so far as reasonably practicable’ to provide and maintain a safe place of work. The question arose whether it was for the plaintiff to prove that it had been reasonably practicable to provide a safe place of work, or for the defendant to show that it had not been so practicable. The House of Lords held that the burden rested upon the employers.\(^2\) Lord Upjohn stated:

I cannot believe that Parliament intended to impose upon the injured workman or, if dead, his widow or other personal representative, the obligation to aver with the necessary particularity the manner in which the employer should have employed reasonably practicable means to make and keep the place safe for him. Although the pursuer can nowadays consult experts, he is at a great disadvantage compared to the employer. He may have little recollection of the accident, or, of course, he may have been killed, and his widow be in an even worse state. Then, on the other hand, it is the duty of the employer to make the place safe so far as is reasonably practicable. It is his duty with his experts to consider the state of the place of work in all its circumstances and to take whatever steps he can, so far as reasonably practicable, to make it safe. He must know and be able to give the reasons why he considered it was impracticable for him to make the place safe. If he cannot explain that, it can only be because he failed to give it proper consideration, in breach of his bounden duty to the safety of his workmen.\(^3\)

5. **Standard of Proof**

At the trial the question of standard of proof may arise at two stages. First, the court must decide whether the party having the evidential burden has satisfied that burden to the required standard, that is, whether that party has established a *prima facie* case. Here the judge is required to consider whether it is reasonably open to the finder of fact to determine the issue in dispute in favour of the party carrying the evidential burden.\(^4\) If so, then that issue will be left to the jury, or in cases tried by judge alone further considered by the judge. Secondly, where an issue is left to

---

\(^1\) [1968] AC 107.

\(^2\) Lord Guest, Lord Upjohn & Lord Pearson; Lord Reid & Lord Wilberforce dissenting.


the jury, the judge must give a proper direction as to the standard that the jury must apply in determining whether the party having the legal burden has satisfied that burden. In cases tried without a jury the judge personally must apply the correct standard.

An appeal court may be concerned with the question of standard of proof in a number of different contexts. First, an appellant may argue that the judge should not have found that the respondent had established a prima facie case. Secondly, an appellant may argue that in the course of her or his summing up the judge misdirected the jury as to the standard of proof or, in cases tried by judge alone, that the judge applied an incorrect standard. Thirdly, an appellant may argue that, on the whole of the evidence, no reasonable jury could have been satisfied of the respondent’s case to the required standard.

A. The Nature of the Legal Standard

In civil cases the legal burden of proof resting on a party is said to be to the standard of ‘the balance of probabilities’ or ‘the preponderance of probabilities’.

This simple statement, however, conceals a significant ambiguity. Must the party carrying the burden of proof merely show that it is more likely than not that the conditions required to establish her or his case exist? Alternatively, must the finder of fact be satisfied that it is more likely than not that the conditions required to establish the party’s case exist? The difference may be crucial. Assume that a plaintiff is trying to show that the brakes of a particular car of a certain make and model were defective. If the plaintiff could establish statistically that 51% of cars of that make and model had defective brakes then, in the absence of other evidence, the plaintiff would have shown that it was more likely than not that the brakes of any particular car of that make and model were defective. If, however, the tribunal is required to be satisfied that the car’s brakes were defective, then the tribunal might regard such evidence as insufficient. A requirement of satisfaction involves something more than an estimate of probabilities; it requires a subjective belief in a state of facts on the part of the body charged with determining the facts.

The latter formulation represents the view that has been adopted in Australia. In Briginshaw v Briginshaw, a husband petitioned for dissolution of marriage on the ground of adultery. The trial Judge felt great difficulty in deciding whether adultery had occurred. He stated that were the case a civil one he might well consider the probabilities were in favour of the petitioner, but that he was not satisfied beyond reasonable doubt that adultery had taken place. The petitioner appealed to the High Court. The court held that the correct standard in divorce cases is the balance of probabilities but that the petitioner was not entitled to a re-trial.

75 A different standard may, of course, be imposed by statute in civil cases, as with the Repatriation Act 1920 (Cth) which imposes the criminal standard on the Repatriation Commission. See Repatriation Commission v Law (1981) 36 ALR 411; Repatriation Commission v Byrne (1981) 40 ALR 296; Repatriation Commission v Perrot (1984) 53 ALR 690.

76 (1938) 60 CLR 336.
The leading judgment was delivered by Dixon J. His Honour held that there exists no third standard of proof intermediate between the criminal and the ordinary civil standard. His Honour held, however, that the jury must be satisfied of those facts which a party must prove, and that what is required to satisfy a jury may vary having regard to the nature of the facts alleged. His Honour stated:

The truth is that, when the law requires the proof of any fact, the tribunal must feel an actual persuasion of its occurrence or existence before it can be found. It cannot be found as a result of a mere mechanical comparison of probabilities independently of any belief in its reality…. Except upon criminal issues to be proved by the prosecution, it is enough that the affirmative of an allegation is made out to the reasonable satisfaction of the tribunal. But reasonable satisfaction is not a state of mind that is attained or established independently of the nature and consequence of the fact or facts to be proved. The seriousness of an allegation made, the inherent unlikelihood of an occurrence of a given description, or the gravity of the consequences flowing from a particular finding are considerations which must affect the answer to the question whether the issue has been proved to the reasonable satisfaction of the tribunal. In such matters ‘reasonable satisfaction’ should not be produced by inexact proofs, indefinite testimony, or indirect inferences. Everyone must feel that, when, for instance, the issue is on which of two dates an admitted occurrence took place, a satisfactory conclusion may be reached on materials of a kind that would not satisfy any sound and prudent judgment if the question was whether some act had been done involving grave moral delinquency.77

His Honour went on to say that the trial judge had found himself unable to arrive at any satisfactory or firm and definite conclusion that adultery had been committed, and that accordingly the petitioner’s appeal should be dismissed.78

The courts have had regard to ‘the nature and consequence of the fact or facts to be proved’ in determining whether they ‘feel an actual persuasion’ of those facts in a variety of circumstances. In Willcox v Sing79 the Supreme Court of Queensland held a trial judge had been entitled to comment that the jury should not


78 The establishment of irretrievable breakdown as the sole ground for dissolution under the Family Law Act 1975 (Cth) meant of course that the ordinary civil standard or proof applies in matrimonial proceedings without the necessity for any gloss on the expression of that standard: see In the Marriage of Pavey (1976) 10 ALR 259.

lightly find against a surgeon where negligence was alleged of such a character as to put his reputation and earning capacity at risk. In *Re Jane*[^80] the Family Court took the view that a finding that it is in the best interests of a physically and mentally disabled woman to order sterilisation requires “something more than a mere tipping of the balance in favour of the proposal.”[^81] In *Shaw v Wolf*[^82] Merkel J held the court should “not lightly make a finding on the balance of probabilities”[^83] that persons were not aboriginals where such finding would result in their election to an ATSIC regional council being held invalid. In *G v H*[^84] however, the High Court held that in maintenance proceedings, where it was established that a particular person could be the father of a child, the question of actual paternity should not be approached on the basis that it involves a grave or serious allegation in the *Briginshaw* sense. In *Clark v NZI Life Ltd*[^85] Thomas J held that while a finding of suicide should not be made lightly, it was not so inherently unlikely or grave as to bring it to the top of the range of the *Briginshaw* test.

The *Briginshaw* approach to standard of proof was affirmed by the High Court in *Neat Holdings Pty Ltd v Karajan Holdings Pty Ltd.*[^86] The plaintiff sued the defendant for deceit, alleging the takings of a business purchased from the defendant had been exaggerated. The defendant claimed the plaintiff had deliberately minimised the takings of the business after acquisition. The trial judge dealt with the issue on the balance of probabilities, but made no express reference to any requirement of clear, cogent or strict proof. The judge found in favour of the plaintiff but that judgment was set aside on appeal to the Full Court of the Supreme Court of Western Australia on the ground that the judge had erred in relation to the standard of proof. On appeal the High Court reinstated the original judgment. In a joint judgment Mason CJ, Brennan, Deane and Gaudron JJ stated:

> The ordinary standard of proof required of a party who bears the onus in civil litigation in this country is proof on the balance of probabilities. That remains so even where the matter to be proved involves criminal conduct or fraud. On the other hand, the strength of the evidence necessary to establish a fact or facts on the balance of probabilities may vary according to the nature of what it is sought to prove. Thus, authoritative statements have often been made to the effect that clear or cogent or strict proof is necessary ‘where so serious a matter as fraud is to be found’. Statements to that effect should not, however, be understood as directed to the standard of proof. Rather, they should be understood as merely reflecting a conventional perception that members of our society do not ordinarily engage in fraudulent or criminal conduct and a judicial approach that a court should not lightly make a finding that, on the balance of probabilities, a party to civil litigation has been guilty of such conduct.[^87]

[^81]: Id at 416. See also *Re L and M: Director-General, Department of Family Services and Aboriginal and Islander Affairs* (1993) 17 Fam LR 357 at 373, 376.
[^83]: Id at 216.
The High Court held, however that where both sides allege serious wrongdoing there is little scope to have regard to the seriousness of the allegations or the gravity of the consequences in reaching a decision. Mason CJ, Brennan, Deane and Gaudron stated:

When an issue falls for determination on the balance of probabilities and the determination depends on a choice between competing and mutually inconsistent allegations of fraudulent conduct, generalisations about the need for clear and cogent proof are likely to be at best unhelpful and at worst misleading. If such generalisations were to affect the proof required of the party bearing the onus of proving the issue, the issue would be determined not on the balance of probabilities but by an unbalanced standard. The most that can validly be said in such a case is that the trial judge should be conscious of the gravity of the allegations made on both sides when reaching his or her conclusion.88

The view of Dixon J in Briginshaw v Briginshaw is, it is submitted, correct in principle.89 It gives the court a degree of flexibility in relation to the issue of standard of proof without requiring the development of intermediate standards between the normal civil standard and the criminal. The Briginshaw approach is taken up in the Evidence Act 1995 (Cth and NSW).90

It may be objected that injustice is done to a plaintiff who may be denied a decision in her or his favour where there is a 51% or more chance that the facts necessary to establish the plaintiff’s claim exist. It is, however, the plaintiff who

---

87 Id at 170–171.
88 Id at 172.
90 Section 140 of the Acts provides:
In a civil proceeding, the court must find the case of a party proved if it is satisfied that the case has been proved on the balance of probabilities.
Without limiting the matters that the court may take into account in deciding whether it is so satisfied, it is to take into account:
(a) the nature of the cause of action or defence; and
(b) the nature of the subject-matter of the proceeding; and
(c) the gravity of the matters alleged.
has chosen to invoke the judicial process and involve the defendant in that process, and when this is remembered it does not seem unjust to require the plaintiff to ‘satisfy’ the tribunal of facts and not merely to demonstrate a statistical probability. A test of statistical probability would itself produce results that would appear unjust. L. Jonathan Cohen takes the case of a rodeo at which a thousand seats were taken but only 499 persons have paid at the turnstiles. Should the proprietor be able successfully to sue 501 patrons chosen at random on the basis that in the case of each person there were 501 chances in a thousand that that person had not paid? It is, it is submitted, correct for the law to require more than a finding of probability in favour of the person bearing the burden of proof in civil cases, and for that something more to vary having regard to the particular issue involved.

Murphy J considered the issue of standard of proof in \textit{TNT Management Pty Ltd v Brooks}. The respondent’s husband was driving a semitrailer that collided with a pantechnicon on a curve in the highway in New South Wales. She sought damages from the employer of the pantechnicon driver. Both drivers were killed and there was no witness to the accident. Under New South Wales legislation the respondent was entitled to full recovery not only if the driver of the pantechnicon was wholly to blame, but also if he was partly to blame. Gibbs J, with whom Stephen, Mason and Aickin JJ agreed, stated that a finding of negligence could not properly be based solely on general considerations as to the expectation or likelihood of negligence, but that a factual basis for such a decision must be found from the evidence in the case. Because the wreckage of the pantechnicon’s cab had been found on the wrong side of the road, his Honour was able to infer on the balance of probabilities that the pantechnicon’s driver had been guilty of some negligence.

Murphy J dismissed the implications from the wreckage as being too slight and based his decision on probability theory. His Honour reasoned that out of the three apparent possibilities — one driver wholly to blame, the other driver wholly to blame, and each driver partly to blame — two stood in favour of the respondent, so that she ought to succeed on the balance of probabilities. His Honour, however, gave no adequate reason for rejecting a fourth possibility, a true ‘accident’ which was the result of neither driver. Thus a proper application of Murphy J’s reasoning gives rise to four possibilities, only two of which favour recovery, with the consequence that the respondent ought to have failed. In any event, it is not legitimate to attempt to decide cases on the basis of general theories about probability unrelated to the circumstances of the particular case. In reality, Murphy

\begin{itemize}
  \item[91] Cohen, \textit{The Probable and the Provable}, above n89 at 75.
  \item[92] This is not to suggest that considerations of probability should not form part of the process by which a court reasons to a conclusion. In \textit{State Government Insurance Commission v Laube} (1984) 37 SASR 31 the court found that it had not been proved that the defendant had been so much under the influence of intoxicating liquor as to be incapable of exercising effective control of his vehicle. It is suggested that in reaching this conclusion the court gave insufficient weight to evidence that the majority of persons with the blood alcohol content the defendant was shown to have had would have been incapable of exercising effective control of a motor vehicle.
  \item[93] (1979) 53 ALJR 267.
  \item[94] \textit{Law Reform (Miscellaneous Provisions) Act} 1965 (NSW) s10.
\end{itemize}
J’s reasoning was an attempt to reverse the burden of proof in negligence actions in New South Wales, requiring the defendant to adduce evidence to defeat the probability argument. The view of Murphy J was rejected by the majority of the High Court in *West v Government Insurance Office of New South Wales*. The court held that where there is insufficient evidence to establish whether the plaintiff or the defendant or both had been negligent, then there should be judgment for the defendant. The majority stated that inferences as to the conduct of either driver should not be drawn in the absence of evidence to support such inferences.

B. The Argument for a Higher Standard in Certain Civil Cases

In the United States a third standard of proof ‘by clear and convincing evidence’ is recognised, resting midway between proof on the balance of probabilities and proof beyond reasonable doubt. A number of judicial statements have been made which might be taken as suggesting that such a third standard exists in English and Australian law. Thus it has been said that where an allegation of professional misconduct involving an element of deceit or moral turpitude is made, a higher standard of proof is called for. Evidence in rebuttal of the presumption of the formal validity of a marriage must show to ‘a high degree of probability’ that the marriage was not valid. An intention to change domicile must be ‘clearly and unequivocally proved’.  

It is submitted, however, that such statements mean no more than that the inherent unlikelihood of the allegation or proposition put by the party bearing the onus of proof is itself a factor to be placed in the balance scales. For example, it may be thought that a well-respected surgeon or counsel is unlikely to engage in misconduct involving deceit or moral turpitude in her or his professional capacity. Accordingly, a party who makes such an allegation must present comparatively weighty evidence to satisfy the tribunal on the balance of probabilities that such misconduct occurred. The standard of proof does not change, but the quantum or quality of evidence required to meet the standard may.

---


96 *McCormick on Evidence*, above n7 at 796–798.

97 *Bhandari v Advocates Committee* [1956] 3 All ER 742; *Cuming Smith and Co v Westralian Farmers Co-Op Ltd* [1979] VR 129 at 147.

98 *Piers v Piers* (1849) 2 HL Cas 331 at 380; 9 ER 1118 at 1136.

99 *Moorhouse v Lord* (1863) 10 HL Cas 272 at 286; 11 ER 1031 at 1036.

100 See *New South Wales Bar Association v Livesey* [1982] 2 NSWLR 231.

101 Above n76.

102 (1965) 112 CLR 517.
In only one Australian case does the court appear clearly to have adopted a distinct and higher standard of proof for a certain class of civil matter. In *Morrison v Jenkins* the applicants, Mr and Mrs Morrison, brought proceedings to obtain the custody of a child they alleged was theirs. They claimed there had been a confusion of babies at the time of birth, and that they had received the child of Mrs Jenkins while Mrs Jenkins had received their child. The children had been born in a country hospital between five and ten minutes of one another. Blood tests established that the child brought home by Mrs Morrison could not have been the child of Mr Morrison. Mrs Morrison testified that she had not had intercourse with anyone other than her husband and this testimony was not challenged. Mr and Mrs Jenkins refused to have blood tests made, although they were given the opportunity of doing so. At the time of trial the children were three-and-a-half years old.

The trial judge, Barry J, set himself a standard of proof depending not ‘upon a bare balance of probabilities, but as the result of the thorough conviction of my mind founded upon a careful and patient attention to all the evidence in the case.’ This test involved, it would seem, no more than the normal standard of the balance of probabilities, regard being had to ‘the nature and consequence of the fact or facts to be proved’. His Honour found in favour of the Morrisons, and the Jenkins appealed successfully to the Full Supreme Court of Victoria. The Full Supreme Court took the view that the standard set by Barry J was not sufficiently strict. Fullagar J, with whom Herring CJ and Lowe J agreed, stated that no order should be made changing the child’s custody ‘unless it is established as a matter of practical certainty that [she] is the child of Mr and Mrs Morrison’. Herring CJ set the standard of proof as ‘beyond all possible doubt’, and Lowe J spoke of a ‘standard of practical certainty’.

The Morrisons appealed to the High Court where, by a majority of three to two (Rich, Dixon and Webb JJ; Latham CJ and McTiernan J dissenting), the decision of the Full Supreme Court was affirmed. With the exception of Latham CJ, all members of the court set the standard of proof at an exceptionally high level. Rich J stated that the Morrisons must exclude ‘every other reasonable hypothesis’. Dixon J expressed his agreement with the view of the Victorian Full Court. McTiernan J stated that ‘all reasonable doubt’ must be excluded. Webb J was not prepared to go so far. His Honour stated that while a court cannot change the

---

103 Above n86. See also *Murray v Murray* (1959–60) 33 ALJR 521 at 524; *Minister for Business and Consumer Affairs v Evans* (1984) 54 ALR 128.

104 [1949] VLR 277; (1949) 80 CLR 626. For discussion of the case, see Peter Brett, ‘Law in a Scientific Age’ in Norval Morris & Mark Perlman (eds), *Law and Crime* (1972) at 80.


106 Briginshaw v Briginshaw, above n76.

107 [1949] VLR 277 at 304.

108 Id at 298.

109 Id at 299.

110 (1949) 80 CLR 626 at 640.

111 Id at 642.

112 Id at 648.
standard of proof, ‘it can and should insist on exact or cogent proofs on issues of grave importance like that of parentage’.

It is submitted that the view of Barry J was preferable to that taken by the Supreme Court and the High Court. The Supreme Court and the majority of the High Court appear to have been overly influenced by the consideration that a verdict in favour of the Morrisons, even if correct, would involve grave disruption to a young child’s life. Such a consideration does not appear legitimate. Whether it is more harmful to disrupt a young child’s life than to order the continued existence of a family situation that is, on the balance of probabilities, a factually unwarranted one, is pure conjecture.

C. Allegations of Crime in Civil Proceedings

Where a party to a civil case is required to establish the commission of a crime as part of her or his case, the standard of proof that must be met is the ordinary civil standard. In *Helton v Allen* a testatrix died of strychnine poisoning. Helton, to whom the testatrix had left the greater part of her property, was tried and eventually acquitted of her murder. The testatrix’s next of kin, in order to prevent Helton from taking under the will, commenced an action seeking to show that Helton had in fact murdered the testatrix. The trial judge instructed the jury that if they thought the probabilities in favour of the opinion that Helton poisoned the testatrix outweighed in any degree, however slight, the probabilities against that opinion, they should find against him. The jury found against Helton, who appealed to the High Court. The High Court held that the standard of proof required to be met by the next of kin was the civil standard and not the criminal. The court, however, ordered a retrial. Their Honours held that the trial judge had misdirected the jury by failing to instruct them that they had to be satisfied that Helton committed the murder, and that in deciding whether they were so satisfied regard must be had to the gravity of the allegations made against Helton. The rule that it is the civil standard that applies when an allegation of crime is made in civil proceedings has been adopted in numerous subsequent cases.

6. Conclusion

Determining which party carries the burden of proof in civil cases may frequently be a difficult task. What is called for is consideration of the substantial nature of the plaintiff’s cause of action and the defences available rather than merely having recourse to the form in which legal rules are stated or causes of action created by

---

113 Id at 654.
114 (1940) 63 CLR 691.
The statute are expressed. The distinction must be drawn between the legal, evidential and tactical burdens of proof, and in cases of uncertainty regard had to considerations of practical convenience. In expressing the standard of proof, the expression 'balance of probabilities', with the gloss placed upon that expression by Dixon J in Briginshaw v Briginshaw, is well understood and correct in principle. The Briginshaw interpretation of the civil standard gives the court a degree of flexibility in applying the civil standard without requiring the development of intermediate standards between the normal civil standard and the criminal.
Abstract

This paper examines the conceptual confusion inherent in the current formulation of the insanity defence. This, it is contended, results from the differing goals and methodologies of law and psychiatry. It is only through the explication of these differences that a reformulation on a more transparent and principled footing may arise. To this end, the main part of the paper seeks to clarify three central tropes or metaphors informing the current law.

Beginning with ‘mental illness’ I show that the law in adopting a ‘disease’ analogy has failed to adequately bring to light the legal and moral nature of the inquiry. Instead, the pathologising of the issue conceals social defence considerations that underlie the law’s concern.

The concept of ‘rationality’ as the characteristic of mental impairment is shown not to be objectively ascertainable. Rather, cognition is to be understood as a process by which pragmatic solutions are affected. The historicity of this purported norm necessitates that the term ‘rationality’ be jettisoned in favour of ‘cognitive competency’.

Criminal responsibility is analysed from the perspective of what lies behind or produces actions. The complexity inherent in human action is seen to involve not only cognitive competence but also an actor’s volitional state. This, it is argued, should be reflected in any ascriptions of criminal responsibility.

In the light of the analysis undertaken it is proposed that in place of the current common law formulation the defence should be put on a legislative footing in the following terms:

A person is not criminally responsible if at the time of the commission of the offence he or she had a mental impairment which included in its symptoms or consequences a loss of cognitive competency to think of the reasons which people are expected to regard as sufficient grounds for refraining from commission of the offence.

It is believed that such a formulation has the advantage of making explicit the claims of both law and psychiatry: the law’s concern with social defence, and psychiatry’s expertise in the assessment of mental impairment.
1. Introduction

This paper examines the nature of the common law defence of insanity, rooted in the M’Naghten Rules, and cognate related defences. It is concerned with one particular goal: the substantive reformulation of the insanity defence. Ancillary to this is a question of methodology: how to proceed with the task of construction subsequent to a critical analysis of the existing defence.

My approach takes as a guiding principle Austin’s remark that ‘words are our tools, and, as a minimum, we should use clean tools’. As such I will analyse the way in which the tropes of ‘mental illness’, ‘rationality’ and ‘criminal responsibility’ not only describe but also inform the insanity defence. In this way it is hoped to reveal the ‘fit’ between words and the realities they describe. It is through a preliminary process of deconstruction that reconstruction is hoped to be brought about.

The substantive part of the paper involves a critical examination of the terms involved. I begin with an overview of the current law. This is followed by an analysis of the three tropes forming the title of the paper. Finally, a reformulated defence of mental impairment is proposed in the light of the foregoing analysis.

Informing both the choice of topic, and methodology, is the work of contemporary theorists in the area of criminal law and philosophy. Such work attempts to understand the nature of the conflicts surrounding liberal conceptions of criminal law, in particular, the contradictions inherent in viewing criminal law as both protecting individual rights and as a system of social control. Social practices within and around the arena of ‘criminal justice’ are viewed by these new critics not, as liberal and critical theorists would have us believe, as contradictory and confused. Rather, policing, prosecuting, sentencing and punishing are seen to be part of a delicately balanced equilibrium. The maintenance of this equilibrium is seen to depend on ‘holding tensions at bay by exploiting logically contradictory discourses but disguising the fact by moving between different levels of, and spheres for, analysis’.

The insanity defence offers in its engagement of two radically different modes of discourse, that of law and psychiatry, a particularly fertile ground for an examination of the nature of this ‘delicately balanced equilibrium’. The defence is shown to operate by means of a strategy of partial conceptual stipulation, (‘disease of the mind’), and an appeal to ordinary usage (‘knowledge of wrong’). While at one level the defence is confused and contradictory, at another ‘it is part of a complex strategy which allows the criminal law to keep various balls of different shapes and colours in the air at once’. It is with the tracing of these shapes and colours that this paper is concerned.

---

4 Id at 641.
2. Insanity and Related Defences

I begin with a discussion of the substantive law. A general remark on the concept of criminal responsibility is followed by a consideration of the nature of the so-called M’Naghten Rules. This leads to an outlining of the concept ‘disease of the mind’, along with the notions of cognition and volition. The issue of automatism is viewed as a means by which the Rules’ inherent shortcomings and dispositional rigidity are circumvented. I finish this Part with a brief summary of the main points dealt with.

A. Criminal Responsibility

The concept of criminal responsibility rests on the presumption that individuals possess the capacity to make rational choices to act or refrain from acting. Where a person, while understanding the significance of the act, performs it voluntarily and intentionally, he or she will be held responsible. 5

There are some mental states that impair an individual’s ability to meet these requirements. So, where a person behaves in a state of automatism, involuntariness may exculpate him or her from liability. 6 Where, due to a ‘disease of the mind’, an individual is incapable of knowing that what they do is wrong, that person may be excused on the ground of insanity. Similarly, a ‘disease of the mind’, short of insanity, 7 may preclude an individual from forming the requisite intention for a particular crime. 8

B. R v M’Naghten

Cases during the eighteenth century interpreted Hale’s metaphor of the insane being, ‘not as reasonable creatures, their actions [being] in effect in the condition of brutes’, as necessitating a complete absence of reason in order for the defence to apply. 9 It was against such a background that the common law was outlined in the following case.

Daniel M’Naghten was acquitted, on the ground of insanity, of murdering the Prime Minister’s secretary, Drummond, whom he had mistaken for Sir Robert Peel. M’Naghten was under the insane delusion that the Tory party was persecuting him and that his life was endangered. The acquittal created such furore that the House of Lords put certain questions to the judges regarding the scope of the insanity defence, the answers to which became known as the M’Naghten Rules. These form the foundations of the modern law of insanity in the common law world.

---

5 Vallance v The Queen (1961) 108 CLR 56.
8 Hawkins v The Queen (1994) 179 CLR 500.
9 (1843) 4 St Tr (ns) 847.
The judges were asked five questions relating to insanity as a defence. The answers to the second and third questions contain the famous ‘knowledge of right and wrong’ test:

[T]he jurors ought to be told in all cases that every man is presumed to be sane, and to possess a sufficient degree of reason to be responsible for his crimes, until the contrary be proved to their satisfaction; and that to establish the defence on the ground of insanity, it must be clearly proved that, at the time of the committing of the act, the party accused was labouring under such a defect of reason, from disease of the mind, as not to know the nature and quality of the act he was doing; or, if he did know it, that he did not know what he was doing was wrong.11

During the trial of M’Naghten, Tyndal CJ told the members of the jury to acquit if they thought that the defendant did not have sufficient reason to apply the moral distinctions of which he or she may have been aware. 12 In this fashion the complete incapacity attendant upon the judicial interpretation of Hale’s test was reformulated. 13

It is to be noted that the Rules provide those accused with two lines of defence: (i) they must be acquitted if, because of a disease of the mind, they did not know the nature and quality of the act; (ii) or, if they did know the nature and quality of the act, they are to be acquitted if, because of a disease of the mind, they did not know it was ‘wrong’. The primary consideration under either limb is whether the accused was suffering from a ‘defect of reasoning from disease of the mind’. If accused persons were unaware that their act was ‘wrong’ for some other reason, this will generally not amount to a defence, for neither ignorance of the law, nor good motive, normally affords a defence. 14 The M’Naghten Rules have been interpreted as confining the insanity defence within very narrow limits, focusing as they do on the purely cognitive aspect of an accused’s behaviour.

C. ‘Disease of the Mind’

According to the Rules, it must be established that cognitive incapacity, ‘defect of reason’, was caused by a ‘disease of the mind’. As such the mere fact that an accused suffers from impaired reasoning powers is not sufficient. A causal link between this and an underlying ‘disease’ is called for.

The meaning of the term ‘disease of the mind’ is a legal rather than a psychiatric question. 15 The issue is whether the accused’s mental faculties were impaired by illness, not whether he or she was suffering from a recognised mental illness. 16 Insanity is a social judgment founded upon but not precisely representing, a medical diagnosis. 17 Whether a particular condition amounts to a

11 R v M’Naghten (1843) 4 St Tr (ns) 847.
12 Ibid.
13 McAuley & McCutcheon, above n10 at 646. n
14 John C Smith, Smith and Hogan: Criminal Law (8th ed, 1996) at 202 (n81).
15 R v Kemp [1957] QB 399 at 406.
17 Glanville Williams, Textbook of Criminal Law (2nd ed, 1983) at 644.
disease of the mind, such that the issue is suitable to go to the jury, is a question of law for the judge. It is up to the jury to decide the question of whether or not the accused was insane at the time of committing the criminal act. The question will involve consideration of whether expert and other evidence establishes the existence of a disease of the mind.

There need be no actual disease in the sense of a deterioration of the material, physical constitution of the mind. All that is required is that the ‘functions of the understanding are through some cause, whether understandable or not, thrown into derangement or disorder’. The expression will cover all conditions clinically recognised as mental illness, including those with an organic base. Hence, purely physical diseases, which are only of incidental concern to psychiatrists, such as arteriosclerosis, epilepsy, and diabetes, may satisfy the requirements of the term. Ultimately however, ‘the condition of the brain is irrelevant as is the question of whether the condition of the mind is curable or incurable, transitory or permanent’.

To this degree the limits to the notion of ‘disease of the mind’ are determined by the concept of pathology. ‘Defects of reason’ stemming from an underlying pathological condition fall within the ambit of the concept. Those ‘conditions of intense passion and other transient states attributable to the fault or to the nature of man’ are excluded from it.

A non-pathological factor may be viewed more as ‘the reaction of a healthy mind to extraordinary external stimuli’.

**D. Cognition and Volition**

There are two limbs to the cognitive component of the Rules (‘defect of reason’): knowing the ‘nature and quality’ of the act, and a knowledge of wrongfulness.

At common law the expression ‘nature and quality’ of the act has been taken to refer to the physical nature and consequences of an offence rather than to its moral aspects. It is important that the accused does not know what he or she is doing, or does not appreciate or foresee the possible effects of his or her conduct. Examples include situations where the accused is mistaken about the physical nature of the act, for example, where an insane person cuts another’s throat.

---

19 *R v Porter* (1933) 55 CLR 182 at 188.
21 *R v Porter* (1933) 55 CLR 182 at 189.
23 *R v Kemp* [1957] QB 399.
26 *R v Kemp* [1957] 1 QB 399 at 407.
27 McAuley & McCutcheon, above n10 at 683.
28 Dixon, above n22 at 260.
believing it is a loaf of bread that is being cut.\footnote{31} Another example is where the accused decapitates someone just to see what the person looks like without a head.\footnote{32} It will be rare that a defendant’s mind is so disturbed as to preclude them from realising the physical nature and quality of the acts and it is normally under the second limb that insanity pleas have been made.

The requisite knowledge of wrongfulness poses an ambiguity: ‘wrong’ according to law, or according to the standards adopted by reasonable people? The question is whether the accused was able to appreciate the wrongness of the particular act at the particular time. It has always been clear that if he or she knew that their act was contrary to law, he or she knew it was wrong for this purpose. The accused is also liable where, while not knowing the act to be contrary to law, knows it was wrong ‘according to the ordinary standard adopted by reasonable men’.\footnote{33}

The High Court, in \textit{Stapleton v the Queen}\footnote{34} considered the situation where the accused knows the act to be contrary to law, but believes that he or she is doing right, and that people would regard the act as right. The Court held that if the accused believed that his or her act was right according to the ordinary standards adopted by reasonable persons, he or she would be entitled to an acquittal. So, wrong means wrong according to the ordinary principles of reasonable people, rather than contrary to law.\footnote{35} The Court held that the main question is whether or not the accused was able to appreciate the wrongfulness of the criminal act, such that he or she was able to ‘reason about the matter with a moderate degree of sense and composure’.\footnote{36} Where the accused could not so reason, it may be said that the accused could not know that what he or she was doing was wrong.

Where an accused satisfies the cognitive aspect of the defence, that is, knew what he or she was doing and that it was wrong, the claim that due to a disease of the mind he or she could not prevent themselves from doing the act in question, will not satisfy the defence.\footnote{37} If as a result of the condition of irresistible impulse the defendant did not know either the nature and quality of the act or that it was wrong, then the defence of insanity is open.\footnote{38} For, as the High Court held in \textit{Attorney-General (SA) v Brown}\footnote{39} ‘uncontrollable impulse … may afford strong ground for the inference that a prisoner was labouring under such a defect of reason from disease of the mind as not to know that he was doing what was wrong.’\footnote{40}

\begin{thebibliography}{99}
\parskip=0ex
\bibitem{33} \textit{R v Codere} (1916) 12 Cr App R 21 at 27.
\bibitem{34} (1952) 86 CLR 358 (approving \textit{R v Porter} (1933) 55 CLR).
\bibitem{35} Contra \textit{Windle v The Queen} [1952] 2 QB 826; See Smith, above n14 at 209.
\bibitem{36} \textit{Stapleton v The Queen} (1952) 86 CLR 358 at 367 citing \textit{R v Porter} (1933) 55 CLR 182 at 189 (Dixon J).
\bibitem{37} \textit{R v Porter} (1933) 55 CLR 182.
\bibitem{38} \textit{Sodeman v The Queen} (1936) 55 CLR 192.
\bibitem{39} (1959) ALR 808.
\bibitem{40} Id at 814.
\end{thebibliography}
appeal to the Privy Council\textsuperscript{41} it was held that for the defence to succeed there must be medical evidence to support the conclusion that irresistible impulse or defect of volition proceeds from mental illness.\textsuperscript{42}

\textbf{E. Automatism}

Strictly meaning ‘action without conscious volition’, automatism has come to denote in the criminal law, ‘conduct of which the doer is not [fully] conscious’.\textsuperscript{43} Acts that occur during this time are considered to be involuntary movements of the body rather than ‘acts’ for the purpose of liability. The question is whether there was an absence of all the deliberative functions of the mind so that the accused person acted automatically. Consciousness is not necessarily fatal to a plea of automatism.\textsuperscript{44}

Automatism requires evidence to be disproved by the prosecution. It has arisen as a result of the combination of two pressures.\textsuperscript{45} First, it has arisen because of the limitation in the M’Naghten Rules, that the exculpatory mental state of the accused must proceed from a disease of the mind. Secondly, it has developed because a verdict of insanity is not an outright acquittal, but in fact subjects the defendant to indefinite detention, a situation clearly not called for in certain situations. In effect, automatism stems from a conceptual failure to define insanity sufficiently widely, along with an unwillingness to adopt a more flexible approach to issues of disposition of those acquitted.\textsuperscript{46}

The distinction between insanity and automatism will turn on whether the state of unconsciousness proceeds from a disease of the mind. If it proceeds from such, the defence is one of insanity or insane automatism,\textsuperscript{47} which is irrelevant to the question of voluntariness. If it does not derive from a disease of the mind automatism is available as a ‘defence’. The implications of this distinction are apparent at both a substantive and procedural level. In the instance of sane automatism, the defendant is unconditionally acquitted, while in the instance of insane automatism the defendant is usually subject to detention in a psychiatric institution.

In the instance of insane automatism the onus will be on the defendant to prove, on the balance of probabilities, insanity at the time of the offence. In the case of sane automatism the defendant will only have to satisfy an evidential burden on the issue of involuntariness, the persuasive burden will then be on the prosecution, to negative the assertion of involuntariness beyond a reasonable doubt.

\begin{footnotesize}
\begin{itemize}
\item[41] [1960] AC 432 at 449–450 (Lord Tucker).
\item[43] \textit{R v Cottle} [1958] NZLR 999 at 1077 (Greeson P).
\item[44] \textit{Ryan v The Queen} (1967) 121 CLR 205 at 214 (Barwick CJ).
\item[46] Id at 719.
\item[47] \textit{Rabey v The Queen} [1980] 2 SCR 513 at 524 (Dickson J); \textit{R v Falconer} (1990) 171 CLR 30.
\end{itemize}
\end{footnotesize}
The conditions giving rise to sane automatism include reflexive reactions, spasms, and convulsions. In such situations, quite clearly, the accused's bodily movements are not the result of an exercise of choice on his or her part but are produced by some other agency. Instances, which are more problematic, are those where the alleged lack of voluntariness stems from a mental state. Some such instances are concussion caused from a blow to the head, sleepwalking, the consumption of alcohol and drugs, hypoglycaemia, or dissociation arising from extraordinary stress and epilepsy. It should be noted that evidence of some degree of control over bodily movements does not preclude automatism. Where the actions of the accused may be said to be involuntary, the cause of the action does not matter.

**F. Summary**

The following points may be noted. First, criminal responsibility is held to be predicated on the capacity to make rational choices to act or refrain from acting. Such a capacity must be exercised in a voluntary and intentional fashion. This paradigm will be considered below in Part 5.

Secondly, in relation to mental illness, the Courts' willingness to interpret the Rules in a more liberal fashion is in fact an acknowledgment of the interrelationship between cognition and volitional impairment. This issue brings to the fore the observation that 'insanity' is a social judgment founded upon a medical diagnosis. The notion of 'disease of the mind', while determined by the concept of pathology, will be shown to be, in more borderline situations, far more contingent. Norms will be seen to be less descriptive than prescriptive. This will be considered below in Part 3.

Thirdly, the concept of the rational, autonomous individual, which underpins ascriptions of criminal responsibility, is in the instance of automatism severely stretched. Necessitated by the rigidity of interpretation accorded to the Rules and strict dispositional orders, sane automatism in fact requires the fiction of 'involuntary goal-directed behaviour'. This has no foundation in scientific research nor does it accord with our moral intuitions. This highlights the conceptual difficulties at the heart of the trope of rationality. This will be considered further in Part 4.

**3. ‘Disease of the Mind’**

In this part I will consider the notion of ‘disease of the mind’. I will begin with the way the law employs the concept of ‘disease’. In particular, I will focus on the way

---

48 *Re Wakefield* (1958) 75 WN (NSW) 55.
in which the law’s failure to provide a definition in terms of moral responsibility has necessitated recourse to the notion of ‘dangerousness’. I will then consider the concepts of ‘disease’, ‘health’, and ‘illness’, and the extent to which these are sociogenic constructs. Finally, I will turn to psychiatry’s ‘biopsychosocial’ model of mental disorder. It will be shown that this model, while an improvement on the legal conception, suffers from conceptual confusion. The model seeks to describe syndromes in psychological terms, while relying on physiological mechanisms to explicate underlying causes. A possible response to these difficulties is put forward by way of conclusion.

A. Law and ‘Disease’

(i) Legal Usage of the Term

For several reasons the law of insanity is not coextensive with the psychiatric conception of mental disorder.\(^{55}\) First, legal insanity is an excuse for wrongdoing, not a diagnosis of the accused’s mental condition. Excuse of an insane individual is predicated on his or her inability to act rationally, that is, to understand what he or she is doing and why, and to act on the basis of that understanding. As such the scope of the insanity defence in not an empirical matter that can be settled by psychiatrists.\(^{56}\)

Secondly, the legal framework begins with the premise that the function of understanding is through some cause, whether understandable or not, thrown into derangement or disorder,\(^{57}\) which may or may not be due to some organic cause. The mere fact of organic damage alone is insufficient. It is the fact of disorder which is decisive, not the cause of it. The legal definition seeks to express a criterion by which an individual may escape responsibility, and further, to exclude those things for which an accused should not be excused, such as violent tempers and so on.\(^{58}\) In contrast, the psychiatric conception of mental illness is tied to diagnostic categories made up of clinically significant behavioural or psychological syndromes or patterns that occur in an individual. A common misconception is that a classification of mental disorders classifies people, when actually what are being classified are disorders that people have.\(^{59}\) The legal enquiry is whether illness exonerates the defendant from blame for his or her actions, not whether he or she is suffering from a recognised mental illness.\(^{60}\)

Thirdly, the law requires that the mental disorder caused the wrongdoing.\(^{61}\) As the defence of insanity functions as an excuse for wrongdoing, not as a diagnosis of mental condition, the behaviour must be shown to be connected with the

---

\(^{55}\) McAuley & McCutcheon, above n 10 at 638–644.

\(^{56}\) Jennifer Radden, Madness and Reason (1985) at 28.

\(^{57}\) R v Porter (1933) 55 CLR 182 at 189 (Dixon J); R v Kemp [1957] QB 399 at 407 (Devlin J).

\(^{58}\) Ian G Campbell, Mental Disorder and Criminal Law in Australia and New Zealand (1988) at 128.


\(^{60}\) McAuley & McCutcheon, above n 10 at 640.

\(^{61}\) R v Kemp [1957] QB 399 at 407 (Devlin J).
psychological state. The psychiatric conception of mental disorder is tied up with the problem of diagnosing and treating severe mental disorders. Psychiatric nosology however, does not carry with it any necessary implications regarding the causes of an individual’s mental disorder. These categories are, moreover, in a constant state of flux, open to revision, reformulation, and challenge.

Fourthly, the common law insanity defence is not equivalent to the so-called ‘product test’ which holds that an accused has a good defence to a criminal charge ‘if his unlawful act was the product of mental disease or mental defect’. The fallacy in this view rests on a conceptual misunderstanding of the way in which diseases affect action. On the basis of the ‘product test’ were a brain lesion to cause a violent impulse, regardless as to whether the defendant could have resisted the impulse, he or she would be free to plead the defence. This however is to confuse causation with excuse. Causation is not an excuse, for all behaviour is caused. If causation were an excuse, no one would be held responsible for any behaviour, criminal or not. Neither is causation the equivalent to compulsion, a species of excuse.

In the instance of physical diseases, to be excused from blame for a disease entails our being excused from blame for actions stemming from that condition. No responsibility accrues here for the simple reason that responsibility presupposes agent causation. An individual’s movement of a limb during an epileptic seizure is morally neutral, and is described as a movement rather than an action. In the instance of a mental disorder however a different type of causation is at work.

Mental ‘diseases’ affect behaviour through the medium of psychological states, such as thoughts, beliefs and desires. Even if one were able for example, in the case of schizophrenia, to uncover an underlying brain state, it would not follow that the person so suffering is not responsible for the ‘symptoms’ of the illness. This would apply only where the claim ‘Madness is not blameworthy’ is held to entail the belief that ‘madness mitigates blame for wrongdoing’. That is, we may disapprove of a person’s misbehaviour even where we recognise that it is entirely in character. While we do not blame an individual for having the characteristics that he or she possesses, we do blame him or her for acting in character, since we assume that ordinary people are capable of keeping the negative elements of their character in check.

---

64 *Durham v United States* 214 F 2d 862 (1954), 874.
65 McAuley & McCutcheon, above n10 at 641–644.
67 See Radden, above n56 at 30–33.
68 McAuley & McCutcheon, above n10 at 643.
(ii) ‘Dangerousness’

Three tests have been formulated judicially in response to the threshold question of did the accused at the time of the commission of the offence suffer from a ‘disease of the mind’?:

- **The recurrence test**\(^70\) which holds that if a mental state is likely to recur it should be considered a disease of the mind;
- **The internal/external test**\(^71\) which provides that if the mental state is ‘internal’ to the accused, as opposed to arising from an external cause, it should be defined as a disease of the mind;
- **The sound/unsound mind test**\(^72\) where a disease of the mind is considered on this test to be evidenced by the reaction of an unsound mind to its own delusions or external stimuli.

It is a policy-oriented question that underlies the legal interpretations of the term ‘disease of the mind’. Namely, is this individual likely to cause harm to others if not confined? If so, then his or her mental condition will be considered to be a disease of the mind for the purposes of the insanity defence.\(^73\)

This is overt in the ‘recurrence test’ with its assessment of the potential dangerousness of an individual as the ground for determining his or her suitability for release. It is also implicit in the other tests, based as they are on the assumption that only those with unsound minds or with mental conditions arising from some internal source should be detained.\(^74\)

As a criterion of mental illness this concern with the disposition of those considered dangerous confuses two issues: the utilitarian issue of what to do with an acquittee who may be dangerous in the future and the moral question of the conditions for ascription of criminal responsibility.\(^75\) It is important to recognise that the question of disposition is irrelevant to what the accused did in the past. As such, dangerousness and treatability are incorrectly considered in relation to issues of responsibility as they relate to past acts.

The law has yet to fashion a construct of disease, which reflects concepts of responsibility. At present ‘disease of the mind’ is understood as a disorder of the reasoning process from some cause (in whole or in part internal to the individual) which vitiates responsibility, if it produces deprivation of any of the capacities. This has required resort to the question of future dangerousness and the disposition of the individual accused as a limiting factor. Campbell argues that while ‘[t]his may be social policy… it is not a legal principle of moral culpability or responsibility’\(^76\).

---

69 Bernadette McSherry, ‘Defining What is a “Disease of the Mind”: The Untenability of Current Legal Interpretations’ (1993) 1 JLM 76 at 82–89.
74 McSherry, above n69 at 89.
75 Campbell, above n58 at 129.
76 Id at 132.
B. Concepts of ‘Disease’, ‘Health’ and ‘Illness’

A theory of medical norms may hold that the body is composed of certain structured systems (forms) each of which has an assignable range of normal functioning. Defect or disorder of such systems relative to such functioning constitutes a sufficient condition of disease. Yet the precise nature of this functioning, in terms of which disease is to be specified, cannot be determined by noting a series of observable form/function correlations. For example, anginal pain may arise from pulmonary hypertension, not coronary disease. The notion of normal or healthy functioning cannot be simply assigned to the localised processes and structures of the body. What is normal must be construed not as a fixed point but as a range of variations.

Both general and psychiatric medicine, insofar as they suppose themselves to be value-neutral sciences, presuppose that human beings have a set of ‘natural’ functions. This set of ‘natural’ functions, physical and psychological, is said to be essential to our respective natures, not merely accidental. However, the seemingly normative nature of medical discourse stems from a failure to distinguish the human animal from the human person. It is impossible to speak of psychiatric disease and disorder exclusively in terms of the condition of the human animal. The disorders are formulable only in terms of the mental processes of culturally emergent persons. As such any ascription of ‘natural functions’ to human persons should not be provided, in the context of psychiatry, in a way that ignores the culturally prepared goals of human societies.

Such goals may be formulated in terms of ‘rational minima’, that is, requirements considered least controversial or objectionable to agents endowed with a minimal measure of rationality and which are more or less presupposed by every significant human society. Such minima may include the norms of health and disease. They amount in effect to prudential values to which persons subscribe: avoidance of death, prolongation of life, restriction of pain, gratification of desires, assurance of security of person, body, property, and the like.

In the field of medicine, disease is whatever is judged to disorder or to cause to disorder, in the relevant way, the minimal integrity of body and mind, relative to these prudential values. Conversely, the criterion of health is the adequate performance of functions, physiological and psychological, in the service of these prudential values.

While performance may be socially assessed, the criteria of health are not primarily social. It is misconceived to equate ill health with social deviation or maladjustment.

---

79 Compare DSM–IV–TR, above n59 at 297, 309.
80 A Lewis, ‘Health as a Social Concept’ (1953) 4 British Journal of Sociology 109 at 122.
81 Gelder, Gath & Mayou, above n62 at 106.
82 Margolis, above n77 at 251.
83 Lewis, above n80 at 124.
mentally ill” implies “Smith is socially maladjusted in the society concerned”, means that the latter claim states only a necessary condition for the former to have relevance. It does not state a sufficient condition, because ‘Smith is mentally ill’ may be false. Nor does ‘Smith is socially maladjusted in the society concerned’ state either a necessary, or sufficient condition, for the truth of ‘Smith is mentally ill’. Smith can be a social misfit in the society concerned without it being either relevant or true to say that Smith is mentally ill.84

C. Psychiatry and ‘Disease’

(i) Methodological Overview

The Diagnostic and Statistical Manual of Mental Disorders (DSM–IV–TR), replaces the term ‘mental illness’ with ‘mental disorder’ which it defines as ‘a clinically significant behavioural or psychological syndrome or pattern that occurs in an individual and that is associated with present distress (for example, a painful symptom) or disability (that is, impairment in one or more important areas of functioning) or with a significantly increased risk of suffering death, pain, disability, or an important loss of freedom’.85 Mental disorders may be defined on various levels of abstraction, for example, syndromal pattern, dysfunction, aetiology, each of which is an indicator for a mental disorder, but none of which is equivalent to the concept.86

Psychiatry has clearly adopted a ‘biopsychosocial’ model, which includes recognition of the importance of social, psychological, and biological factors in the aetiologies of mental disorders.87 In the instance of an individual with schizophrenia, characteristic signs and symptoms include both ‘positive’ symptoms, reflecting a distortion of natural functions, and ‘negative’ symptoms, reflecting a diminution or loss of normal functions, persisting for a period of at least six months (Criterion A and C). No single symptom is pathognomonic of schizophrenia; the diagnosis involves the recognition of a constellation of signs and symptoms. The individual must experience dysfunction in one or more areas of functioning (Criterion B: for example, interpersonal relations, work or education, or self-care). Further, the disturbance must not be due to the direct physiological effects of a substance or a general medical condition (Criterion D and E).88

As discussed above, a theory of medical norms may hold that the body is composed of certain structured systems. These ‘regulative principles’ require viewing a human being as an entity whose operating mechanisms conform, more or less, to the ‘natural’ design of the type of organism to which it belongs.89 Where such functioning is found to be outside what is accepted as the normal range, then this is taken as prima facie evidence of internal malfunctioning. This is expressed

84 Farrell, above n78 at 24.
85 DSM–IV–TR, above n59 at xxxi.
86 Id at xxx.
as something having gone wrong with the organic machinery of the system. When we speak of psychological reactivity, we are committed to suppose that certain sorts of machinery are at work, and responsible for these reactions. However, unlike the somatic case the details of their operations, or what these mechanisms are, remain uncertain.  

Models may be used where underlying causes of relationships, said to exist, are still unknown; that is, when a real theory is not available. Reliance on a statistical model, with its inherent concept of deviancy from a norm, provides psychiatry with the necessary support to buttress complex theories and hypotheses. Specific mechanisms are postulated which are formal, or merely structural, in nature. In any ascription of schizophrenia, for example, an empirical statement is being asserted that not only describes and classifies an individual’s clinical manifestations, it also helps to explain them. For the diagnosis maintains that the clinical picture is the manifestation of the impaired or morbid working of the individual’s psychological machinery.

With mental illness, such formal systems tend to be metaphorically identified. Norms assign operative characteristics to human nature rather than to somatic entities. In this regard, the norms of health and disease tend to correspond with putative norms of ‘happiness’ and ‘non well-being’.

Mental illness is not a straightforward, empirical concept of the sort we find in natural science. Nomological (causal) explanation is roughly the standard in the natural sciences while some kind of hermeneutic understanding seems crucial for social sciences. However, in psychiatric settings theoretical constructs tend to be confused with clinical observations, ‘understanding’ mistaken for ‘explanation’.

---

88 DSM–IV–TR, above n59 at 298–302:
- **Criterion A: Characteristic Signs and Symptoms**
  - Positive: distortions in thought content (delusions) perception (hallucinations) language, and self monitoring of behaviour
  - Negative: affective flattening, alogia, and avolition
- **Criterion B: Marked social or occupational dysfunction**
- **Criterion C: Persistence of symptoms for at least 6 months**
- **Criterion D: Disturbance not better accounted for by other Disorder**
- **Criterion E: Not due to other physiological effects**
  - Not due to the direct physiological effects of a substance or a general medical condition

89 Farrell, above n78 at 27.
91 Id at 17.
92 RE Kendell, ‘The Concept of Disease and its Implications for Psychiatry’ (1975) 127 *British Journal of Psychiatry* 305 at 309.
93 Moore, above n66 at 281.
94 Gelder, Gath & Mayou, above n62 at 100.
(ii) Szasz’s Critique

Szasz’s work has sought to expose the categorical error involved in the conflation of sets of dichotomous terms such as, ‘disease’ or ‘illness’ on the one hand, and, on the other, what he considers to be descriptions of personal behaviour, such as ‘schizophrenia’. Szasz contends that the former terms, by definition, require a demonstrable histopathology or pathophysiology. The suggestion that schizophrenia is a ‘mental illness’, in the absence of an aetiological basis, which, according to Szasz, does not, and cannot exist, gives rise to the claim that psychiatric nosology is used to aid and abet a ‘myth of mental illness’. Various responses have been made to these claims, two of which will be considered here.

First, a belief that mental illness is a disease process, with an organic or biochemical underlying cause, with certain symptoms, and specific treatments and prognosis, underlies the ‘medical model’ of psychiatric aetiology. The model is best suited to organic states with discoverable underlying causes of physical malfunction. Epistemologically, neither individual symptoms nor a syndrome is sufficient to infer the presence of a particular disease state. As discussed in the preceding section, there is no precise form/function correlation for all diseases allowing abnormal structure and abnormal function to be compared. Hence, it is a presumption that underlying states of malfunction are present on the basis of symptomatology.

The medical model when applied to psychological manifestations or symptoms presupposes that, for example, the ‘disease’ of schizophrenia refers not only to the psychological states making up its symptoms but also to unknown physical states of malfunction producing them. In fact, even in the instance of ‘operational’ conditions, wanting any explanation in organic terms, characterised rather by a set of psychological and behavioural manifestations, the expression ‘disease’ is employed. The confusion inherent in such use of metapsychological language is succinctly stated by Radden who observes that ‘a metaphorical extension has come to acquire a literal meaning’. In effect, psychiatric explanations are held to give causes of mechanical movements or processes, rather than giving reasons for actions. A more cautious convention would view the medical model as presupposing an analogy, not an identity, between madness and physical disease: madness is a ‘disease’ not a disease.

99 DSM–IV–TR, above n59 at 305.
100 Radden, above n56 at 18.
101 Moore, above n66 at 281.
102 Radden, above n56 at 18.
Secondly, McAuley views Szasz as seeking to exploit a conceptual difficulty lying at the heart of contemporary discourse in psychiatry, by means of a false dichotomy held to exist between the terms ‘organic’ and ‘behavioural’.

Diseases need not be characterised by a clearly discontinuous change in bodily functioning, traceable to some discrete cause. They may be systemic, having their origins in the internal state of the organism, and arise as a transformation and breakdown of otherwise normal functions. In the domain of mental impairment, we may refer to genetically influenced variations in brain organisation as dispositions to varying forms of mental disorder. Where these biological dispositions appear at the extreme end of the spectrum of abnormality, we find the ‘disease’ entities of psychiatry becoming clearly defined. This re-conceptualisation of disease broadens it to include, in addition to physical lesions and their symptoms, a constitutional proneness or predisposition. Even in the absence of an underlying brain dysfunction, there may exist an operative tendency-to-disorder in addition to the prospect that an organic substructure may eventually be discovered.

It is only a wider concept of mental illness which is called for at this descriptive phase of psychiatry’s development, one not tied down to known organic pathology and/or aetiology. Szasz’s dichotomy of ‘organic’ and ‘behavioural’ is, on this view, a misdescription of modern psychiatric thought and practice.

Fulford argues that viewing mental impairment along the lines of a scientific model of bodily disease, a ‘science-half-field-view’, is not so much wrong as incomplete. A ‘full-field’ view, by contrast, would recognise, in addition to facts and failure of functioning, value and failure of action.

The evaluative connotations of mental illness reflect differences in our evaluations of mental phenomena, in contrast to the relatively uncontentious value judgments in physical illness. Yet these connotations represent legitimate variations in what counts as mental illness. Rather than leading to the medicalisation of life’s problems, as Szasz would contend, this non-descriptivism makes explicit the evaluative logical element in the diagnosis of mental illness. As such, it sounds a warning for any immediate descriptive attitude sought to be advanced.

In the full-field view, mental impairment is to be understood not as defective cognitive functioning, but rather as a defect in ‘reasons for action’, that is, a defect in ‘practical reasoning’, or a defect in the reasons we have for the things we do. Actions are defined by means of the reasons for which they are done. If the reasons themselves are defective, there is a constitutive, rather than merely executive, failure of action. There is, in this sense, no action at all. This will be considered in Part 4.

103 McAuley, above n97 at 68–73.
104 Roth, above n97 at 321–322.
105 McAuley, above n97 at 67.
106 Id at 67–68.
107 Id at 290–291.
108 Id at 298–299.
109 Id at 298–299.
D. Reform

The above discussion of the term ‘disease of the mind’ has brought to light the following key issues. First, differences in the nature of the enquiry with which law and psychiatry are concerned were noted. In particular the fact that legal insanity is an excuse for wrongdoing, not a diagnosis of the accused’s mental condition. Policy considerations were seen to inform the legal definition of ‘disease of the mind’. These social defence considerations lead to dispositional issues being decided on grounds of ‘dangerousness’. That is, what the accused is likely to do in the future, not what he or she may have done in the past.

The psychiatric conception replaces ‘mental illness’ with ‘mental disorder’. Adopting a biopsychosocial model, diagnostic categories serve both a descriptive and explicative purpose. While an improvement on the legal explanation, psychiatry also adopts an essentially scientific model of bodily disease. Szasz’s argument attacks the conceptual confusion inherent in the formulation of mental processes in terms of mechanistic function.

While the concept of ‘disease’ is readily applicable to physical states, being a defect or disorder of a structured system from an assignable range of normal functioning, it has only a metaphoric application to mental disorders. It is for this reason that in any reformulation of the insanity defence the term ‘disease of the mind’, or its cognate predicate ‘mental illness’, be jettisoned. The legal or moral question is whether or not mental responsibility was substantially impaired. Functional impairment is the feature characterising those we describe as insane. The nature of this dysfunction is best understood in terms of a defect in practical reasoning, or reasons we have for the things we do. Accordingly, any reformulation of the insanity defence should incorporate the following definitional element:

‘Mental Impairment’ is to be understood as a defect in practical reasoning, characterised by a distortion of natural functions and a diminution or loss of normal functions.

‘Natural’ denotes functions with either an organic or non-organic basis. ‘Normal’ is to be defined in terms of societal ‘prudential values’ as discussed above. In this way both the evaluative and normative aspects of psychiatry’s diagnostic categories are made explicit, without recourse to mechanistic explanations. Additionally, our moral intuition that agents have some capacity to act for reasons is clearly stated.\(^{110}\) The concept of ‘practical reasoning’ will be explored in the next section.

4. ‘Rationality’

In its focus on symptoms, the medical model views the M’Naghten Rules formulation, ‘defect of reason’, as applying to the intellectual operations, or ‘cognitive functions’, of the mind.\(^{111}\) ‘Mental illness’ has been characterised as an

\(^{110}\) Moore, above n66 at 244–245.

\(^{111}\) Moore, above n66 at 244–245.
extreme and prolonged inability to know and deal in a rational and autonomous way with oneself and one’s social and physical environment. ‘Autonomy’ is defined as ‘having and freely actualising a capacity for making one’s own choices, managing one’s own practical affairs and assuming responsibility for one’s own life’.\footnote{Rem B Edwards, ‘Mental Health as Rational Autonomy’ in Edwards, above n98 at 70.} ‘Autonomy’ will be examined in Part 5 in relation to ‘criminal responsibility’. In this Part, I will consider the meaning of the term ‘rationality’.

The first section will seek to outline the process of ‘practical reasoning’, which several authors have sought to employ in explicating the concept of rationality. The second will attempt to show the precise manner in which the concept is used in the context of mental impairment. Finally, I will turn to a critical examination of the concept of ‘rationality’.

A. ‘Practical Reasoning’

The ability to reason has come to be associated with that of decision-making. This accords with the interdependence of the two abilities, their ‘mutual recursiveness’.\footnote{Philip N Johnson-Laird & Eldar Shafir, ‘The Interaction Between Reasoning and Decision Making: An Introduction’ (1993) 49 Cognition 1 at 2.} This account is also integral to ‘folk psychology’, the view most individuals accept about mental life. Most people believe that their conscious feelings and judgments control their actions, and that these are readily available to them through introspection.

People use practical reasoning to try to achieve their goals in the actions they perform, and when they do this they display what has been termed, a ‘rationality of purpose’.\footnote{Jonathan StBT Evans, David E Over & Ken I Manktelow, ‘Reasoning, Decision Making and Rationality’ (1993) 49 Cognition 165 at 169.} This pattern of practical reasoning can be expressed as follows:

1. x desires q (for example, let it be the case that the wine be chilled).
2. x believes that doing p will produce q (for example, if I open the window, then the wine will be chilled).
3. x does p (for example, I open the window).

If in fact x does p because of a desire for q and in the belief that doing p will produce q, then we can say that p was an intentional action explained by x’s belief/desire set. When a person acts according to this type of practical syllogism, he or she causes an action, which is rationalised by his or her desires and beliefs.\footnote{Moore, above n66 at 13–14.} An explanation rationalises an act when it portrays that act as the rational thing to do given the actor’s beliefs and desires.

A concern of those supporting this analysis is to show that the legal conception of a person is the same as the moral one. Both conceptions are instantiated by the notion of the ‘person as practical reasoner’. Acts are explained in terms of personal agency, a person bringing about some state of affairs, rather than through purely

\footnotesize
\begin{itemize}
\item \footnote{Herbert Fingarette, ‘Insanity and Responsibility’ (1972) in Edwards, above n98 at 504.}
\item \footnote{Rem B Edwards, ‘Mental Health as Rational Autonomy’ in Edwards, above n98 at 70.}
\item \footnote{Philip N Johnson-Laird & Eldar Shafir, ‘The Interaction Between Reasoning and Decision Making: An Introduction’ (1993) 49 Cognition 1 at 2.}
\item \footnote{Jonathan StBT Evans, David E Over & Ken I Manktelow, ‘Reasoning, Decision Making and Rationality’ (1993) 49 Cognition 165 at 169.}
\item \footnote{Moore, above n66 at 13–14.}
\end{itemize}
causal accounts.\textsuperscript{116} What results is a theory of rational agency, serving legal and moral interests, framed in terms of the actor’s reasons for acting.

This theory seeks to attribute both logical and causal force to practical reasoning.\textsuperscript{117} The logical force is found in the wants represented in the first premise, in conjunction with the beliefs about the effective means to fulfilling those wants included in the second premise. These provide the reasoner with good reason (rationalisations) to want to perform the action indicated by the conclusion. The causal force is found in the agent’s wanting to achieve the end stated in premise 1 and he or she believing that the action identified in premise 2, will provide an effective means to that end. This belief will cause him or her to decide to act on that want or directive, represented by 3, as a means to satisfying the want stated in 1. The process represents a pattern of inference by which the individual reasons from his or her beliefs and desires for acting to an intentional act. The act is identified as an effective means, to the end represented, by his or her reasons for acting.

In essence, practical reasoning’s central premises are: (1) that whether a bodily movement is an action depends on what a person knows he or she can do and on what he or she knows him or herself to be doing; and (2) that reasons (belief/desire sets) can both cause and rationalise behaviour independently of other factors.

I will now turn to an examination of the implications that this analysis of action has for questions of mental impairment.

\textbf{B. ‘Practical Reasoning’ and Mental Disorder}

\textit{(i) The Nature of Mental Impairment}

Affective disorders without substantial cognitive dysfunction will not ground a mental impairment defence.\textsuperscript{118} As such this enquiry will focus on the paradigmatic case of psychotic disorders such as schizophrenia. DSM-IV-TR defines the differential diagnosis of these disorders to include psychotic symptoms. Psychosis ‘refers to delusions, any prominent hallucinations, disorganised speech, or disorganised or catatonic behaviour’.\textsuperscript{119} Delusions are defined as ‘a false belief based on incorrect inference about external reality…[in the face of] incontrovertible proof or evidence to the contrary.’\textsuperscript{120} An ‘hallucination’ is defined as ‘a sensory perception that has the compelling sense of reality of a true perception but that occurs without external stimulation of the relevant sensory organ.’\textsuperscript{121}

\begin{flushleft}
\textsuperscript{116} Herbert Fingarette, \textit{The Meaning of Criminal Insanity} (1972) at 196–197. \\
\textsuperscript{117} Robert F Schopp, \textit{Automatism, Insanity and the Psychology of Criminal Responsibility} (1991) at 118–119. \\
\textsuperscript{119} DSM–IV–TR, above n59 at 297. \\
\textsuperscript{120} Id at 821. \\
\textsuperscript{121} Id at 823.
\end{flushleft}
Inferences about disorganised thinking are based primarily on the individual’s speech. This may manifest as ‘loose associations’, slipping from one topic to another; ‘tangentiality’, where answers to questions are either obliquely related or completely unrelated; and in severe cases, disorganised ‘word salad’. Catatonic symptoms, however, are non-specific and may occur in other mental disorders.

Major disorders of thought processes involve disturbances in three areas: cognitive focus, reasoning and concept formation. Cognitive focus is the failure to effectively select relevant aspects of a stimuli field or to adjust attention in response to changing situations. Disturbed reasoning involves over-generalised thinking, the drawing of conclusions without evidence and combinative thinking, an inappropriate condensation of impressions and ideas into beliefs and conclusions. Concept formation disturbance involves ‘over-inclusiveness’: ‘the inability to conserve conceptual boundaries with the result that there is an incorporation of irrelevant ideas’. This failure to distinguish between false and veridical perceptual experiences and beliefs has been termed a failure of ‘reality testing’.

(ii) Rationality and Thought Disorder

Mental impairment has been viewed as constituting a failure of epistemological justification. This maintains that beliefs are justified if they are the product of correct cognitive processes, where the correctness of these foundations for beliefs is an inherent feature of the cognitive processes involved. Delusional thought processes will, in the absence of such justification, generate ‘false’ beliefs. It is the way a belief is formed and held, not its truth or falsity, which will determine its irrationality.

It will be noted that this criterion relates knowledge not to the content of an individual’s beliefs but rather to cognitive process. An effective practical reasoner is one who reasons, from wants and beliefs, to necessary and sufficient conditions for satisfying those wants. Such reasoning requires at least some unspecified degree of competence at recognising intuitive implication and inconsistency. In light of this, consider the position of an individual experiencing severe mental impairment.

A person experiencing lack of cognitive focus may attend to peripheral rather than central features of an event or situation, and thereby fail to consider implications, inconsistencies, or the significance of an act for his or her wants.
Impaired reasoning may result in an unwarranted interpretation of situations leading to an irrelevancy of action. Alternatively, it may lead to combinative thinking giving rise to not merely contradictory thoughts, but cognitive processes allowing inconsistency to guide the individual’s actions. The thought processes that allow the individual to hold inconsistent beliefs, prevents him or her from responding appropriately, to the immediate implications and inconsistencies in thinking about their conduct.\footnote{132}{Compare Radden, above n56 at 66.}

Impairment of concept formation involves two types of distorted thought. First, abstract concepts may have no meaning, or the individual may not be able to accurately apply those concepts to concrete events. Secondly, idiosyncratic interpretations and symbolism can lead on one to vest events with an uncalled-for significance.

A failure of ‘reality testing’ involves both a failure to accurately perceive the external world and one’s relationship to it. This inability is not confined to hallucinations. It denotes a cognitive failure, which includes thought disorder leading to faulty reasoning and judgment, as well as purely perceptual errors (hallucinations). A person with schizophrenia is just as likely to draw erroneous conclusions from accurately perceived objective evidence, as he or she is to mistake the flux of his or her own consciousness, for such evidence. Further, he or she may rely on private sensations, with no objective correlative in the real world, as a guide to action.\footnote{133}{McAuley, above n97 at 35–36.}

This failure to distinguish between inter-subjective and intra-subjective experience and meaning is itself a ‘fundamental want of reason’.\footnote{134}{Radden, above n56 at 66.} It amounts to more than a mere failure to reason well. It precludes reasoning itself.

The objects of the person with schizophrenia’s thought have such a radical privacy that they are not connected in an apparent and predictable way with public, external events or stimuli. Nor does the person with schizophrenia share with others the responses required for a shared language. It is the absence of these qualities that deprive the individual with schizophrenia’s speech, while psychotic, of language status. For it is language based upon shared meaning which grounds communication.

Hence, thought within such speech cannot be described as ‘rational’ or ‘reasonable’, for this requires communicability. Such a shared communicable thought presupposes the ability to test reality, and is the basis upon which logic proceeds.

### C. Critique of ‘Rationality’

In everyday life we do not reason in order to be logical, but are logical in order to achieve our goals.\footnote{135}{Evans, Over & Manktelow, above n114 at 170.} This would seem to suggest that practical reasoning somehow accords with our rational designs. Yet practical reasoning has been
subjected to analysis showing that ‘normative theory and psychological facts pass each other by’.136 The correlation between practical reasoning and rationality will be considered in this section.

Effective practical reasoning requires at least three types of cognitive competency.137 First, the actor must be able to form accurate beliefs about the relevant wants and circumstances. Second, the actor must, through an associative process, select an act likely to promote his or her target want from amongst a set of other relevant wants and beliefs. Finally, the actor requires an accurate reasoning process that will allow him or her to draw warranted conclusions about the probable relationships among the various wants, acts and consequences.

The notion of rationality has no ideal definition.138 However, any working definition should be multi-level and incorporate the following:139

- Intelligible desires and beliefs, that is, not wildly inconsistent or incoherent;
- Beliefs and desires which are consistent with one another;
- Beliefs and desires that reflect ‘transitivity’, that is, those that are intelligible and free from direct contradiction, logically cohere, and are implied by one another.

On this basis, to be ‘insane’ is to act on reasons that are not only unintelligible, but inconsistent and incoherent as well.

No one conception of rationality may be said to prevail over any other in support of the above.140 Even so, most advocates of practical reasoning view rationality as objectively ascertainable. For example, Moore writes that those who fail to act for rational ends ‘lack an essential attribute of being a person’.141

Empirical evidence however, suggests that in fact people, due to cognitive limitations, are more likely to achieve goals by satisficing.142 That is, individuals find adequate solutions rather than optimise or maximise utility across all possible choices and outcomes. The search for norms of verification is illusory, as all that may be attained is an account of how people in fact behave.

The idea of rationality operates as an unqualified condition.143 A person is either fully rational or only minimally so. The acquisition of desires and beliefs are judged independently of their role in practical reasoning when ascertaining the threshold question of minimal rationality.144 This however is to preclude the

136 Johnson-Laird & Shafir, above n113 at 3.
137 Schopp, above n117 at 188.
138 Morse, above n66 at 782.
141 Moore, above n66 at 178.
142 Evans, Over & Manktelow, above n114 at 169, 185.
144 Radden, above n56 at 73.
possibility of discriminating between psychotic and non-psychotic behaviour, which any criterion should be able to perform. 145

Focusing on cognitive processes, rather than content, we may say that even sane people entertain unreasonable or irrational beliefs. Faulty judgment, an everyday weakness, finds expression in an insane individual as delusional thinking. Yet delusional thinking need not be an indicator, on its own, of irrationality. For a person may suffer from an apparently delusional belief and maintain this, despite evidence to the contrary, yet not be considered to hold anything other than a false belief, not an irrational one. The belief may be maintained as well grounded where it is based upon idiosyncratic but countervailing ‘evidence’ provided by hallucinations, though these need not go together. 146 The presence of delusions cannot be appealed to in any general way to demonstrate irrationality in insanity. 147

Practical reasoners deny the holding of metaphysical and religious beliefs as fit subjects for consideration with regards to questions of rationality. 148 The orthodox belief in human immortality differs from, for example, an individual with schizophrenia’s messianic belief, only in respect of its widespread acceptance. The difference between the two lies not in the content of the beliefs, but in the way the beliefs are formed and held.

A reliance on expressions such as ‘widespread belief’, ‘widely shared’, and ‘reasonable view’ as criteria precluding analysis lacks intellectual rigour. 149 This is to propose that rationality be confined to an arbitrarily defined area of ideas, and that only uncontested ideas and beliefs count. 150 It is also to ignore the philosophic or religious roots of beliefs that people have about themselves and their environment. In this manner rationality is seen to be the prerogative of a majority.

The concept of rationality conceals a normative and moral judgment. It is the contents of the observed person’s motives and beliefs, which are comparatively assessed. Rationality is both socially and historically contingent. 151 This applies equally to our formal modes of rationality, as well as to our conceptions of rationality. That is, we effectively determine what kind of thought and action will be rational, and in this sense constitute rationality as a mode of human activity.

Norrie has noted the way in which the reasoning legal subject, within a world of legal rights, protected by the rule of law, provides a mechanism of ‘ideological legitimation’. 152 By positing a test of insanity on the basis of rationality, ‘[l]egal discourse … decontextualises[s]…social acts and individualises[s] social

146 Radden, above n56 at 68–69.
147 DSM–IV–TR, above n59 at 821.
148 Radden, above n56 at 73; Morse, above n66 at 784; Moore, above n66 at 105.
149 Ibid.
150 Pols, above n140 at 76.
152 Alan Norrie, Crime, Reason and History (2nd ed, 2001) at 176.
conduct. This is to conceal the link between social context and behaviour and portray conduct as the result of exclusively individual, rather than social, pathology.

D. Reform

Irrationality, as the characteristic of mental impairment, accords with our intuitive notion of what it means to be insane. There must be a direct relationship between the rationality of one’s reasons and the ability to perceive and understand the nature or wrongfulness of what one is doing. It is these latter capacities, which ensure that our beliefs and desires are indeed rational.

However, a belief in an objectively ascertainable view of rationality is illusory. Empirical studies reveal that the way people do reach decisions has less to do with practical reasoning as with pragmatic decision-making. Further, the concept of rationality is seen to be socially and historically contingent, it being the content of a person’s motives and beliefs, which are scrutinised.

The argument that to be insane is to act on reasons, which are not only unintelligible, but also inconsistent and incoherent, denies clinical reality. For many psychotics have hallucinatory belief systems which are internally coherent.

If the reasons for behaviour are unintelligible, this should be sufficient to ground an insanity defence. Yet the concept should be limited to desires and beliefs inconsistent with those held by others in the individual’s societal group. In this fashion, incorrect but intelligible desires, such as jealousy-motivated acts, or beliefs such that stealing while you are poor is justified, will not be deemed to be irrational. Moreover, such limitation will make clear that it is the person’s mental state, and not his or her behaviour, which is the object of the enquiry.

It is suggested that a reformulation of the insanity defence should incorporate the following definitional element:

‘Cognitive competency’, as indicative of practical reasoning, is measured by the extent to which the desires or beliefs, which motivated the offence, are unintelligible in light of what the individual’s society commonly desires and believes.

It will be noted that ‘cognitive competency’ is used in place of either ‘irrationality’, or Foucault’s ‘unreason’ (‘deraison’). I avoid ‘irrationality’ due to the breadth of pejorative connotations it carries, as well as the prevailing belief in its normative value. Foucault’s ‘unreason’ has been contrasted with ‘madness’ (‘folie’). The former represents an understanding of insanity as a form of illogicality or unreasonableness allied to the ordinary, imperfectly rational thought of sane people. Madness, on the other hand, is seen as a medical condition, and

153 Id at 189.
154 Moore, above n66 at 100–108.
156 Radden, above n56 at 2.
as such, alien and beyond comprehension. Yet insanity is not a status, as for example childhood is,\textsuperscript{157} but rather an excuse. It is not allied to ordinary rational thought, but is in fact the total absence of rational thought. While an insane person may have periods of remission, a child cannot be said to have intermittent periods of being ‘doli capax’. A particularly infelicitous expression, ‘unreason’, is best avoided.

5. Criminal Responsibility

I now turn to an examination of the concept of responsibility and its implications for the mental impairment defence. I begin with a consideration of two conflicting accounts of what grounds ascriptions of responsibility, choice and character. An examination of the nature of mental impairment, however, reveals that the conflict between the two theories is spurious. For what underlies criminality is neither character nor choice per se, but wrongful action, which manifests an incorrect attitude to the law. Finally, the interrelationship between cognition and volition will be analysed, and certain conclusions drawn.

A. Choice, Character and Culpability

Modern legal theory, while acknowledging the importance of physical actions, seeks to understand ascriptions of culpability in terms of what lies behind or produces the actions. In particular theorists view criminal liability as being founded on either ‘choice’ or ‘character’.\textsuperscript{158} That is, what justifies conviction and punishment is viewed as being either the defendant’s wrongful choice, or some defect of character that the criminal conduct reveals. Both these notions of responsibility are instantiated in respective theories of excuse.

Representative of the ‘choice’ or ‘capacity’ theory of excuse is the work of H LA Hart.\textsuperscript{159} Hart holds that actors may be punished only where they had ‘the normal capacities, physical and mental, for abstaining from what [the law] forbids, and a fair opportunity to exercise these capacities’.\textsuperscript{160} The theory is but an instance of the more general principle of moral responsibility ‘that a person is not to be blamed for what he [sic] has done if he could not help doing it’.\textsuperscript{161} The justification for the choice theory offered by Hart is grounded in considerations of fairness and justice, constraints ‘which civilized moral thought places on the pursuit of the utilitarian goal by the demand that punishment should not be applied to the innocent’.\textsuperscript{162}

An actor is responsible where that actor’s choice is not made impossible by factors over which he or she has no control, that is, where the actor’s choice is a

\textsuperscript{157} Ronnie D Mackay, \textit{Mental Condition Defences in the Criminal Law} (1995) at 81–85.
\textsuperscript{159} Herbert LA Hart, \textit{Punishment and Responsibility} (1968).
\textsuperscript{160} Id at 152.
\textsuperscript{161} Id at 174.
\textsuperscript{162} Id at 80.
cause of his or her behaviour. Such ‘impossible’ factors are incapacity in the agent or the lack of fair opportunity to use a non-defective capacity.

As to why criminal liability depends on choice, Hart holds that the law is ‘a choosing system’ that aims ‘to guide individual’s choices as to behaviour’.\(^{163}\) This the law does by presenting him or her with reasons for exercising choice in the direction of obedience, but leaving them to choose. Such a system maximises ‘the efficacy of the individual’s informed and considered choice in determining the future’ by making his or her choice ‘one of the operative factors determining whether or not’ he or she is punished.\(^{164}\) Further, choice is viewed as the central condition of responsibility. Cardinal to Hart’s case for legal excuses is the belief that where excuses are absent, offenders act in accordance with their own values, attitudes and beliefs.\(^{165}\)

Character theorists argue that choice theorists’ contentions are defensible only on the basis of two essential premises: that we can separate the choosing-self from what motivates its choices, and that we can take a behavioural view of a person’s character traits.\(^{166}\) Such views, it is contended, do not accord with either our common sense, nor with current tenets of moral psychology.

First, the character theorist argues that the proper focus of the criminal law is not on the action of the defendant. Conviction is warranted given the validity of the inference we may make to an undesirable character trait. It is that character trait which the law condemns and punishes. Where there is an acquittal, given the absence of a fault element or some defence or excuse, the inference from act to character trait is not available. Choice is neither a necessary nor a sufficient condition of criminal liability.\(^{167}\)

Secondly, it is agreed that some of our character traits can constitute motivations to act, because of a particular set of values, feelings, desires, or aversions. Yet, a moral character trait may also refer to far more than a simple disposition to do the act required by some moral rule. It may also encompass dispositions to feel the appropriate emotions and moral motivations to do what the governing moral norm requires. The moral agent cannot simply make a rational choice, to experience the appropriate moral sentiments when he or she wishes to act morally. These moral emotions are the product of prior moral education, experience, and deliberation by the agent about what he or she should strive to be.\(^{168}\)

Objections to the character theory may be in large part due to the way it has been formulated by advocates.\(^{169}\) First, there is a supposition that the point of the

---

163 Id at 44.
164 Id at 46–47.
167 Duff, above n158 at 364.
168 Arenella, above n166 at 79.
criminal law and punishment is not directly to deter and punish harm-causing actions, but to deter and punish bad character, in and of itself. Yet this is to elevate character theory from a theory about moral culpability into a theory of criminal liability. The latter, however, must be underpinned by some version of the ‘harm principle’: liability is only predicated on harm caused. In the absence of any harm attributable to character per se, there can be no argument grounding liability on the basis of the formation or display of bad character.

Secondly, the assumption that culpability is attributable to a character trait renders the theory unable to explain liability for negligently caused harm, for example, the uncharacteristic ‘slip’. Even where the harm was intentionally inflicted, the theory would hold that this would go to the question of culpability, rather than, as is the case, towards the issue of sentencing.

Finally, the theory asks the criminal law to inquire into the accused’s character. That is, the character theory places a gloss on the law, asking the law to go beyond the fault element in ascriptions of culpability, and consider possible motives where no defence is offered.

Such criticism presents the character theorist with an opposite problem to that faced by the choice theorist. The latter’s account of criminal liability is seen to be too ‘thin’: there is more to being criminally responsible than the making of choices. On the other hand, the character theorist, in finding his or her account on the moral significance of character, provides a far too rich account of criminal responsibility.

**B. Mental Impairment and Responsibility**

Not all who raise a defence of mental impairment fail to make a choice. In fact it may be the case that the accused may have acted differently had he or she so chosen. In response to this, a choice theorist may attempt to excuse mental impairment by arguing in one of two ways. First, that such responses are not really choices. The mentally impaired individual is deemed to be unable to engage in practical reasoning, or to conform his or her conduct to any judgments he or she might make.

Alternatively, the choice theorist may hold that mental disorder should excuse only where the agent could not have chosen to act otherwise than they did.

Yet, capacity to make the necessary choice, as a criterion of excuse, fails to take into account the fact that such capacity is a presupposition of moral responsibility and criminal liability, not the object of either. Further, to argue that liability rests on instrumental rationality, or practical reasoning, is not adequate. While some mental disorder manifests a lack or impairment of cognitive capacity, this is not the case in all instances.

---

171 Horder, above n169 at 208.
173 Duff, above n158 at 355.
174 Mackay, above n157 at 82.
Controversially, a person with an Anti-Social Personality Disorder (ASPD) may not appreciate the normative aspects of his or her actions. A person suffering severe depression may fail to have insight into the way in which his or her normative understanding is distorted by the pathological affective state. And a person suffering from a compulsive volitional disorder may have his or her ability to conform to conduct seen to be right effectively precluded. Any attempt to explain these kinds of incapacity must involve an appreciation of what is entailed in the concept of a person.

Self-conception relies on our ability to integrate the personal, impersonal, and non-personal points of view. That is, an integrated individual is able to look at themselves as an ‘I’, a ‘someone’, and a ‘something’, and express these perspectives in thought and action. This integration involves the ability to shift back and forth from one perspective to another, comparing and evaluating one’s beliefs about other people and one’s environment as well as criticising and correcting one’s beliefs and intentions to act in certain ways. This capacity to adopt and integrate these different perspectives is not merely a cognitive or intellectual capacity. Paralleling this development of logic is the development of cooperation and personal autonomy in the moral sphere.

As we have seen, it is a structuring principle of a person’s motivational system that beliefs and intentions are inter-subjectively supportable. Where such a system is closed, the idiosyncratic procedures forming the individual’s belief system are said to lack epistemic justification. Not subject to external or internal checks, the individual is unable to evaluate his or her beliefs, intentions or desires. Such an individual has lost his or her capacity to conform their behaviour to the law. The reasons for this incapacity stem either from the holding of false beliefs, incorrect rules of inference, or the maintaining of inconsistent beliefs. The agent’s conduct is therefore ascribed to the disorder, and the offender excused as not autonomous, as not having entered into the process of forming intentions and beliefs. A character theorist would contend that the person’s action does not manifest any character at all.

It could be argued however, that the criminal law is concerned not with character per se, but with criminal action which is sufficient for liability. Even acts ‘out of character’ result in liability. Criminal action is not merely evidence from which guilt, as some undesirable disposition, is to be inferred, but constitutes

---

179 Radden, above n56 at 141–146.
criminal guilt.\textsuperscript{181} This applies equally to the choice theorist’s claim that an action’s
criminal character is determined purely by what the agent ‘chose’ to do.

This observation gives force to the argument that, in fact, the conflict between
choice and character theorists rests, ultimately, on a spurious distinction between
what a person ‘is’, and what a person ‘does’. As Silber has argued, ‘what a man
does is a function of what he, in the context of his situation, is, and what he is
within this context is revealed by what he does’.\textsuperscript{182} What makes a person
criminally liable is not choice or character, but wrongful action, which, as the
action of a responsible moral agent, manifests in and by itself some inappropriate
attitude towards the law and the values the law protects.\textsuperscript{183}

C. Non-Cognitive Impairment

It was noted above that a person with a compulsive disorder is precluded from
controlling his or her behaviour. Such behaviour is referred to as ‘akrasia’ and is
characterised as ‘acting against one’s better judgement’.\textsuperscript{184} The individual’s
beliefs and desires all support one cause of action, and yet, the person persistently
does another. In such cases, the agent cannot refrain from doing, or repeating, a
particular action. Such cases of non-cognitive impairment\textsuperscript{185} are to be
distinguished from those in which the agent’s desires coincide with his or her
behaviour.\textsuperscript{186} In the latter instance, Hall argues, there is no real way of determining
whether the actor ‘could not’, or ‘did not’, refrain from action.\textsuperscript{187} In such cases the
attribution of responsibility is considered necessary as a socio-political
premise.\textsuperscript{188}

Much of the problem relating to controversial non-cognitive impairments
stems from confusion as to the relationship between volition and cognition.
Traditional volitional standards have been misguided because they have sought a
volitional basis for exculpation that stands separately from the cognitive one. Yet,
the relationship is not that of two distinct processes.\textsuperscript{189} Volition is the exercise of
the faculty, or function, by which one engages in conscious and intentional action,
as a result of decision or choice, through deliberation. Volitional impairment
involves some disruption of the practical reasoning process through which an actor
selects a goal. Severe cognitive disturbance distorts this process of practical
reasoning. Major cognitive psychopathology therefore, disrupts the volitional
process. It does so by distorting the causal process through which the actor’s

\textsuperscript{181} Duff, above n158 at 379.
\textsuperscript{182} John R Silber, ‘Being and Doing: A study of Status Responsibility and Voluntary
Responsibility’ (1967) 35 UChiLR 47 at 65 (emphasis in original).
\textsuperscript{183} Duff, above n158 at 380.
\textsuperscript{184} DSM–IV–TR, above n59 at 667.
\textsuperscript{185} Radden, above n56 at 133.
\textsuperscript{186} McAuley, above n97 at 50.
\textsuperscript{187} Jerome Hall, ‘Mental Disease and Crime’ (1945) 45 Columbia LR 677 at 702–703.
\textsuperscript{188} Pols, above n140 at 79.
\textsuperscript{189} Schopp, above n117 at 202.
reasons for acting serve as structuring causes for his or her decision to perform the act he or she selects.\textsuperscript{190}

It follows that a cognitive interpretation of a mental impairment defence that addresses distortion of cognitive process, rather than content, provides an accurate account of both cognitive and volitional aspects of major psychopathology. A similar observation was made by Stephen, who defined self-control as ‘a power to attend to distant motives and general principles of conduct, and to connect them rationally with the particular act under consideration’.\textsuperscript{191} This link, between cognition and volition, allows Stephen to hold ‘that a man who by reason of mental disease is prevented from controlling his own conduct is not responsible for what he does’.\textsuperscript{192}

Stephen’s formulation makes clear that, on this interpretation, the crucial question becomes whether a mental disorder prevented the accused from controlling his or her actions.\textsuperscript{193} The fact that an accused person could control his or her behaviour, in the sense of being able to engage in practical reasoning, however, does not vitiate a defence of volitional insanity. Such would be the case where the derangement of an accused’s mind precludes his or her accessing the kind of moral information required for the development of a sense of self, as discussed above. In such an instance we would say that, while the cognitive component is functioning, the individual’s motivational-affective state is not.\textsuperscript{194}

D. Reform

It has been argued that while some mental disorder may be described in terms of a defect in instrumental rationality, this does not account for all cases. In particular, non-cognitive disorders, such as the compulsive’s inability to control his or her behaviour, where this does not accord with his or her desire, pose conceptual difficulties.

The difficulty can be seen to stem from a failure to note the interrelatedness between cognition and volition. Severe cognitive impairment impacts on volitional behaviour by distorting the causal process through which an actor selects a goal.

Cognitive ability to engage in practical reasoning is a necessary, not a sufficient, condition of volitional behaviour. For a sense of personal agency is predicated not only on an intellectual, but also on an inter-subjectively supportable, system of beliefs and intentions. This latter has been shown to stem from an integration of differing viewpoints.

In the light of the complexity inherent in human action, any conception of responsibility should take into account more than simply the actor’s cognitive competence. It should equally seek to comprehend the manner in which this cognitive element interrelates with the actor’s volition or ‘motivational-

\textsuperscript{190} Id at 203, 246.
\textsuperscript{191} Stephen, above n177 at 170.
\textsuperscript{192} Id at 167.
\textsuperscript{193} McAuley & McCutcheon, above n10 at 663, 672.
\textsuperscript{194} JM Livermore & PE Meehl, ‘The Virtues of M’Naghten’ (1967) 51 Minn LR 789 at 807.
affective,' state. A person may distinguish between right and wrong and yet not be able to choose between them.

Such a conception would accord with the test stated by Dixon J in _R v Porter_: criminal responsibility will attach according to whether ‘D could not think rationally of the reasons which to ordinary people make his act morally right or wrong’. Yet in stating such a test, two refinements should be made. The expression ‘rational’ should be replaced by ‘cognitive competence’, and the enquiry should turn on whether the conduct amounted to an offence, not its morality. This is a purely legal question.

A reformulation of the insanity defence should include the following element in relation to the question of criminal responsibility:

A person is responsible for his or her actions where he or she had adequate cognitive competency to think of the reasons which people are expected to regard as sufficient grounds for refraining from commission of the offence.

Such reasons are: the prohibited nature of the conduct; the risk of punishment and the absence of offsetting advantage; and the ability to exercise choice.

Such a formulation, rooted in a principle of understanding, effects two immediate advantages. First, the ‘knowledge of wrongness’ rule is displaced by a test of competency for understanding the imperatives of compliance. So where individuals are deluded into thinking that they are satisfying a divine command, or higher objective, even though they know that the conduct is legally or morally forbidden, they will have a defence. For, such individuals may be said to lack the capacity to think competently about whether there is an overriding advantage in so acting. They may also be unable to think of their selves as capable of exercising choice.

Secondly, the issue is one of competency, not mental impairment. Understanding is thus not vitiated where the accused’s state was one of dissociation from extreme stress, or substance induced hallucinations. In such instances policy considerations become explicit in the law’s concern with public security. This concern would find expression in evaluations of ‘dangerousness’ and the need for preventive detention. This would be best handled under mental health legislation providing involuntary commitment and mandatory supervision.

195 Id at 808.
196 McAuley & McCutcheon, above n10 at 649.
197 (1933) 55 CLR 182 at 189 approved _Stapleton v R_ (1952) 86 CLR 358.
199 Bernadette McSherry, ‘Criminal Detention of those with Mental Impairment’ (1999) 6 JLM 216 at 220
6. Conclusion

It was noted in Part 2 (C) that insanity is a social judgment founded upon, but not precisely representing, a medical diagnosis. In establishing the defence it has been held that what constitutes a ‘disease of the mind’ is a question of law for the judge. Whether the accused was insane is a question for the jury, aided by expert and other evidence. Usually, the expert evidence is that given by a psychiatrist. It is my contention that what results is a conceptual confusion stemming from law and psychiatry’s differing methodologies. It has been the analysis of this confusion, that has formed the subject of this paper. The submissions for a proposed reformulation have sought to make explicit what is often only implicit in the Rules.

Criminal responsibility has been reformulated in terms of a ‘principle of understanding’. A legal question, it operates as a threshold to the defence of insanity. It makes no reference to mental impairment but re-conceptualises the issue as one of competency to meet an expected standard of behaviour. Where such is not met the accused is deemed not responsible. Yet, in order to preclude the release of potentially dangerous individuals, the law is to engage the notion of ‘dangerousness’. In this fashion social defence concerns will be met.

In relation to the jury question, the law seeks the aid of experts. It is argued that it would be conducive to the transparency, though not necessarily surface clarity, of the law that psychiatry’s evaluative implications be made explicit. In this way psychiatric nosology may be shown to be both societally and historically contingent.

Morris has remarked that the M’Naghten Rules while ‘woolly, semantically confused, psychologically immature nonsense’ are ‘means whereby juries work rough justice in a difficult peripheral area of law and morality’. What has been attempted above is a restatement of the Rules in accordance with a clearer understanding of the foci of law and psychiatry. Accordingly it is proposed that a restatement of the insanity defence be legislatively enacted in these terms:

A person is not criminally responsible if at the time of the commission of the offence he or she had a mental impairment which included in its symptoms or consequences a loss of cognitive competency to think of the reasons which people are expected to regard as sufficient grounds for refraining from commission of the offence.

---

Definitions

‘Mental Impairment’ is to be understood as a defect in practical reasoning, characterised by a distortion of natural functions and a diminution or loss of normal functions.

‘Cognitive competency’ as indicative of practical reasoning, is to be measured by the extent to which the desires or beliefs, which motivated the offence, are unintelligible in light of what the individual’s society commonly desires and believes.

‘Reasons’ include, the prohibited nature of the conduct; the risk of punishment and the absence of offsetting advantage; and the ability to exercise choice.

The history of the insanity defence could be recapitulated as the struggle between two different imperatives: social defence, the law’s concern, and scientific objectivity, purportedly measured by a notion of value-free truth existing in the domain of psychiatry. It is believed that the underlying methodological differences between law and psychiatry will be rendered, by being made explicit, more suited to the fulfilment of a single agenda. This agenda must include as a goal the explication of both principle and policy issues as they appear in forensic settings.
Blue Asbestos and Golden Eggs: Evaluating Bankruptcy and Class Actions as Just Responses to Mass Tort Liability

PETA SPENDER*

Abstract

Mass torts give rise to complex legal questions and invidious moral choices. The asbestos litigation has shown that corporations manufactured asbestos decades after its dangers had been publicly recognised. Later, when faced with spiralling claims, firms in the US such as Johns-Manville were permitted to use bankruptcy procedures without proving insolvency thereby forcing tort claimants into a limited fund. In the late 1990s asbestos defendants sought wider powers to collectivise the claims through class actions although this attempt was ultimately unsuccessful. This article provides case studies of US firms and shows that similar strategies are now being adopted in Australia and the UK.

Certain privileges flow from bankruptcy such as the moratorium on claims and the right to distribute entitlements pro rata. However, in the context of mass torts these privileges have frequently led to under-compensation of tort victims, wealth transfers to shareholders and bewilderment about how to protect future claims. The article will explore these problems and consider how they may be ameliorated by effective monitoring.

1. Introduction

‘James Hardie Industries yesterday took its biggest step yet in shedding the baggage of the past by quarantining almost $300 million in current and future asbestos liabilities to allow it to focus on its international ... operations. [Due to] the plan to create a “foundation” to take its asbestos liabilities off its balance sheet ... James Hardie will no longer bear any costs associated with asbestos. While there remains some doubt about James Hardie’s ability to stave off any future liabilities, $72.4 million in asbestos provisions will be removed from the company’s balance sheet in favour of the foundation and no future provisions are expected. The move was welcomed by investors, who pushed the shares 17c higher to $3.80.1’

‘Mr Gerry Gardiman of law firm Turner Freeman ... is suspicious of the offer. “My concern would be what happens to the victims when the $293 million runs out?” he said.2

* Reader in Law, Faculty of Law, ANU. The author wishes to thank Kent Anderson, Juliet Behrens, Stephen Bottomley, the two anonymous referees and the participants in the CLTA Conference at Monash University on 12 February 2002 for their helpful and insightful comments on this article and the issues raised by it.


The purpose of this article is to examine corporate responses to mass tort liability. The article will explore techniques used by defendants to aggregate tort creditors through bankruptcy and class actions to reduce compensation.

Mass tort liability is generally very large, involves high transaction costs and in the case of long-tail torts may extend over decades and may threaten the viability of the defendant enterprise. It creates incentives for corporations to minimise payouts to tort creditors. Although there are many examples of mass torts due to defective products, asbestos liability is an archetypal example and it will form the subject matter of this article.

This article sets out to do a number of things. First, it will provide a narrative of the asbestos litigation featuring defendants in the US, Australia and the UK. Judgments for personal injury and death caused by exposure to asbestos have been occurring in Australia and the US since the late 1970s. Since then the volume of asbestos claims continues to spiral, surpassing most predictions.

---

3 ‘Mass torts’ may be defined as ‘encompassing any (negligently or strictly) tortious systematic risk-taking by business that exposes some population of individuals to injury in person or property or both. All product and occupational liability, as well as all business-generated environmental hazards, are mass torts.’ David Rosenberg, ‘Mass Tort Class Actions: What Defendants Have and Plaintiffs Don’t’ (2000) 37 Harv on Legis 393 at 393.

4 A problem of terminology is already arising because we need to think about two different concepts — bankruptcy and insolvency. In this article I will use the word ‘bankruptcy’ in a general sense to describe the capacity to invoke the legal machinery of insolvency such as reorganisation procedures and in a particular sense to refer to US procedures. The distinction between bankruptcy and insolvency is important, but the terminology is slippery. ‘Insolvency’ is defined under s95A of the Corporations Act 2001 (Cth) as the inability to pay debts as and when they fall due. In Australia and the United Kingdom, the term ‘bankruptcy’ generally refers to the end point of a process undertaken under the Bankruptcy Act and specifically refers to individuals. In the USA, the term refers to both corporate and individual bankruptcies under Chapter 7 of the Bankruptcy Code and reorganisations under Chapter 11 of the same code. By comparison, in Australia ‘corporate insolvency’ is dealt with under Chapter 5 of the Corporations Act 2001 (Cth).

5 In a long-tail tort there is a long latency period between a person’s use or exposure to a harmful product and the first manifestation of harm.

6 Such as Agent Orange, Dalkon Shield and silicone gel breast implants.

7 Asbestos will have the greatest number of claims and liability will come in waves for many years to come. For example, in CSR v Young (1998) Aust Torts Reports ¶81–468 at 64952, the New South Wales Court of Appeal found CSR liable in the first mesothelioma case involving the townspeople of Wittenoom. The plaintiff was not an employee of CSR or its subsidiaries.

8 The judgment in Borel v Fibreboard Paper Prods Corp 493 F 2d 1076 (1973) by the US Fifth Circuit that asbestos manufacturers could be held strictly liable for injuries resulting from asbestos exposure is regarded as the point at which asbestos litigation began in the USA.

9 In 2001 Behrens and Parham quote statements by two asbestos defendants (GAF and Owens Corning) who had settled more than 740 000 claims. See Monica Parham & Mark Behrens, ‘Stewardship For The Sick: Preserving Assets For Asbestos Victims Through Inactive Docket Programs’ (2001) 33 Texas Tech Law Review 1 at 1–2. A study by the RAND Institute in August 2001 showed that although more than 500 000 claims have been filed, this may be less than half of those that ultimately come forward: Deborah Hensler, Stephen Carroll, Michelle White & Jennifer Gross, Asbestos Litigation in the US: A New Look at an Old Issue (RAND Institute, 2001).
The second purpose is technical — the article will detail the positive law relied upon by defendants to aggregate and under-compensate asbestos victims, primarily via corporate law and civil procedure. A number of US firms have used reorganisation under Chapter 11 of the US Bankruptcy Code. The use of reorganisation in this context challenges traditional bankruptcy norms because the firms involved are not necessarily insolvent, yet they are able to allocate a pool of funds from which the tort creditors' claims must be satisfied. Part Two of the article will examine the history of one of the largest US manufacturers of asbestos, the Johns-Manville Corporation, the evidence of its complicity in the concealment of the dangers of asbestos and the reaction to its Chapter 11 filing.

Mandatory class actions will be considered in Part Three of the article. Although many US firms managed their mass tort liability with the use of Chapter 11, other defendants attempted to use mandatory settlement class actions under the US Federal Rules of Civil Procedure ('FRCP') to achieve a wider aggregation and discharge. The discharge is effected by operation of res judicata which arises upon certification of the class action. A similar role may be played by allocating a 'limited fund' which must be apportioned amongst the members of the class.

Part Four will examine the application of this issue to Australia where one asbestos defendant has set up a limited fund ‘in order to absolve itself for all liability with respect to future asbestos claims’ and another has been party to several settlement class actions. The final case study in Part Five will investigate the circumstances leading up to the administration order made in October 2001 over T&N Ltd, which was the largest British manufacturer of asbestos. The administration order flows from the Chapter 11 filing of T&N’s US parent, Federal-Mogul, who acquired T&N in 1997.

The third purpose of the article is normative. Part Six will attempt to tease out the dilemmas created by these corporate responses to mass tort liability. This analysis will weigh up the right of corporations to defray enterprise-threatening liability through aggregating devices with the obligation of the tort system to achieve corrective justice and individual compensation. The article will conclude with a critical examination of the ‘golden egg’ arguments raised by corporations in this context.

10 For example, Federal Court Act 1976 (Cth) s33ZA; Federal Rules of Civil Procedure (US), Rule 23(b)(1)(B).


12 Aesop tells the story of the Goose with the Golden Eggs. In that story a farmer was astonished to discover that his goose had laid an egg of solid gold. He took the golden egg to the market and sold it for a good price. Every day thereafter the farmer found an egg of the purest gold and he sold them one by one. He became a rich man. One day he thought, ‘Why must I be content with only one golden egg a day? If I kill the goose and cut her open, I can take all the treasure at once!’ Whereupon he killed the goose but there was no gold at all inside. ‘Foolish man!’ cried his wife, ‘if only you had understood that those who are greedy for too much sometimes lose all.’ Ann McGovern, Aesop’s Fables (9th ed, 1974) at 72.
2. The Johns-Manville Corporation

From the 1920s until the 1970s Johns-Manville was both the largest manufacturer and the largest supplier of asbestos products in the United States.13 As early as 1933, litigation by 11 employees alleging injuries caused by exposure to asbestos placed the company on formal notice of the health risks for workers.14 According to the minutes of a board of directors meeting on 24 April 1933, the suits were settled by Johns-Manville on the express condition that the plaintiffs’ attorneys would not bring similar claims against the company in the future.15

Correspondence produced on discovery in a South Carolina case included letters about the suppression of publication of articles in Asbestos magazine at the direction of industry executives. According to Judge James Price:

[T]he correspondence reveals… that Raybestos-Manhattan and Johns-Manville exercised an editorial prerogative over the publication of the first study of the asbestos industry which they sponsored in 1935… [I]t further reflects a conscious effort by the industry in the 1930s to downplay, or arguably suppress, the dissemination of information to employees and the public for the fear of promotion of lawsuits.16

Asbestos workers began forming unions and the presence of unions facilitated the transition from individual workers reporting asbestos-related injuries to the extensive epidemiological data available today. Unions also provided workers with access to lawyers who developed an expertise in the area. Although the early cases were often unsuccessful, as plaintiff lawyers developed further evidence and expertise, there was a sudden explosion of asbestos litigation.17

---

14 Concerns about asbestos have some antiquity. In 1st Century AD, Pliny the Elder noted that slaves working in asbestos mines were dying at a young age of lung disease. By the 20th century the concerns were gaining momentum. For example, in 1926 the first of hundreds of successful claims for compensation for asbestos disease were settled by the Massachusetts Industrial Accidents Board and in 1930 the British Home Office reported widespread asbestos disease in UK factories and introduced regulations controlling dust particle density. For a history, see Paul Brodeur, Outrageous Misconduct: The Asbestos Industry on Trial (1985).
15 Brodeur, id at 113–114.
By 1982 the pace of litigation against Manville had increased to an average filing of three cases per hour, every hour of the business day.\(^\text{18}\) The company projected its total asbestos liability at more than $1 billion.\(^\text{19}\) On 26 August 1982, Manville Corporation filed for reorganisation under Chapter 11 of the US Bankruptcy Code.

Despite this massive contingent liability, Manville was clearly not insolvent. The Wall Street Journal pointed out that Manville was ‘far from broke’ and in fact was financially strong. The company ran a series of full-page advertisements in the Wall Street Journal that, despite the filing, proclaimed its financial strength.\(^\text{20}\)

In the advertisements, Manville declared to the world, ‘[n]othing is wrong with our businesses.’\(^\text{21}\) The company did not need to ‘fudge or falsify’ the allegations of its financial strength because the Bankruptcy Code had been amended in 1978 to remove any requirement for a voluntary petitioner to be insolvent or unable to pay its debts as they fell due. Kennedy stated:

> This seemed a glaring gaffe to a number of commentators on the legality of the filing, but voluntary relief under the bankruptcy laws has long been available without the bootless expenditures of time and money to determine whether the debtor was in sufficiently dire financial straits to require or deserve bankruptcy relief.\(^\text{22}\)

The 1978 amendments had effected what Zipes has called ‘bankruptcy’s paradox’, that is, that the debtor need not be insolvent to gain the finality of discharge.

> A debtor’s ability to repay debts in whole or in part is not adequate cause for dismissal of a case under the Bankruptcy Code. Until a debtor proposes a plan, it is too speculative to conclude that reorganisation is not in the best interests of the debtor and debtor’s creditors simply because the debtor is solvent. Further, if a debtor is able to fully recover after confirmation of a plan, perhaps becoming wildly successful, it has no duty to pay pre-petition creditors.\(^\text{23}\)

---

\(^\text{18}\) Kevin Delaney, Strategic Bankruptcy: How Corporations and Creditors Use Chapter 11 to Their Advantage (1992) at 61. Coffee reports that in 1982 over 16 500 cases were outstanding against the company. Manville had either settled or judgment had been entered in more than 3570 claims at an average disposition cost of $20 000. John Coffee, ‘Class Wars: The Dilemma of the Mass Tort Class Action’ (1995) 95 Columbia Law Review 1343 at 1387.

\(^\text{19}\) Coffee, ibid.

\(^\text{20}\) Frank R Kennedy, ‘Creative Bankruptcy? Use and Abuse of the Bankruptcy Law — Reflection on Some Recent Cases’ (1985) 71 Iowa Law Review 199 at 202. One author has stated that the reorganisation made front page news throughout the US because at the time Manville was the largest company ever to file under this legislation, and, with assets of more than $2 billion, one of the richest companies ever to take such action. At the time it was 181\(^\text{st}\) on Fortune’s list of the US top 500 industrial corporations. London Hazards Centre, Asbestos: The Worst Industrial Killer (1994): [http://www.lhc.org.uk/members/pubs/books/asbestos/ashb09.html#_Toc401990477\(^\text{2}\)](http://www.lhc.org.uk/members/pubs/books/asbestos/ashb09.html#_Toc401990477\(^\text{2}\)) (12 Jan 2003).

\(^\text{21}\) Delaney, above n18 at 62.

\(^\text{22}\) Kennedy, above n20 at 202.

By 1991, 11 of the 25 major asbestos manufacturers had sought bankruptcy protection. These bankruptcies were controversial because their novel premise was that the corporation was entitled to bankruptcy protection not because it was insolvent, but because continuing trends in asbestos litigation made a bankruptcy reorganisation the best way to manage the payment of present and future claims.24

The Manville Plan of Reorganisation provided for the creation of a trust that was designed to satisfy fully Manville’s asbestos-related liability. During the six years of bankruptcy, no asbestos-related claims were paid. Upon the approval of the plan in November 1988, the trust became the exclusive entity from which tort creditors could seek compensation for injuries caused by exposure to Manville’s asbestos products. An injunction insulated the Manville Corporation from any asbestos related liability.25

Shareholders retained 20 per cent of the equity in the company,26 which although low, still offends bankruptcy priorities.27 Commercial creditors were granted full payment on the monies owing to them and a package of shares, warrants and bonds in lieu of interest.28 The uncertainty of the tort claims was replaced by a ‘compensation board system’.29 A bankruptcy judge, Lifland J, commented:

[T]he Trust is guaranteed an ‘evergreen’ source of funding. This funding of the Trust will continue until the last asbestos victim is found and paid. Thus an effort to determine the number and amount of [asbestos] claims with exactitude is not imperative. The imperative rather, is to ensure to the greatest degree possible the continuing viability of the reorganised corporation, which will fund the Trust, whatever the number and amount of claims happen to be.30

---

25 In Re Joint Eastern & Southern, above n13 at 752 (Weinstein J).
26 Manville’s shareholders lost up to 80 per cent of the value of their stock in a ‘reverse split’ in which six shares were traded for a single share. Delaney, above n18 at 72.
27 It is a fundamental premise of corporate law that shareholders’ interests are deferred to creditors. Upon winding up, shareholders should receive nothing until creditors are paid in full. In US bankruptcy law the principle is referred to as the ‘absolute priority rule’. As applied to a reorganisation it means that when a firm owes more than its assets are worth, the shareholders receive nothing unless the creditors consent. See Douglas Baird & Thomas Jackson, ‘Bargaining after the Fall and the Contours of the Absolute Priority Rule’ in Jagdeep Bhandari & Lawrence Weiss (eds), Corporate Bankruptcy: Economic and Legal Perspectives (1996) at 113.
28 Delaney, above n18 at 71.
29 Delaney, id at 79.
30 In re Johns-Manville Corp 68 B R 618 (1986), 621–622. As of 31 December 1995, the trust had received over 280,000 claims. The Plan predicted the trust would receive between 83,000 and 100,000 claims during the life of the trust. Manville Personal Injury Settlement Trust — History: <http://www.mantrust.org/history.htm> (12 January 2003). A report issued by the trust in May 2001 indicated that the number of claims was still increasing. During 2000 new claims totalling 59,200 were filed with the trust which represented an 82 per cent increase in volume over the 32,500 claims filed during 1999, and was the greatest volume of claims received in any year since 1989, the first full year the trust received claims.
Within two years of commencing operation the trust was out of funds.  

It was necessary for the company to restructure again and this time it used a class action device by declaring the fund allocated under the Chapter 11 reorganisation as a ‘limited fund’. This device will be discussed shortly. Ultimately the value of the fund declined so that now claimants are paid 5 per cent of the liquidated value of their claims.

3. The Movement to Defendant Class Actions

A. Class Actions as Defendant Exocet

[T]he class action has landed like a 600-pound gorilla in the arena of tort reform, where there has of late been increasing interest in replacing tort litigation with scheduled benefits like those provided in these class action settlements. One reaction to using class actions to accomplish this reform might be ‘They can’t do that, can they?’

By 1990 in the US, judicial concern about the problems caused by the asbestos litigation and the long-term financial viability of defendants led to the establishment of an ‘ad hoc nationwide coordinating committee’ to resolve the asbestos case load crisis through case consolidation and management. The report of the Ad Hoc Committee on Asbestos Litigation in 1991 concluded that ‘this litigation impasse cannot be broken except by aggregate or class proceedings’. The report also urged a national system for resolving asbestos claims to ‘ensure uniform consistent recovery to claimants’. Later that year eight federal district judges petitioned the Judicial Panel on Multidistrict Litigation (JPML) to consolidate all asbestos personal injury cases in the federal system into a single forum. The JPML transferred all pending personal injury asbestos cases to Judge Charles R Weiner of the Eastern District of Pennsylvania, who resisted plaintiff attempts to transfer matters back. As a result, plaintiff firms became more receptive to the idea of a global settlement.

Then, in what has been described as ‘a surprising development’, the largest industry consortium of defendants (CCR) approached two large plaintiff firms to negotiate a separate settlement with their firms covering ‘all future personal injury claimants in the United States who had been exposed to asbestos’. The next step,

31 In Re Joint Eastern & Southern, above n13 at 754 (Weinstein J).
32 Pursuant to the Court’s decision in In re Joint Eastern and Southern, Findley v Falise 878 F Supp 473 (1995), the trust is required to undertake a re-estimation of the pro rata share at least every three years. On 19 June 2001 the trustees decided to reduce the pro rata payment percentage from 10 to 5 per cent in view of the expanding number of claims being filed. See: <http://www.mantrust.org/FTP/REDUMEM2.pdf> (18 April 2003).
35 Coffee, above n18 at 1389.
36 Coffee, id at 1392.
according to Coffee, was ‘even more extraordinary’. In January 1993, the two sides simultaneously filed the complaint, the answer, a joint motion for class certification and a settlement in Georgine v Amchem Products, Inc in a federal district court. In other words, they had devised a global settlement and were simultaneously filing the complaint and the settlement in order to obtain class certification. Class certification was sought in order to attain the benefit of res judicata which would preclude those covered by the settlement and certification from bringing a claim. Thus the coverage of the settlement was critical; and amazingly the settlement was framed to catch all future claimants. In August 1994 Judge Reed issued an order that certified the class action under Rule 23(b)(3) FRCP and approved the proposed settlement as fair and adequate.

The court enjoined all class members from ‘initiating or maintaining any asbestos-related personal injury or death claim(s) or lawsuit(s) against any CCR defendant’ except as expressly permitted by the Stipulation of Settlement.

Shortly after the Amchem settlement was certified, but before it was scrutinised by the US Supreme Court (which will be discussed later), the case of Ahern v Fibreboard was filed. Ahern was similar to Amchem but it is more pertinent to the arguments in this article.

Fibreboard was a long-term subsidiary of Louisiana Pacific Corporation which had been spun off by its parent as an independent publicly owned company in 1988 at a point when its potential insolvency due to asbestos liabilities was already a clear danger. By that time, over 50000 asbestos-related actions had been filed

---

37 Georgine v Amchem Prods, Inc 157 F.R.D. 246; Coffee, id at 1393.
38 This occurs because class actions under Rule 23 FRCP are based upon an opt-out procedure. An opt-in procedure is typified by the joinder provisions which are available in most courts, eg, ACT SCR O19 r 10, NSW SCR Pt 8.13, FCR O 6 r 2. These provisions are more restrictive than class action procedures because the parties must have the same interest and the words ‘same ... series of transactions’ have involved some fatal constraints, such as where applicants participated in only one of the series of transactions. But the basic quality of these procedures is that the parties who are represented must take positive steps to be part of the proceedings and to be bound by the judgment. Contrast that with an opt-out procedure. In this procedure the represented parties are described rather than named. Therefore the class will be formed by reference to a particular description, eg, ‘people who purchased X Ltd’s peanut butter products from x date to y date and who suffered personal injury’. By virtue of the doctrine of res judicata, where an opt-out procedure is used, everyone who comes within the description is part of the class and will be bound by the judgment unless he or she takes positive steps to not be part of the class, ie, to opt-out. Section 33E of the Federal Court Act 1976 (Cth) states that no consent is required to be a group member with certain exceptions, eg, the Commonwealth, Ministers of the Crown etc. See Carnie v Esanda Finance Corporation Limited (1995) 182 CLR 398.
39 Coffee, above n18 at 1393. The class was defined to include ‘all persons ... who have been exposed in the United States or its territories ..., either occupationally or through occupational exposure of a spouse or household member, to asbestos ... for which one or more of the defendants may bear legal liability and who ... have not, as of January 15, 1993, filed a lawsuit for asbestos-related personal injury or damage ... against the defendant(s)’: Georgine, 157 F.R.D. at 319, discussed in Ortiz v Fibreboard 527 US 815 (1999) at 827.
40 Preliminary Injunction, Joint Appendix 1049 & 1050, in Coffee, above n18 at 1393.
41 1995 US Dist LEXIS 11062 (E D Tex, 27 July 1995). Note that the Ahern v Fibreboard proceedings were named Ortiz v Fibreboard before the US Supreme Court.
Fibreboard began negotiating with plaintiff firms to secure a global settlement. In December 1992, Fibreboard and the large plaintiff firm Ness, Motley reached an inventory settlement covering approximately 20,000 existing asbestos personal injury claims. Pursuant to this agreement, Ness, Motley agreed to recommend the settlement provisions to future claimants that it represented. The class action was filed on 9 September 1993 and on the same day Parker J certified, for settlement purposes only, a future class of persons with asbestos-related personal injuries who had not already asserted such claims against Fibreboard. His Honour appointed Ness, Motley to act as negotiating counsel for the future claimants even though they were also acting for present claimants.

The most interesting part of the certification was Parker J’s decision to certify the class as a mandatory class under Rule 23(b)(1)(B) FRCP. Under the settlement, Fibreboard’s two principal insurers had deposited $1.525 billion into a trust fund for future claimants but Fibreboard contributed only $500,000. Coffee commented that ‘this token contribution might be understandable if Fibreboard were otherwise insolvent, but, apart from its asbestos liabilities, Fibreboard appears to be a thriving company.’ Expert opinion at the hearing valued Fibreboard’s assets at between $230–$300 million.

Certification of an action as a mandatory class has the practical effect of denying class members the right to opt out. Moreover, the finding of a limited fund allows defendants to limit tort creditors to an asset pool, which must necessarily be allocated pro rata. Such a result would generally flow from bankruptcy, but not where the company is solvent. In the result, Fibreboard’s share price increased by over 300 per cent above its highest closing price in the month before the settlement. Coffee expressed the unfairness succinctly: ‘Such a dramatic stock
price movement suggests that there has been a wealth transfer. Fibreboard’s shareholders have gained, and its tort creditors … have lost.\textsuperscript{49}

B. Due Process Strikes Back

[W]e do need to recognize the tension between the limited fund class action’s pro rata distribution in equity and the rights of individual tort victims at law.\textsuperscript{50}

\textit{Amchem} and \textit{Fibreboard} were both appealed to the US Supreme Court. In \textit{Amchem Products, Inc v Windsor}\textsuperscript{51} the court sharply circumscribed the use of mandatory settlements in mass tort litigation and found that the ‘sprawling class’ which had been certified by the District Court in that case did not satisfy the requirements of Rule 23 FRCP.\textsuperscript{52} In \textit{Ortiz v Fibreboard},\textsuperscript{53} the court rejected the company’s limited fund settlement and reversed the grant of class certification for the purpose of a global settlement. The court found that the petitioners had failed to demonstrate that the fund was limited except by the agreement of the parties and that the fund lacked the structural protections of the FRCP.

The majority judgment in \textit{Ortiz}\textsuperscript{54} found that when the US Federal Civil Rules Advisory Committee was drafting Rule 23(b) FRCP in 1966, they sought to catalogue ‘in functional terms those recurrent life patterns which call for mass litigation through representative parties’. Rule 23(b)(1)(B) FRCP covered ‘situations where lawsuits conducted with individual members of the class would have the practical if not technical effect of concluding the interests of the other members as well, or of impairing the ability of the others to protect their own interests’.\textsuperscript{55}

Among the traditional suits encompassed by Rule 23(b)(1)(B) FRCP were those involving property which needed to be distributed among a class or liquidated to provide payments to members of a class. One example was where claims were aggregated against a fund insufficient to pay all claims, that is, a limited fund.\textsuperscript{56} Classic limited fund class actions ‘include claimants to trust assets, a bank account, insurance proceeds [and] company assets in a liquidation sale’.\textsuperscript{57}

The majority in \textit{Ortiz} found that a limited fund has the following characteristics:

\begin{itemize}
\item \textsuperscript{49} Coffee, above n18 at 1402.
\item \textsuperscript{50} Ortiz, above n39 at 845.
\item \textsuperscript{51} 521 US 591 (1997).
\item \textsuperscript{52} In particular the predominance requirements under Rule 23(b)(3) and adequate representation under Rule 23(a)(4).
\item \textsuperscript{53} Ortiz, above n39 at 815.
\item \textsuperscript{54} Souter J delivered the opinion of the court, in which Rehnquist CJ, and O’Connor, Scalia, Kennedy, Thomas, and Ginsburg JJ joined. Rehnquist CJ filed a concurring opinion, in which Scalia & Kennedy JJ joined. Breyer J filed a dissenting opinion, in which Stevens J joined.
\item \textsuperscript{55} Ortiz, above n39 at 833 (Souter J).
\item \textsuperscript{56} Ortiz, above n39 at 834 (Souter J).
\item \textsuperscript{57} Newberg on Class Actions §4.09 at 4–33.
\end{itemize}
1. The totals of the aggregated liquidated claims and the fund available for satisfying them, set definitely at their maximums, demonstrate the inadequacy of the fund to pay all the claims. ‘The concept driving this type of suit was insufficiency, which alone justified the limit on an early feast to avoid a later famine.’

2. The whole of the inadequate fund was to be devoted to the overwhelming claims.

3. Third, the claimants identified by a common theory of recovery were treated equitably among themselves. The cases assume that the class will comprise everyone who might state a claim on a single or repeated set of facts, invoking a common theory of recovery, to be satisfied from the limited fund as the source of payment.

Overall, mandatory class treatment through representative actions on a limited fund theory was justified ‘with reference to a ‘fund’ with a definitely ascertained limit, all of which would be distributed to satisfy all those with liquidated claims based on a common theory of liability, by an equitable, pro rata distribution’.

This looks very much like a description of bankruptcy or a Chapter 11 reorganisation and demonstrates the functional equivalence between the two procedures. This is particularly the case where certification is sought under Rule 23(b)(1)(B) FRCP because there is no opportunity for members of the class to opt out of the proceedings. The opt-out power has been a significant sweetener to allay the due process concerns raised by class actions. Removal of this option brings the class action process very close to bankruptcy proceedings. This was recognised by the majority judgment, which quoted Monaghan as follows: ‘No draftsmen contemplated that, in mass torts, (b)(1)(B) “limited fund” classes would emerge as the functional equivalent to bankruptcy by embracing “funds” created by the litigation itself.’

The Supreme Court rejected this functional equivalence. The first basis was the poor fit between the original concept of the limited fund and that which was being argued by Fibreboard. The majority quoted Amchem, emphasising that the ‘rules of procedure “shall not abridge, enlarge or modify any substantive right”’. 

---

58 Ortiz, above n39 at 838.
59 Ortiz, above n39 at 841.
60 Henry Monaghan, ‘Antisuit Injunctions and Preclusion Against Absent Nonresident Class Members’ (1998) 98 Columbia Law Review 1148 at 1164. Souter J also quoted Richard Marcus who had stated ‘the original concept of the limited fund class does not readily fit the situation where a large volume of claims might eventually result in judgments that in the aggregate could exceed the assets available to satisfy them’. Later in the quoted article, Marcus draws further on the bankruptcy analogy by arguing that the justification for the use of Rule 23(b)(1)(B), when the defendant’s assets were insufficient to cover all the claims, has the attractiveness of inviting something resembling a bankruptcy cram-down to cope with the expected shortfall of assets. Marcus, above n33 at 877. For an interesting discussion on bankruptcy cram-downs and future claims see Jeffrey Davis, ‘Cramming Down Future Claims in Bankruptcy: Fairness, Bankruptcy Policy, Due Process and the Lessons of the Piper Reorganisation’ (1996) 70 American Bankruptcy Law Journal 329.
Importantly the limited fund was not to be used to effect a pro rata distribution to tort victims.

However, the major justification for the rejection of functional equivalence was procedural, that is, constitutional due process rights. First, the certification of a mandatory class followed by settlement of its action for money damages jeopardises the Seventh Amendment jury trial rights of absent class members.\(^{62}\)

Secondly:

and no less important, mandatory class actions aggregating damage claims implicate the due process ‘principle of general application in Anglo-American jurisprudence that one is not bound by a judgment in personam in a litigation in which he [sic] is not designated as a party or to which he has not been made a party by service of process,’ (Hansberry v. Lee, 311 US 32) … , it being ‘our “deep-rooted historic tradition that everyone should have his own day in court,”’ (Martin v. Wilks), 490 US 755, 762…. Although we have recognized an exception to the general rule when, in certain limited circumstances, a person, although not a party, has his interests adequately represented by someone with the same interests who is a party…. [T]he burden of justification rests on the exception.\(^{63}\)

In the result, the Supreme Court found that the respondents had not satisfied their ‘burden of justification’, stating that it had raised due process concerns a decade earlier in Phillips Petroleum v Shutts\(^{64}\) when the court had stated that an absent class member’s right of action was not extinguishable unless due process requirements were complied with.

4. The Position in Australia

No one has a reliable estimate of how many tens of thousands of asbestos disease cases… have already occurred in Australia…. World authority on asbestos mortality, Professor Douglas Henderson… has estimated Australia will see about 13,000 cases of mesothelioma by 2020, and another 30,000 to 40,000 cases of asbestos-related lung cancer.\(^{65}\)

We have already witnessed a range of corporate responses to asbestos liability in Australia. The most common has been the use of separate legal personality. Australian asbestos manufacturers arranged their corporate groups after the dangers of asbestos had already been recognised overseas. Therefore, from its inception, the mining and manufacture of asbestos in Australia tended to be conducted by the undercapitalised subsidiaries of firms. The following discussion will focus on the two major private sector defendants — CSR and James Hardie Industries.

---

62 Ortiz, above n39 at 845–846. The Seventh Amendment provides: ‘In Suits at common law, where the value in controversy shall exceed twenty dollars, the right of trial by jury shall be preserved’.

63 Ortiz, above n39 at 846.


A. The Australian Asbestos Defendants

(i) James Hardie Industries

James Hardie opened its first Australian asbestos factory in 1916. Although the dangers of asbestos were documented in the US and the UK before the 1920s, the first known death of a Hardie employee due to asbestos occurred about 1960. In 1964 a safety officer employed by the company wrote a long memo to senior management warning that ‘asbestos dust is one of the most dangerous of all industrial poisons’. Although James Hardie experimented with fibres to replace it, asbestos was far more profitable. The first asbestos compensation claims were filed against CSR and James Hardie in 1977. However, Hardie did not stop manufacturing asbestos until 1987. It is estimated that approximately 2000 claims have been made against James Hardie by the beginning of 2001.

By 1998, Hardie was interested in listing on the New York Stock Exchange. It sold its core businesses to a Dutch-based company in order to ‘maximise the tax benefits to shareholders’. A new holding company, James Hardie NV, was created ‘to hold all core assets with asbestos liabilities left in the Australian company’. The plan was for Hardie’s Australian shareholders to own 85 per cent of NV and the remaining 15 per cent to be offered to new US investors in the US float. However, even after the ‘troublesome asbestos liabilities [had] been left behind in the Australian-listed vehicle’ there was weak demand on the US market for the Hardie shares. On 5 January 2000, the *Australian Financial Review* reported that ‘as the international number crunchers run the rulers over Hardies, they are likely to find some poison pills primarily relating to asbestos liabilities’.

On 16 February 2001, James Hardie announced the creation of the Medical Research and Compensation Foundation. The CEO of James Hardie, Mr Peter Macdonald, stated that ‘the establishment of the … foundation provides certainty for people with a legitimate claim against the former James Hardie companies which manufactured asbestos products. Its establishment has effectively resolved James Hardie’s asbestos liability and this will allow management to focus on growing the company for the benefit of all shareholders.’ Mr McDonald also stated that the $293 million was not derived from

---

67 Hill, ibid.
70 Hughes, ibid.
73 James Hardie Australia Pty Ltd, above, n11.

James Hardie is confident that the creation of the trust has eliminated its potential liability for asbestos. Accordingly, no provision was made in its 2001 and 2002 annual reports, revealed by the notes accompanying the financial statements.

With the establishment and funding of the Medical Research and Compensation Foundation... the Company no longer owns or controls two Australian companies which manufactured and marketed asbestos-related products prior to 1987.... As these two former subsidiaries are no longer a part of James Hardie, and all such claims have been successfully defended by JHIL, no provision has been established in the Company’s accounts at March 31, 2001 or 2002.\footnote{James Hardie, Notes to Consolidated Financial Statements, 2002 Annual Report: <http://www.ir.jameshardie.com.au/JH_ir/pdf/2002/Financials/09Notes_to_Financials.pdf> (12 January 2003) at 59. See also James Hardie, Notes to the Financial Statements, 2001 Annual Report: <http://www.ir.jameshardie.com.au/JH_ir/pdf/AR2001_NOTES TO THE_FIN_STATEMENTS.pdf> (12 January 2003) at 55.}

(ii) CSR Limited

CSR is another major defendant in Australia that was involved in asbestos mining and the sale of asbestos-containing products. Most importantly, it conducted the Wittenoom mine in Western Australia through its subsidiary, Australian Blue Asbestos Ltd (ABA).\footnote{The deposits at Wittenoom had been discovered by Lang Hancock and his companies began pick and shovel mining there in 1936.} CSR held all the shares\footnote{With the exception of a few shares issued to the directors of ABA.} and on 23 June 1943 ABA appointed CSR ‘to Act as its Managing Agent and Sole Distributor with full and absolute authority to do all things necessary for the proper management and control of the business and undertaking of [ABA]’.\footnote{CSR v Young, above n7 at 64952 (Handley JA).}

Wittenoom has particular importance because the liability of CSR is potentially extensive and likely to move in waves as new classes of plaintiffs establish liability. For example, the liability of CSR for injuries suffered to the townspeople of Wittenoom was established in 1998 by the New South Wales Court of Appeal in CSR v Young.\footnote{Ibid. In that case, the plaintiff's family had moved to Wittenoom in May 1959. The plaintiff (Mrs Olson) was born in September 1959 and was exposed to blue asbestos from birth until the age of 27 months. Her father was employed by ABA as a draftsman and frequently returned from work with asbestos fibres on his clothing. The back yard where the children played was spread with tailings which contained asbestos fibre. Mrs Olson was diagnosed as suffering from malignant mesothelioma at the age of 34 years. She died nine months later.} The Court of Appeal found that CSR owed a duty of care to the people living in the town which was coextensive with that owed by ABA. The court documented the knowledge possessed by CSR of the dangers of asbestos
prior to and during the mining operations and found that it had caused the injuries suffered by the plaintiff.

The finding is significant because crocidolite tailings from the tailings dump of the mine were spread around the town as a cheap gravel and sand substitute, particularly around buildings and public housing. As stated by Snell:

The tailings were used … in public works on local roads around the town, at the golf course, racecourse, airport, drive in cinema and caravan park. Tailings were also used for concrete slabs and pads for houses. Residents spread tailings around domestic driveways, gardens and yards. Domestic activities like mowing the lawn, children playing on the ground or pets digging stirred up the dust. The airport was very hazardous because of the dust from aircraft taking off and landing.

Although the mine was closed in 1966, Wittenoom’s contamination with crocidolite appears to be unique. The extensive use of large quantities of milled crocidolite as land fill and in construction has not occurred anywhere else in the world.

Another facet of CSR’s liability concerns injuries suffered by US citizens in handling asbestos mined at Wittenoom. Prior to 1990, employees of Johns-Manville commenced the vast majority of the asbestos claims against CSR. Manville was the principal purchaser of the asbestos sold by CSR. Subsequent claims were for injuries caused by the finished products of Manville where the plaintiff alleged that CSR asbestos had been used during manufacture. CSR believes that both the claims in both categories — worker and finished product — are attributable to Manville’s bankruptcy.

Despite a potentially burgeoning liability, on 21 May 2001 CSR announced that it did not wish (at this stage) to adopt the same model as James Hardie to deal with the problem: ‘We obviously watch the situation carefully,’ Mr Brennan said. ‘We don’t see the James Hardie solution, in inverted commas, as something we want to pursue. ‘We might do something similar at some stage but certainly not now.’

CSR’s annual report for 2001 included a $65 million increase in its product liability provision which covers asbestos claims. The company stated:

Management is of the opinion that asbestos litigation in the United States and Australia will not have a material adverse impact on CSR Group’s financial condition. However, an adverse outcome in a single case may negatively affect CSR Group’s earnings in a particular reporting period and there can be no

81 CSR v Young, above n7 at 64942–624947 (Handley JA).
83 Snell, ibid.
84 CSR Ltd, Annual Report to Securities and Exchange Commission, above n68 at 52.
assurance that the litigation will not have a material adverse impact on the CSR Group’s financial condition.  

Accordingly, CSR has entered into a number of global settlements of the Manville worker claims with various claimants’ counsel.

B. Relevant Australian Insolvency Law

When faced with the need to reorganise, defendants in Australia would generally choose either a voluntary administration under Part 5.3A or a scheme of arrangement under Part 5.1 Corporations Act 2001 (Cth). Voluntary administration is generally the procedure of choice as it does not involve the leave of the court and is therefore less costly. The effect of the provisions is to impose a short ‘breathing space’ within which an insolvency practitioner (the administrator) may examine the company’s position and make an informed recommendation to the company’s creditors. Over the last few years the procedure has been used for large complex insolvencies, exemplified by Ansett Australia Ltd.

The other question about the voluntary administration procedure in this context is whether companies must be insolvent to use it. If so, the procedure would be distinguishable from Chapter 11 of the US Bankruptcy Code. Under Corporations Act ss436A–436B, an administrator is generally appointed by a resolution of the board that, in the opinion of the directors voting for the resolution, the company is insolvent or is likely to become insolvent at some future time and an administrator should be appointed. According to Crutchfield, the administration procedure is open to companies which are not insolvent at the time when the resolution is passed, provided that the directors are satisfied that there is a likelihood that the company will become insolvent at some future time. However, Keay has stated that it could be argued that many companies might become insolvent in the future and it may be preferable to provide a definite requirement or else merely state that insolvency in the future is likely.

Australian judges could adopt the American approach and interpret the ‘future time’ requirement flexibly to facilitate reorganisation without proof of insolvency or they might require insolvency as a precondition to the use of the procedure. The latter approach has been taken in the United Kingdom. In the UK an administration order may be made under Insolvency Act 1986 (UK) s8 if the elements of s123 of the same act are satisfied. Pursuant to the tests in s123, the company must be unable to pay its debts as they fall due. Either a ‘cash flow’ test

86 CSR Ltd, Annual Report to Securities and Exchange Commission, above n68 at 10.
87 CSR Ltd, id at 59.
88 The main task of administrators is to convene a series of creditors’ meetings which decide the company’s future. These meetings must be convened within a short period. Section 439B(2) of the Corporations Act 2001 (Cth) provides that the creditors’ meetings should take place within 60–88 days from the date of appointment.
or a ‘balance sheet’ test achieves proof of insolvency. A cash flow test, which is concerned with the existing situation, will be of limited utility because it will only be concerned with debts which presently fall due. However, on the ‘balance sheet’ test the company will be deemed to be insolvent if it is proved to the satisfaction of the court that the value of the company’s assets is less than the amount of its liabilities, taking into account contingent and prospective liabilities. Brown suggests that contingent and prospective liabilities must be taken into account, though it is not clear whether s123 requires that a value be attached to them.\(^92\) Section 8(1) of the Insolvency Act 1986 (UK) also allows an administration order to be made where the ‘company is likely to become unable to pay its debts’. This provision gives rise to the time frame problems discussed above. The predictive element in s8 in combination with the uncertainties surrounding the valuation of contingent and prospective liabilities suggests that the exercise ‘cannot be approached too scientifically’.\(^93\)

However, both Australian and UK insolvency regimes have supplanted procedures that require the leave of the court. The practicalities of insolvency management have favoured the evolution of cheap, expeditious procedures that may be activated by companies and creditors. An undesirable consequence of this development is the removal of the judicial gatekeeping role.\(^94\) Under previous regimes judges could scrutinise the proposed reorganisation to check if it is desirable in all the circumstances. The current regime in Australia does not contemplate any routine monitoring by the courts. The UK system now requires that certain documents be filed in court, for example, a statutory declaration evidencing insolvency\(^95\) and administrators’ reports,\(^96\) but it is difficult to ascertain whether the courts will actively monitor insolvent administrations or merely act as information registries. Although both the Australian and UK regimes provide opportunities for creditors to challenge the administrations in court,\(^97\) the difficulties of ensuring effective representation of present and future claimants abound. This will be discussed in further detail below.

### C. Australian Class Action Procedures

The basic scenario for class actions in Australia has been that opt-in procedures occur in the Supreme Courts and the Federal Court has opt-out procedures under Part IVA of the Federal Court Act 1976 (Cth). Additionally, opt-out procedures are

---

92 David Brown, Corporate Rescue (1996) at 57.
93 Brown, id at 58.
95 Enterprise Act 2002 (UK) s248 para 27.
96 Enterprise Act 2002 (UK) s248 paras 53, 54, 55.
97 In Australia the court has power to make orders where the proceedings are being abused (Corporations Act 2001 (Cth) s447A(2)(b)), where the administrator is managing the company’s affairs in a manner prejudicial to the company’s creditors (Corporations Act 2001 (Cth) s447E(1)), or generally to protect creditors (Corporations Act 2001 (Cth) s447B). In the UK, creditors may apply to the court to challenge an administration where the administrator is acting so as to ‘unfairly harm the interests of the applicant’ Enterprise Act 2002 (UK) s248 para 74.
available in South Australia and in the Victorian Supreme Court. Australia has had a few mass torts based on product liability and major accidents such as the explosion at the Longford gas plant in 1998 which deprived Melbourne of gas for several days. However the asbestos litigation in Australia has been overwhelmingly conducted on an individual basis. This is probably because, firstly, the quantum of damages that may be awarded to an individual litigant in an asbestos claim is often significant and therefore class action procedures are not justified on a cost-benefit basis. Secondly, a significant portion of the asbestos work is conducted in New South Wales and Western Australia, which do not have opt out procedures.

There is a limited fund rule in Part IVA of the Federal Court Act and the Supreme Court Act 1986 (Vic) which are both modelled on Rule 23 (b)(1)(B) US FRCP. There have been settlements of group proceedings effected upon a limited fund model, for example in Lopez v Star World Enterprises Pty Ltd.

98 For example, Bright v Femcare [2002] FCAFC 243 where a group of 61 women have sued the manufacturer of the Filshie clips. Another example is the action brought against a group of three tobacco manufacturers in Philip Morris (Australia) Ltd v Nixon [2000] FCA 229. The applicants claimed that they had contracted a smoking-related disease in consequence of being influenced by the respondents’ conduct to begin or continue smoking. However, the Full Federal Court found that the action could not continue as a group proceeding because Federal Court Act 1976 (Cth) Part IVA does not contemplate group proceedings against multiple defendants.

99 See, for example, Johnson Tiles Pty Ltd v Esso Australia Ltd [1999] FCA 56.

100 Compare mass torts to securities class actions or consumer class actions where the number of affected people may be large but each person’s loss is too small to justify individual proceedings. In relation to securities class actions in Australia, see Peta Spender, ‘Securities Class Actions: A View from the Land of the Great White Shareholder’ (2002) 31 Common Law World Review 123.

101 Both sections 33ZA (ie, s33ZA of the Federal Court Act 1976 (Cth) and s33ZA of the Supreme Court Act 1986 (Vic)) state as follows:

Constitution etc of fund

(1) Without limiting the operation of subsection 33Z(2), in making provision for the distribution of money to group members, the Court may provide for:

(a) the constitution and administration of a fund consisting of the money to be distributed and

(b) either:

(i) the payment by the respondent of a fixed sum of money into the fund; or

(ii) the payment by the respondent into the fund of such instalments, on such terms, as the Court directs to meet the claims of group members; and

(c) entitlements to interest earned on the money in the fund.

(2) The costs of administering a fund are to be borne by the fund, or by the respondent in the representative proceeding, as the Court directs.

(3) Where the Court orders the constitution of a fund mentioned in subsection (1), the order must:

(a) require notice to be given to group members in such manner as is specified in the order; and

(b) specify the manner in which a group member is to make a claim for payment out of the fund and establish his or her entitlement to the payment; and

(c) specify a day (which is 6 months or more after the day on which the order is made) on or before which the group members are to make a claim for payment out of the fund; and

(d) make provision in relation to the day before which the fund is to be distributed to group members who have established an entitlement to be paid out of the fund.
However in that case the defendant was in liquidation by the time that settlement was achieved and the Court ordered that members of the class could opt out of the settlement. Similarly, the judgment of Goldberg J in *Williams v FAI Home Security Pty Ltd (No. 4)* shows considerable caution in dealing with the res judicata effects of potential settlement between the plaintiff and defendant firms. To my knowledge, global mandatory settlements along the lines attempted in *Amchem* and *Ortiz* have not been attempted in Australia, but there is considerable potential for defendants to do so.

High Court decisions on class actions indicate that the High Court is willing to facilitate the use of class actions, although the thorny issues about individual rights in the class action context are yet to be considered. Constitutional issues, such as the individual right to notice and right to have damages assessed individually, have been considered by the Full Court of the Federal Court and the Victorian Court of Appeal respectively. Both jurisdictions have taken a cautious but pragmatic approach. However, in the view of the writer, a genuine controversy about due process entitlements in class actions has not yet surfaced at the appellate court level in Australia.

5. **The United Kingdom — The Manville Option Crosses the Atlantic**

On 1 October 2001 Federal-Mogul Corporation announced that the company had voluntarily filed for financial restructuring under Chapter 11 of the US Bankruptcy Code in order ‘to separate its asbestos liabilities from its true operating potential’. In addition, Federal-Mogul subsidiaries in the United Kingdom had filed jointly for Chapter 11 and administration under the *Insolvency Act 1986* (UK). The Press Release assured readers that it was business as usual and opened with the following list:

103 The sum of $750,000 was to constitute a fund to be distributed amongst the members of the group who did not opt out of the proceeding and who could establish that they had suffered an injury as a result of consuming contaminated food products manufactured by the respondent (a hot bread bakery). The fund was to be administered by Mr Ian Hone, a solicitor employed by Slater & Gordon, the solicitors for the applicant. Each member of the group was required to submit a written claim verified by medical reports or medical certificates. The quantum of damages to which each group member was entitled would then be assessed and the fund distributed *pari passu*. Any decision made by Mr Hone under the settlement scheme was not subject to review by the court.
106 *Femcare Ltd v Bright* [2000] FCA 512 (19 April 2000).
107 *Schutt Flying Academy (Australia) Pty Ltd v Mobil* [2000] VSCA 103 (8 June 2000).
• Business Will Continue to Operate Without Interruption
• No Job Losses or Facility Closures Directly Resulting From the Filings …
• Supply Chain Remains Strong

A Press Release issued less than two weeks later boasted that Federal-Mogul had won four contracts valued at more than $20 million annually.109

One of the subsidiaries seeking an administration order was T&N Ltd,110 one of the largest asbestos defendants in the UK.111 Further, like the other asbestos defendants discussed above, the history of T&N demonstrates the usual pattern of first-hand knowledge of the risks of asbestos exposure accompanied by concealment of the risks for commercial ends.112

By 1994, the cost to T&N of asbestos litigation was spiralling. In that year, the company announced a net loss of £16.5 million due to an increase in asbestos-related costs.113 In 1996, the share price was 138p and the market value of the company was depressed by the huge cost of settling asbestos liabilities.114 The company responded by taking out an insurance policy which effectively ‘collared’ the asbestos liability at about £1.2 billion, of which T&N contributed £300 million.115

By October 1997 the share price of T&N had risen to around 240p on the basis of takeover speculation.116 A formal bid was made by Federal-Mogul on 16 October 1997 at the price of 263p per share.117 Although the City had been worried about the drain on T&N’s cashflow from asbestos litigation, the Chairman of T&N argued that ‘the City got it wrong and [Federal-Mogul] got it right.’118 The Times commented that ‘the City had allowed its paranoia about asbestosis liabilities to obscure the true value’ of T&N.119

110 Formerly Turner & Newall Ltd.
111 Between 1920 and 1976, T&N was the largest British employer in the asbestos industry. By 1950 T&N had achieved a virtual monopoly position in the UK asbestos industry; in 1955 T&N’s market share of the sales of asbestos cement products was 75 per cent and 50 per cent of friction materials. For a time T&N’s sales overtook Johns-Manville. In 1959, while Johns-Manville had sales of $304.1 million, T&N had sales of $450 million. L Kazan-Allen, ‘T&N Ltd’: <http://www.btinternet.com/~ibas/Frames/f_lka_tn_dev.htm> (12 January 2003).
113 London Hazards Centre, above n20 at Chapter 9.
117 Carl Mortished, ‘Big investors unhappy over Pounds 1.5bn T&N bid’ The Times of London (17 October 1997) at 25.
118 Carl Mortished, ‘City lets home-grown success story fall to the Americans’ The Times of London (17 October 1997) at 29.
However, by October 2001, the value of Federal-Mogul had slumped from US$5 billion to a low of $50 million.\textsuperscript{120} The Times had changed its tune somewhat and with the wisdom of hindsight stated:

[the curse [of asbestos] seems, however, to have beaten all calculations. The London stock market never really rated Turner & Newall as it might a company without a fatal flaw. Federal-Mogul did. The US motor components group never made asbestos products but bought into asbestosis by taking over T&N and others. Now it too is seeking protection from its asbestosis creditors as well as challenging their claims in court.]\textsuperscript{121}

The administration order had the effect of freezing present and future asbestosis claims. A lawyer acting for plaintiffs said that several of the company’s settlement cheques had bounced and they could not now say when, if ever, the money due would be paid since there was some doubt about the company’s insurance coverage.\textsuperscript{122} In an Early Day Motion lodged in the House of Commons on 14 November, 41 MPs deplored ‘the fact that the company can remain trading and protect the interest of its shareholders but can abnegate its responsibilities towards its former employees who are now suffering from chronic and terminal asbestos-related disease[s]’.\textsuperscript{123}

6. Teasing Out the Dilemmas

[The court] feared on the one hand that an early reorganization might be misused by a firm to escape its tort liability. Yet it feared on the other hand that a failure to reorganize early could cripple the firm and leave later tort victims with empty claims against a corporate charter.\textsuperscript{124}

A. Introduction

The asbestos litigation established that many defendant manufacturers in the US, Australia and the UK knew at an early stage about the dangers of asbestos and made a commercial decision to keep producing it, thereby jeopardising lives. The tort system has a role to play to deliver corrective justice by providing compensation to those who were injured by their acts. However, whilst an individual tort victim could collect his or her golden egg, the escalation of the number and size of claims threatened to ‘kill off’ defendant enterprises. The

\begin{itemize}
  \item \textsuperscript{120}James Moore, ‘Federal-Mogul forced into administration’ The Times of London (2 October 2001) at 26.
  \item \textsuperscript{121}Patience Wheatcroft, ‘Asbestosis’ The Times of London (2 October 2001) at 25.
  \item \textsuperscript{122}Frances Gibb, ‘Asbestos firm freezes payouts for victims’ The Times of London (12 November 2001) at 13. However, on 10 May 2003 the English High Court ruled that claims for asbestos-related diseases were covered by T&N’s insurance policy. Robert Orr and Nikki Tait, ‘RSA and Lloyds Face Millions in Asbestos Payments’, Financial Times (10 May 2003) at 1.
  \item \textsuperscript{123}Early Day Motion No 423, Turner & Newall and Compensation Claims For Asbestos-Related Disease, House of Commons, 14 November 2001: <http://edm.ais.co.uk/weblink/html/motion.html/EDM1_SES=01/ref=423> (12 January 2003).
\end{itemize}
defendants’ response was to use insolvency and class action procedures to aggregate present and future claims and to isolate those claims from the main enterprise through a limited fund. This led to undercompensation of the tort victims, whereas the enterprise remained viable or even thrived.

In the following discussion the dilemmas raised by this scenario will be explored. It is clear that the right of the tort victim to receive compensation is relational to the entitlement of the corporation to collectivise claims and provide less compensation. But when should the corporate entitlement be exercised? I will argue that it should only be exercised upon proof of genuine insolvency, in order to protect the norms underpinning insolvency law, for example, debtor autonomy and fresh start. This precludes the class action alternative. As the US Supreme Court made clear in Ortiz, a corporation cannot act unilaterally to create a limited fund. Thus the need for caution in relation to the arrangements made by James Hardie. As stated above, the power to collectivise must be based on a genuine insolvency. If we respond honourably to the golden egg argument and decide that present sacrifices should be made for future claimants, we must ensure an equitable distribution between present claimants, future claimants and corporate stakeholders. In particular, we must be wary of strategic behaviour and wealth transfers, as discussed below.

B. Responsibility in Tort

A citizen’s obligation to compensate an injured person when the citizen has committed a tort is clear. Honoré argues that the tort system is one means by which the state, on behalf of the community, seeks to reduce conduct that it sees as undesirable. If the state is justified in making certain types of conduct criminal and attaching to it penalties that may include prison, it must also be justified in marking conduct as tortious and attaching to it the lesser sanction of compensation. Legal sanctions respond to legal liability and more importantly, to historic responsibility. Arguably, the tort system must deliver corrective justice in the asbestos context. ‘Corrective justice’ refers to a system of rights and obligations of repair. Here asbestos defendants concealed the dangers of asbestos and thereby injured successive generations of workers and others. Corrective justice requires those who have without justification harmed others by their conduct ‘to put matters right’.

However, the obligation of defendants to put matters right operates within a complex web of practices about taking responsibility and holding persons

---

126 Peter Cane, Responsibility in Law and Morality (2002) at 44.
128 Honoré, above n125 at 73. Clearly the obligation to put matters right is complex. However, in the present context it would include the payment of individual loss by way of damages or agreed settlement. Cane argues that the process of deciding what sanction ought to attach to any particular finding of legal liability provides important opportunities for the law, by establishing ‘scales’ or ‘degrees’ of responsibility, to modify and fine-tune responsibility judgments made in adjudicating upon liability. Cane, above n126 at 44. Clearly the aggregation of claimants will undermine such responsibility judgments.
responsible, as well as relationships which flow from these practices.\(^{129}\) By way of example, the obligation of the tortfeasor to put matters right is not infinite and it is well accepted that the solvency of the tortfeasor marks the outer limits of this obligation in practical terms. This is an important aspect of litigation practice — the search for the ‘deep pocket’ defendant or for a pool of funds provided by an insurance policy. The capacity to seek the benefits of insolvency is an important right of the tortfeasor which protects his or her autonomy.\(^{130}\)

C. How and Why Insolvency and Class Action Procedures Ameliorate Enterprise Threatening Liability

In this part of the article I will primarily focus upon the principles of insolvency. However, defendant class actions and insolvency procedures are often functionally equivalent because under both procedures the debtor/defendant is granted the power to:

1. collectivise the claims of the tort creditors (generally through a limited fund); and
2. discharge the responsibility to pay compensation because of the operation of the ‘fresh start’ principle under bankruptcy law and res judicata under class action law.

The first power listed above gives rise to two independent effects that occur in long tail mass torts:

a) The collectivisation will almost inevitably lead to under-compensation of tort creditors. Not only does this lead to inequity,\(^ {131}\) it also results in a shift to other tortfeasors\(^ {132}\) who may have less culpability. In the class action context it leads to complex questions about maintaining the collectivisation when an individual asserts his or her right to opt out of the proceedings.

b) Over the long term operation of the collectivisation, the class will have a fluctuating population and further members of the class will make claims. Very important questions arise about how we can protect those future interests.

---

\(^{129}\) Cane, above n126 at 281.

\(^{130}\) With regard to corporate insolvency, some have argued that tortious liability should be regarded as something different in kind to other debts and that shareholders should be liable in particular circumstances for tort claims. See, for example, Henry Hansmann & Reinier Kraakman, ‘Towards Unlimited Shareholder Liability for Corporate Torts’ (1991) 100 Yale LJ 1879 at 1919; Robyn Carroll, ‘Corporate parents and tort liability’ in Michael Gillooly (ed), The Law Relating to Corporate Groups (1993) 91. See also the statements of Rogers AJA in Briggs v James Hardie & Co Ltd (1989) 7 ACLC 841 at 863–864.

\(^{131}\) For example, as discussed above, the Manville Trust only pays 5 per cent of a claimant’s entitlement. See <http://www.mantrust.org/FILINGS/Q3_01/3RDQTR01.HTM> (12 January 2003).

\(^{132}\) The example provided above was the shift by the employees of Johns-Manville from their employer to CSR, ie, from the manufacturer to the mining company.
Although functionally equivalent, the procedures are morally distinct. Although both are based on a limited fund model, defendant class actions (and the James Hardie foundation) are voluntary and unilateral. Insolvency procedures are beneficial and designed to overcome involuntary adversity. Ultimately a person must earn the power to collectivise and a power to mandate a collectivisation must be carefully scrutinised. This anticipates increased monitoring by the judiciary. These issues will be discussed in more detail below.

D. The Power to Collectivise

Bankruptcy law prevents a costly and destructive race to the firm’s assets by offering a collective proceeding that freezes the rights of all investors in the firm, values them, then distributes them according to a priority scheme that the parties agreed would be used in [this] event…  

Insolvency law at its core is concerned with collective action and is designed to keep individual actions against the insolvent debtor from interfering with the best use of assets for stakeholders as a group. Economic analysts argue that an ‘agreement’ arises which mirrors the bargain that creditors would have made amongst themselves if they were able to negotiate before the onset of insolvency. However, its power to explain the allocation of entitlements of involuntary creditors is debatable, as will be discussed below.

The need for the collectivisation arises because there is a limited pool of assets to which each of the claimants has an equal title. Since the titles are equal, no one of the creditors should be able to use his superior strength or swiftness to get more of the common pool than his just share. Thus a moratorium is placed on claims.

Kilpi describes the ethical underpinning of the collectivisation in the following way:

135 For example, Douglas Baird, Thomas Jackson, Lucien Bebchuk & Barry Adler.
136 Baird, above n133 at 339.
137 Veli Jukka Kilpi, The Ethics of Bankruptcy (1998) at 13. Note that in the Manville reorganisation, the moratorium placed upon claims against the corporation (as opposed to the Trust) was enforced by channelling injunctions. A channelling injunction has been described in the following way:

A channelling injunction steers claimants toward a trust or pool of assets to compensate claimants as it simultaneously steers those claimants away from the reorganized entity. … The channelling injunction reinforces the effect of the discharge while it clearly directs claimants toward a specific fund…. Without the imposition of a channelling injunction, claimants might attempt to pursue individual suits against defendants, unwinding the benefits of collective action and equality of treatment …

[I]nterference in individual debt collection is justified because the target of individual seizure efforts is actually a common pool to which no sole person should have an exclusive right. Ethically it is recognition that the debtor’s insolvency has extinguished her moral right to her property and the joint right of the creditors has taken over. The institution of bankruptcy, by transferring the property to the creditors’ joint control, gives legal expression to this ethical idea.138

Keay argues that the collectivisation allows the general group of unsecured creditors to be the primary beneficiaries of the bankruptcy, but the collectivisation must be accompanied by the equal treatment of creditors who are similarly situated.139 Therefore, in a reorganisation, all general unsecured creditors should share available proceeds pro rata. Future claimants and other tort creditors should be treated on par with other unsecured creditors.140

E. A Fresh Start

Kilpi adopts a Kantian perspective by arguing that insolvency law is an important protection of autonomy. Although it is important that we keep our promises or fulfil our legal responsibilities, the protection of autonomy allows us to break our promises or fail to repay our debts without becoming a slave to the promisee or going to debtors’ prison. Moreover, the concept of autonomy protects us from the full consequences of error, or as Kilpi puts it, ‘going broke, breaking promises’.141

Bankruptcy must allow a person who is in a hopeless financial position to obtain relief from the pressure of creditors and to be given an opportunity to make a fresh start. In the past, the emphasis of bankruptcy laws was to deter default and to punish defaulters. The policy underlying bankruptcy is now rehabilitation rather than retribution.142

Modern insolvency law does not decree that every last cent of the debtor must be spent on paying back debts. There is no rule of law or of morality that requires the tortfeasor to exhaust all of her resources for the sake of compensating tort creditors. This principle applies with particular force when the tortfeasor is undergoing reorganisation because a primary goal of reorganisation is amelioration of “the devastating effect that a huge liability may have on the worth of a business”.143 As stated by Resnick, “the protection of the business enterprise by preserving its going concern value, thereby maximizing value for distribution to creditors, is central to the reorganization process.”144

---

138 Kilpi, id at 13.
139 Keay, above n90 at 17.
141 Kilpi cites a range of responses to ‘going broke and breaking promises’ such as forgiveness and utilitarianism. Kilpi, above n137 at 68–72.
142 Keay, above n90 at 499.
143 National Bankruptcy Review Commission, above n137 at 318–319.
F. Subverting Individual Interests and Protecting Future Claimants

The Manville reorganisation exemplifies the unfairness of persons in the present taking for themselves resources that ought to be reserved for the future. Many of the controversies of our time — the use of natural resources and the preservation of the environment, the status of social security … for example — involve the potential misappropriation of resources rightfully belonging to future persons.145

Protecting the defendant from unrestrained litigation requires the altruism called for by Smith (above) — that we should not misappropriate resources rightfully belonging to future persons.146 The ultimate problem about submitting individual action to collective interests in this context is one of incentive. Although individual plaintiffs may be prepared to take up Smith’s challenge to subvert their own interests to preserve resources for future claimants, there are significant disincentives created by the self-interested behaviour of corporate participants. This is exemplified by strategic bankruptcy and wealth transfers to shareholders which will be discussed later.

A further question arises which is both practical and normative. Can we truly look after the interests of future claimants in this system? In relation to the normative questions, it could be argued that the Supreme Court has foreshadowed its approach by imposing the due process requirements in *Amchem* and *Ortiz*.147 Going beyond the actual decisions in *Amchem* and *Ortiz*, Gibson suggests that there are wider due process and fairness concerns which animate the Court’s reasoning.148 She concludes:

*My enthusiasm for a bankruptcy solution … is tempered by the recognition of the concerns expressed by the Supreme Court in other mass tort contexts about forcing a judicial resolution on claimants who are not active participants in the lawsuit leading to the global solution.*149

The concerns of the Supreme Court in *Amchem* were elegantly expressed. Their Honours raised questions about the efficacy of notice to persons who ‘may not even know of their exposure, or realize the extent of the harm they may incur’ and wondered whether ‘notice sufficient under the Constitution … could ever be given to legions so unselfconscious and amorphous’.150

Roe focuses on the inability of future claimants to effectively negotiate and raises the considerable problem of undercompensation. He states that:

*… the norm of accurate compensation is much more significant in tort reorganizations than in financial reorganizations; indeed, it may be central. In financial reorganizations, the actors — banks, insurance companies, the public shareholders — can bear the risk of a disparate outcome. They have either*

145 Smith, above n24 at 370.
146 Smith, ibid.
149 Gibson, id at 2117.
assumed that risk of miscompensation or can be viewed as having assumed it, and they often can diversify their holdings to offset its impact. None of this can be said about the individual tort claimant. He is far less able to bear such a risk. He cannot easily avoid the risk by marketing his claim, did not assume the risk, and usually cannot diversify.\footnote{Roe, above n\textsuperscript{124} at 877–878.}

There is of course a myriad of practical problems which beset attempts by the courts to assess future claims. Occasionally the courts have appointed an officer to represent future claimants but there are still process hazards in this plan.\footnote{As stated by her Honour Edith Jones in ‘Rough Justice in Mass Future Claims: Should Bankruptcy Courts Deliver Tort Reform?’ (1998) 76 Tex LR 1695 at 1713: In bankruptcy, as in future claims class actions, the claimants are absent, invisible, and passive, creating room for exploitation in several ways. The class representative … operates without the supervision or control of real clients. Because the claims themselves are not concrete, but rather amorphous and conjectural, the representative’s bargaining position is inherently weak. The representation of future claims thus carries with it a tendency toward conflicts of interest. There is no vigorous check on a class representative’s accepting a settlement that provides generous fees for the representative but modest relief for the class. A conflict may arise if the representative undertakes to settle claims of both present and future ‘future’ claimants. Compare Gregory Bibler, ‘The Status of Unaccrued Tort Claims in Chapter 11 Bankruptcy Proceedings’ (1987) 61 American Bankruptcy Law Journal 145 at 150–161 with Ralph Mabey & Jamie Gavrin, ‘Constitutional Limitations on the Discharge of Future Claims in Bankruptcy’ (1993) 44 South Carolina Law Review 745 at 755–759.}

Furthermore, the concept of a ‘future claim’ has been poorly developed in US law\footnote{Matthew Stiegler, ‘The Uncertain Future of Limited Fund Settlement Class Actions in Mass Tort Litigation After Ortiz v. Fibreboard Corp’ (2006) 78 North Carolina Law Review 856 at 886. See also J Tidmarsh, Mass Tort Settlement Class Actions: Five Case Studies (1998) at 77–78 (discussing the failure of a $4.23 billion multi-defendant settlement of breast implant claims caused by a massive underestimation of claim volume).} and there is a natural reticence to delve too far into future uncertainties as well as grappling with present ones. Judicial estimation of mass tort liability has not been an unqualified success,\footnote{A judge of the US Bankruptcy Court captured the problem with the following statement: Many commentators now assume that mass tort bankruptcy cases must follow a certain script: first estimate aggregate tort claims; then negotiate a fund to satisfy it; and finally, devise a procedure for fairly apportioning the funds. … But there are good reasons to question the premise upon which the standard mass tort model for reorganization rests. The model rests on the shaky foundation that judges can accurately estimate the results of a series of extremely speculative problems. While one court ‘seems’ to have pulled it off (A.H. Robins), at least another has predictably failed (Johns-Manville): Spector J, In Re Dow Corning 211 BR 545 (1997), 562.} as evidenced by the enormously inaccurate estimate made by Lifland J of the needs of future claimants in the Manville reorganisation.\footnote{Thomas Jackson, The Logic and Limits of Bankruptcy Law (1986) at 193–194.} As stated above, an amount that was intended to last the lifetime of the trust was exhausted in 18 months.

\textbf{G. Strategic Behaviour}

Jackson suggests that whenever bankruptcy law changes the relative value of ‘nonbankruptcy rights’, incentives are created for inappropriate uses of the bankruptcy process.\footnote{Thomas Jackson, The Logic and Limits of Bankruptcy Law (1986) at 193–194.} Strategic insolvency arises where the bankruptcy is
invoked due to strategic decision-making rather than being a passive response to market forces. Delaney argues that a strategic bankruptcy arises when a firm mobilises the bankruptcy process in order to transform ties with other institutions or groups of individuals.157

Clearly there are enormous incentives for the asbestos defendants to use reorganisation to evade paying compensation to tort victims. As discussed above, Smith considered that the most controversial aspect of the Manville bankruptcy was the premise that the corporation was entitled to bankruptcy protection not because it was insolvent, but because a bankruptcy reorganisation the best way to manage the payment of present and future claims.158

Smith continued:

> Despite its novelty, the premise of the Manville bankruptcy petition made sense as policy. It was difficult to see how the problem of looming future liability could be fairly solved outside of bankruptcy, whatever was the technical legal status of this future liability. … From the outset, fair treatment of future claimants was a major justification for the Manville bankruptcy.159

With respect, the logic of these statements is hard to follow. The arguments demonstrate the danger of allowing the power to collectivise to deviate too far from its normative underpinning. The norms underpinning bankruptcy are discussed above — to protect the autonomy of the debtor, to apply the pari passu principle and to rehabilitate in the case of reorganisations. The power must only be triggered by insolvency because otherwise a coercive collectivisation is unjustifiable. The point is made elegantly by her Honour Edith Jones, a former US Bankruptcy judge and currently Judge of the US Court of Appeals for the Fifth Circuit:

> The question again arises whether the mass future claims proposals have anything to do with bankruptcy, or whether they are a contrivance to shoehorn mass tort litigation into a coercive, collective settlement that preserves management control and shareholder equity.160

As regards bankruptcy, the minimum safeguard must be the need to prove insolvency. In this respect, the UK system is superior to the Australian and US systems because it requires proof of insolvency before the power to collectivise can be exercised. But in addition, judges should act as affirmative gatekeepers of the power to collectivise. In a comparison between US and Japanese insolvency laws, Anderson has stated that ‘Chapter 11 debtors tromp without hesitation through automatic doors into the waiting protections of reorganization, while Japan’s debtors must pass a doorman’s inspection before they are allowed in’.161

Removal of the need to prove insolvency under Chapter 11 of the Bankruptcy Code (US) should not be emulated in other jurisdictions and the power to collectivise should be carefully monitored.

157 Delaney, above n18 at 59.
158 Smith, above n24 at 373.
159 Smith, ibid.
160 Jones, above n152 at 1722.
161 Anderson, above n94 at 403.
The alternative to threshold judicial scrutiny is monitoring by creditors. In Australia, as stated above, a voluntary administration may be challenged on several grounds — inter alia because the company is solvent or because the provisions of Part 5.3A are being abused. The potential of voluntary administration procedures to facilitate strategic behaviour has been recently recognised in Australia, most notably in the MUA v Patrick Stevedores litigation. In the US, Chapter 11 of the Bankruptcy Code provides similar mechanisms to challenge the reorganisation on good faith grounds, but effective representation of tort creditors is almost impossible to achieve.

Perhaps these problems may be resolved by conferring preferred priority status upon these claims. However, the preferred priority rules which currently apply in Australia, the UK and the US at best only apply to injury compensation arising out of an employment relationship; there is no priority for general personal injury tort creditors such as the townspeople of Wittenoom.

II. Wealth Transfers
In each of the case studies discussed above, a jump in the share price accompanied a reorganisation or the announcement of a class action certification. Coffee argued that the 300 per cent increase in Fibreboard’s share price upon certification of the mandatory class action amounted to a wealth transfer. Similarly, the share price of T&N doubled during 1996–1997 and it could be argued that a wealth transfer arose by Federal-Mogul’s speculation upon T&N’s asbestos liabilities. The shareholders clearly profited from the transaction, but the asbestos claims are now in jeopardy.

162 Corporations Act 2001 (Cth) s447A.
166 See generally Susan Cantlie, ‘Preferred Priority in Bankruptcy’ in Jacob Ziegel (ed), Current Developments in International and Comparative Corporate Insolvency Law (1996) at 413.
167 In Australia s556(1)(f) of the Corporations Act 2001 (Cth) provides that priority on liquidation is accorded to ‘amounts due in respect of injury compensation’. However, ‘injury compensation’ is defined in s9 of the same Act to mean ‘compensation under any law relating to workers compensation’. The priority rules in s556 must be applied by an administrator in a deed of company arrangement, pursuant to para 4, Schedule 8A of the Corporations Regulations (Cth). In the UK, s175 of the Insolvency Act 1986 (UK) only extends priority to workers compensation and sick leave payments. This provision applies to administration by virtue of para 65, s248 of the Enterprise Act 2002 (UK). In the US, s507 of the Bankruptcy Code only allows a priority of up to $4650 for sick leave pay.
168 Coffee, above n18 at 1402.
Although shareholders entitlements should be subordinated to creditors upon a winding up or a reorganisation, a study by LoPucki and Whitford found that shareholders of insolvent companies nearly always shared in the distribution under the reorganisation plans of large, publicly held companies. They also found that the relative size of shareholders’ recovery was more attributable to the quality and aggressiveness of equity’s representation than the financial condition of the company.

Given the difficulties with representing the interests of future claimants, the wealth transfer seems insuperable. However, the Australian courts have buttressed the procedural rights of claimants over the last decade. There have been developments in two areas which might be of interest. Firstly, the High Court held in 1994 that a plaintiff’s claim for damages due to negligence causing personal injuries was ‘property’ for the purposes of s51(xxxi) of the Constitution and thus only able to be acquired by the Commonwealth on just terms. The second development involves expansion of Mareva orders. A Mareva order is an interlocutory order which restrains the defendant from removing assets or disposing of assets with such effect that if the plaintiff should obtain judgment against the defendant there will be insufficient assets from which the plaintiff can obtain satisfaction of his judgment. The effect of a Mareva order is to preserve the defendant’s assets (or so much of them as would be necessary to satisfy the plaintiff’s claim) so they will be available to satisfy the plaintiff’s judgment if she obtains it. The Mareva order does not give the plaintiff any lien or other proprietary interest in the assets so preserved. In Cardile v LED Builders P/L, the majority of the High Court stated that the purpose of the Mareva order is to avoid frustration of the court’s process. Whilst neither development would currently allow future claimants to restrain a wealth transfer, they might be a source of further doctrinal development.

7. Conclusion

Asbestos litigation has provided our community with complex legal questions, but more importantly, invidious moral choices. The history of asbestos reveals that many corporations cynically disregarded the health and ultimately the lives of others by making commercial decisions to mine and manufacture asbestos decades after its dangers were publicly recognised. Due to the long latency of diseases like asbestosis and mesothelioma and the early difficulties in establishing liability, the defendants only became answerable under tort law in the 1970s. However, by that
time the number of injured had spiralled, so that by 1982 the largest asbestos manufacturer in the USA, Johns-Manville, decided it needed to protect itself against the escalation in claims.

A twist in the morality story then occurred because Manville argued that it needed to survive in order to pay future asbestos claims and it was critical not to kill off the goose that laid the golden egg. Due to a quirk of US insolvency law, Manville was permitted to reorganise under Chapter 11 of the US Bankruptcy Code, even though it was not insolvent.

Fundamental to the Manville reorganisation was the power to collectivise the asbestos claims, forcing plaintiffs to divert their claims from the corporation itself to a fund set aside for that purpose. This fund proved to be inadequate and after 20 years the sum available to claimants is 5 per cent of their entitlements. Most importantly, the goal of forcing the claimants to collectivise and to satisfy their claim out of a limited fund had been achieved without proof of insolvency.

In the late 1990s, asbestos defendants sought wider powers to collectivise their claims through class actions. They argued that a limited fund should be allocated which should be distributed to asbestos claimants across the US. The class would include future claimants and a moratorium upon claims against the companies would be achieved by operation of the doctrine of res judicata. Although there was no suggestion of insolvency, Manville had already established that insolvency was not critical to the inquiry, rather, it was a question of ‘enterprise threatening liability’.

The attempt to use the class action to coercively collectivise without insolvency was unsuccessful. The Supreme Court in the Amchem and Ortiz cases denied limited fund status purely for a sum that had been set aside by the defendant. The Supreme Court also raised doubts about whether the procedural rights of future claimants could realistically be protected under the proposals.

Recently, the largest asbestos manufacturers in Australia and the United Kingdom rearranged their assets to limit their asbestos liabilities. In the case of T&N, the company was subject to an administration order which flowed from a Chapter 11 reorganisation of its parent company. The effect of the order was draconian — the settlement cheques of tort victims immediately bounced.176

Mass torts in general throw up formidable problems; attempts to solve the problems have consumed a vast amount of law journal space in the last few years. Some commentators suggest that wide-ranging institutional redesign involving the

---

176 In a similar vein, a settlement was finally effected between the South African victims of another major UK asbestos manufacturer, Cape plc, in March 2003. Settlement agreements totalling £10.7 million will be divided among 7500 members of a group action brought in England by people who were occupationally or environmentally exposed to asbestos in South Africa, reported in Lubbe v Cape plc [2000] 4 All ER 268. Although it was conceded that once the money is divided up, it will amount to ‘next to nothing’, the plaintiffs’ solicitor Richard Meeran, agreed that a smaller settlement had to be negotiated with Cape because of the company’s precarious financial situation. See Peter Muchlinski, ‘Corporations in International Litigation: Problems of Jurisdiction and the United Kingdom Asbestos Case’ (2001) 50 ICLQ 1; Laurie Kazan-Allen, ‘Cape PLC To Compensate Foreign Plaintiffs’: <http://www.binternet.com/~ibas/Frames/f_lka_cape_comp_forgn_pla_0303.htm> (18 April 2003).
tort, litigation and insolvency systems might be necessary. For example, Abraham argues that the liability rules in tort need to be remodelled to recognise collective responsibility.\textsuperscript{177} Clearly, there is a danger in attempting to force the mass tort problem into traditional frameworks, but this danger is accompanied in the present debate by a lag in the theoretical development of the law in this area compared with its empirical development.\textsuperscript{178}

In the meantime, in the view of the author, the right to coercively collectivise and to make a fresh start is only morally defensible in situations of genuine insolvency in order to protect the autonomy of the debtor. The power to collectivise and to create a limited fund for tort claims is far reaching and must be carefully monitored by the courts. The case studies reveal that the creation of a limited fund generally results in the under-compensation of tort victims, particularly future claimants. This may be acceptable upon a genuine insolvency, but not by unilateral acts of the corporate defendant. The situation becomes particularly inequitable when wealth is transferred from claimants to shareholders.

Overall, it is a considerable challenge. Commenting on the T&N administration, John Battle, MP for Leeds, reminds us:

> This evasion is a scandal. Making the polluters pay obviously means not only proving responsibility, not only winning the moral and legal arguments, but taking on high level international corporate gamesmanship that continues to write off the lives of asbestos victims.\textsuperscript{179}


Case Note

Western Australia v Ward & Ors

KATE STOECKEL*

1. Introduction

On 8 August 2002 the High Court delivered its judgment in Western Australia v Ward,1 a seminal decision regarding the extinguishment of native title under the Native Title Act 1993 (Cth).2 The central finding of Ward is that the Native Title Act mandates partial and permanent extinguishment — something not established in either Mabo [No 2] v Queensland3 or Wik Peoples v Queensland.4 This judgment establishes how the test for partial extinguishment may be applied in future determinations of native title. The concept of partial extinguishment did not lead, as it did in the earlier Full Federal Court decision, to an extensive debate about the characterisation of native title. However, the High Court did note the usefulness of the ‘bundle of rights’ analogy in native title determinations.

In Ward, the Court made it clear that, for native title determinations under the Native Title Act, primary regard is to be had to the Act.5 Common law principles, such as those established in Mabo and Wik, are only to be used for interpretative purposes. As the application in Ward was brought under the Native Title Act, the case revolved around construing certain terms of the Act, particularly those relevant to full and partial extinguishment. To determine the extent to which partial extinguishment occurred, the Court favoured an ‘inconsistency of incidents’ test before alternatives.6

Several of the justices also found that international law should play only a limited role in determining the rights of Australia’s indigenous people. This approach has wide ramifications, demonstrated by the majority finding that there is no native title right to resources, nor a right to protect indigenous cultural knowledge. However, there are conflicting views from the Bench on the relevance of international law in this area. Justice Callinan’s dissenting judgment is vociferous in rejecting any application of international law or precedent. Justice Kirby, however, pays particular regard to the connection between indigenous rights and international human rights.

* BEd/LLB (Hons). I would like to thank Peter Butt and Andrew Stoeckel for their advice and patience with this note. Any errors remain mine.
1 Western Australia v Ward; Attorney-General (NT) v Ward; Ningarmara v Northern Territory [2002] HCA 28 (8 August 2002) [hereinafter ‘Ward’].
2 Hereinafter ‘the Act’. The Native Title Act was amended by the Native Title Amendment Act 1998 (Cth) [hereinafter the ‘1998 Amendments’].
3 (1992) 175 CLR 1 [hereinafter ‘Mabo’]
4 (1996) 187 CLR 1 [hereinafter ‘Wik’].
5 Ward, above n1 at [25] (Gleeson CJ, Gaudron, Gummow & Hayne JJ); at [564], [567] (Kirby J).
6 Id at [79].
This article outlines the Ward decision, how the Court applied the statutory scheme of the Native Title Act to full and partial extinguishment, and the ramifications of disentangling the ‘bundle of rights’ for Australia’s indigenous people in native title cases. The sheer length of this case demonstrates the complexity of the statutory framework established by the Native Title Act, of which partial extinguishment is but one aspect. This complexity adds both time and cost to all native title applications, decided as they are on a case-by-case basis. Consequently, it is imperative that the current system be reviewed to simplify the law and enhance its expediency.

2. Background to the Case

This case concerned a native title application by three groups of indigenous claimants\(^7\) over a 7900 square kilometre area of land and waters in the East Kimberley region of Western Australia extending into the Northern Territory. Included in the claim area were two substantial economic entities: the Ord River irrigation project and the Argyle Diamond Mine. The area under consideration is a remote area of Australia where ‘a large part of the land resumed or acquired … remains in unaltered form’.\(^8\) Various native title rights were claimed. Some were very wide, such as the right to ‘possession, occupation use and enjoyment of the land’ and the right to ‘speak for country’.\(^9\) Others were narrower, for example sustenance rights, the right to hunt and gather food, and the rights to perform ceremonies and protect cultural knowledge.\(^10\)

The claim was first heard in 1995 in the Federal Court before Justice Lee.\(^11\) At that time the Native Title Act was in its original form. For native title to be recognised under the Act, it must satisfy the elements set out in s223(1).\(^12\) A determination over a particular area is then made under s225. The primary judge held that native title, as determined under s225 of the Native Title Act as it was at the time of the hearing, occurred in respect of ‘a very large portion of the claim area’.\(^13\) Important to the subsequent appeals was that Justice Lee recognised that the nature and extent of native title included the right to control access to the determination area.\(^14\)

---

\(^7\) The first claimants (Ben Ward & others) applied on behalf of the Miriuwung and Gajerrong People, the second claimants comprised Cecil Ningarmara and others, and the third claimants (Delores Cheinmora & others) applied on behalf of the Balangarra Peoples: Ward, above n1 at [33] (Gleeson CJ, Gaudron, Gummow & Hayne JJ).

\(^8\) Id at [36].

\(^9\) The right to ‘speak for country’ is an imprecisely defined Aboriginal concept in the Australian legal system. The concept embodies the broad notion that indigenous people are custodians of the physical and spiritual attributes of the land. It includes several specific rights such as granting permission to enter upon the land. See Ward, id at [14], [89] (Gleeson CJ, Gaudron, Gummow & Hayne JJ); at [592] (Kirby J).

\(^10\) Id at [47] (Gleeson CJ, Gaudron, Gummow & Hayne JJ); at [563] (Kirby J).


\(^12\) See below at text accompanying footnote 55.

\(^13\) Ward, above n1 at [38] (Gleeson CJ, Gaudron, Gummow & Hayne JJ).

\(^14\) Ward v State of Western Australia (1998), above n11 paragraph 3(d) of the determination.
The primary decision was appealed against and the Full Federal Court substituted a new determination. This decision was also appealed. Altogether, four appeals were heard before the High Court. The Aboriginal claimants submitted that the determination of native title should have been more extensive, to reflect their ‘communal native title’ as being ‘practically equivalent to full ownership’ vested in the community. The native title claimed, it was submitted, should also include the right to resources and the right to ‘maintain, protect and prevent the misuse of cultural knowledge’, found by the trial judge but overturned by the Full Court. The opposing parties responded with contrary submissions.

3. The High Court Judgments

Four judgments were handed down. Chief Justice Gleeson and Justices Gaudron, Gummow and Hayne issued a joint judgment. Justice Kirby generally concurred with the majority, but added some minor qualifications. To provide ‘a clear statement of the applicable law’ in the legal quagmire of native title disputes, Justice Kirby put aside his personal preferences on certain issues in accordance with his view that the ‘avoidance of unnecessary doubt and confusion is a proper objective of land law’. The reservations he records are discussed below in the relevant sections. Separate dissenting judgments were delivered by Justice McHugh and Justice Callinan.

3.1 The Majority

(a) Characterisation of Native Title

The High Court did not devote much analysis to the Full Federal Court’s substantial debate regarding the proper characterisation of native title, as the principles of extinguishment focused on by the High Court do not turn on characterisation. However, the High Court did state that any determination on native title must first understand the meaning of the term.

(i) The Spiritual Connection with the Land

There will always be a substantial degree of difficulty in reconciling two very different systems of law. For example, the complex spiritual and ‘organic’

---

16 Two appeals were brought by indigenous claimants, and two by the respondents at trial (the State of Western Australia and the Northern Territory): Ward, above n1 at [54]–[55] (Gleeson CJ, Gaudron, Gummow & Hayne JJ).
17 Id at [83].
18 Id at [57].
19 These qualifications do not, however, affect the final orders of the majority judgment. Ward, id at [599] (Kirby J).
20 Id at [565].
21 Wik, above n4 at 221 (Kirby J).
23 Milirrpum v Nabalco Pty Ltd (1971) 17 FLR 141 at 167 (Blackburn J).
relationship between Aboriginal people and their land defies characterisation as a ‘right’ or an ‘interest’ by its very nature. So how can it be recognised by Anglo-Australian land law? Yet, as noted by the High Court, that is precisely what the Native Title Act requires:

The spiritual or religious is translated into the legal. This requires the fragmentation of an integrated view of the ordering of affairs into the rights and interests which are considered apart from the duties and obligations which go with them.24

The disempowering ramifications of the fragmentation of traditional rights can be seen in the Court’s determination on cultural knowledge.25 The primary judge found the existence of a ‘native title right to maintain, protect and prevent the misuse of cultural knowledge’.26 The Full Federal Court overturned this finding, and the High Court upheld the Full Court’s view by majority. The reason given by the High Court joint judgment is that a native title right to prevent the misuse of cultural knowledge goes beyond the ‘connection’ required by s223(1)(b) of the Native Title Act, discussed below, as it would be ‘akin to a new species of intellectual property’.27 Although the majority noted that the native title right to restrict access to culturally sensitive sites, where not extinguished, could be used to protect cultural knowledge, the majority view was that a regime to protect cultural knowledge must be found under general law and statute.28

Justice Kirby, in disagreement with the majority on this point, found that the connection required under s223(1)(b) could be established through the spiritual connection with the land.29 This view is supported by reference to Australia’s ratification of international human rights instruments,30 which, in Justice Kirby’s opinion, include ‘the right of indigenous people to have “full ownership, control and protection of their cultural and intellectual property.”’31 Justice Kirby’s finding is a rare instance of the practical application of the sui generis nature of native title. To date this concept has been dormant in Australian law, for ‘so far

25 This is particularly a problem when combined with the evidentiary disadvantages that Aboriginal claimants face. Much of indigenous cultural knowledge is secret. Various internal rules apply, such as restricting access to strangers, and even restricting access to sectors of their own community, such as members of the opposite gender: Ward, id at [576] (Kirby J).
26 Ward v State of Western Australia (1998) above n11 paragraph 3(j) of the determination.
27 Ward, above n1 at [59].
28 Ibid.
31 Ward, above n1 at [581] (Kirby J).
claims have generally related physically to land or waters in a way analogous to common law property concepts.\(^{32}\) The possibilities arising from the uniqueness of native title have not yet been fully realised or tested in Australia.

\__(ii)\ The Right to ‘Speak for Country’ and the ‘Bundle of Rights’ Debate\n
The right to ‘speak for country’ was accepted by the High Court as ‘a core concept in traditional law and custom’, that does not, however, exhaustively capture the rights and interests conferred by traditional law.\(^{33}\) The concept is expressed most closely in the common law as the ‘right to possess, occupy, use and enjoy the land to the exclusion of all others’.\(^{34}\) However, defining the right to ‘speak for country’ in this way gives rise to a tendency to ‘break the expression into its constituent elements’.\(^{35}\) This, in turn, may see the right to speak for country misconstrued as a ‘bundle of rights’ rather than as a relationship with the land.\(^{36}\)

Whether native title can be seen as a ‘bundle of rights’ was comprehensively discussed before the Full Federal Court. Although the \textit{sui generis} nature of native title was recognised in \textit{Mabo},\(^{37}\) Justices Beaumont and von Doussa in the Full Federal Court took the view that native title constitutes a ‘bundle of rights’.\(^{38}\) Even though this view was qualified as ‘not to deny the possibility that in a particular case the rights and interests may be so extensive as to be in the nature of a proprietary interest in land’, it does not adequately reflect the indigenous relationship with the land.\(^{39}\) This view was in contrast to the primary judge’s characterisation of native title as ‘a communal “right to land” arising from the significant connection of an indigenous society with land under its customs and culture’.\(^{40}\)

Justice North in the Full Federal Court agreed with the primary judge’s characterisation, regarding communal native title rights as ‘[reflecting] the traditional law of the Aboriginal people’.\(^{41}\) There is precedent to support this view. For example, Justice Brennan in \textit{R v Toohey; ex parte Meneling Station Pty Ltd} held that ‘Aboriginal ownership is primarily a spiritual affair rather than a bundle of rights’.\(^{42}\) The relevance of this debate becomes apparent in the present case when considering extinguishment. If the \textit{sui generis} characterisation of native title as a relationship to the land were followed, it would be difficult to establish a rationale for the partial extinguishment of native title rights. However, the High Court took the view in this case that the ‘bundle of rights metaphor’ was ‘useful’:

\(^{32}\) \textit{Ward}, id at [578] (Kirby J).
\(^{33}\) \textit{Ward}, id at [88],[90] (Gleeson CJ, Gaudron, Gummow & Hayne JJ).
\(^{34}\) Id at [88].
\(^{35}\) Id at [89].
\(^{36}\) Id at [89], [93].
\(^{37}\) \textit{Mabo}, above n3 at 89 (Deane & Gaudron JJ). This approach is taken from Dickson J in the Canadian case of \textit{Guerin v The Queen} [1984] 2 SCR 335 at 382.
\(^{38}\) \textit{Western Australia v Ward}, above n15 at [90].
\(^{39}\) Id at [96]–[97].
\(^{40}\) \textit{Ward v State of Western Australia}, above n11 at 508 (Lee J).
\(^{41}\) \textit{Western Australia v Ward}, above n15 at [784].
\(^{42}\) (1982) 158 CLR 327 at 358 (emphasis added).
it highlights the possibility that more than one right may exist, and more than one type of right may be present. This approach warrants extinguishment of one or more rights whilst leaving others intact.

(b) Principles of Extinguishment

(i) Prior to the 1998 Amendments

Mabo established that, regardless of earlier precedent, native title exists in Australia and is recognised by the common law. Further, it continues to exist until extinguished by clear and plain legislative intent. Mabo’s principles meant that suddenly many native title issues arose for urgent determination. Of relevance to this case was how the pastoral lease — a peculiar statutory creature covering large swathes of territory — interacts with native title. The High Court decision of Wik held that pastoral leases did not necessarily confer the right of exclusive possession upon the grantee, and could thus coexist with native title. Ward takes up the issue of the extinguishment of native title under the pastoral lease, and particularly whether partial extinguishment is possible.

It is the very source of native title, namely traditional Aboriginal laws and customs, that also gives rise to its central weakness — its susceptibility to extinguishment. As native title is not derived from Crown sovereignty and is not held of the Crown, it does not enjoy the same protection and priority as Crown tenures. It is extinguished by ‘a valid exercise of sovereign power inconsistent with the continued enjoyment or unimpaired enjoyment of native title’, including Crown actions that evince a clear and plain intention to extinguish such title. The purpose of extinguishment, as noted by Justice North in the Full Court:

is to ensure that inconsistent rights or interests have priority over native title and override native title ….Where those rights or interests can be enjoyed without abolition of native title there is no place for extinguishment.

Extinguishment, once effected, is irreversible. Justice Kirby offered four explanations for the irreversibility in Fejo, a case where the inconsistent grant was a statutory estate in fee simple. Firstly, native title is a fragile legal property right recognised by the common law as highly susceptible to extinguishment. Secondly, the granting of an estate is an ‘assertion of the sovereign rights of the grantor to establish its power in respect of the land and to exclude any claim not

43 Ward, above n1 at [95] (Gleeson CJ, Gaudron, Gummow & Hayne JJ).
46 Wik, above n4 at 85 (Brennan J).
47 Western Australia v Ward, above n15 at 330-331 (North J).
49 Fejo, id at 155 (Kirby J).
2003] NOTE 261

specifically granted by it. 50 Thirdly, for native title to exist after extinguishment, a common law rule would be required for its reinstatement. As native title is not derived from the common law but from traditional Aboriginal laws and customs, ‘[t]he conferral of… new rights by common law would be completely incompatible with the notion that native title rights have their origin in Aboriginal custom: not in the Australian legal system.’ 51 Finally, non-revival of extinguished native title recognises the legal history of Australia’s land law and the inability of the courts to interfere with its central principles. 52 As property jurisprudence is an area of law where ‘there is a very high premium on certainty’, 53 the principle of non-revival reflects both this history, and practical and policy considerations.

However, the irreversibility of extinguishment for native title has dire consequences for indigenous communities for, as seen above, the concept of land within Aboriginal culture is intrinsic to every aspect of indigenous culture. Arguably, the denial of native title rights has a further-reaching and more serious effect than the denial of non-indigenous interests.

(ii) Statutory Framework for Extinguishment

One of the primary functions of the Native Title Act is to provide certainty for non-indigenous Australians who have been granted interests in land that is potentially subject to native title. Certainty in tenure is at the heart of Australian land law, providing strong incentives for economic growth and prosperity. However, the indigenous perspective reveals that ‘certainty’ is really a euphemism for ‘extinguishment’.

The High Court made it clear that extinguishment is to be governed by the statutory scheme of the Native Title Act. 54 Section 11(1) of the Native Title Act states that native title cannot be extinguished contrary to the Act. In order for native title to be protected, the rights and interests it encompasses must first come within the definition of native title in s223(1): 55

(a) The rights and interests are possessed under the traditional laws acknowledged, and the traditional customs observed, by the Aboriginal peoples or Torres Strait Islanders; and

(b) The Aboriginal peoples or Torres Strait Islanders, by those laws and customs, have a connection with the land or waters; and

(c) The rights and interests are recognised by the common law of Australia.

Regarding s223(1)(b), it is not necessary that the connection with the land or waters be of continual usage or physical occupation, 56 although s13(1) of the Act

50 Ibid.  
51 Ibid.  
52 Id at 156.  
53 Ward, above n1 at [565] (Kirby J).  
54 This consists of the Native Title Act, the related State and Territory Acts that supplement it, and the Racial Discrimination Act 1975 (Cth).  
56 Id at [63].
provides that any judicial determination of native title can be varied or revoked if the continued connection with the land is disrupted.\textsuperscript{57} Rather, ‘the absence of evidence of some recent use of the land or waters does not, of itself, require the conclusion that there can be no relevant connection’.\textsuperscript{58} However, it is lamentable that the High Court did not provide guidance on what the nature of the connection must be, and particularly, when a spiritual connection might suffice.

Regarding ‘recognition’ under s223(1)(c), the majority note that ‘the case law does not purport to provide a comprehensive understanding of what is involved in the notion of “recognition”’.\textsuperscript{59} Justice Kirby adds that ‘recognition’ of native title has no effect on the underlying Aboriginal law, as traditional laws and interests remain separate and continue to operate ‘regardless of any … extinguishment’.\textsuperscript{60}

Section 237A of the Act provides that native title, once extinguished, cannot be revived. However, s238 embodies a ‘non-extinguishment principle’, whereby native title rights and interests are suspended in certain circumstances for the duration of the inconsistent interest. Once the non-native interest ceases to operate, native title again has full effect.\textsuperscript{61}

Before the 1998 amendments, which were the legislative response to \textit{Wik},\textsuperscript{62} there was little in the \textit{Native Title Act} regarding extinguishment.\textsuperscript{63} Divisions 2A and 2B were part of the 1998 amendments, and now govern extinguishment of native title in conjunction with Div 2. Firstly, however, an act must be valid before the extinguishment question arises. Divisions 2A and 2B then confirm past extinguishment of native title by certain valid or validated acts attributable to the Commonwealth.\textsuperscript{64}

The timing of an act is important for validation purposes. Acts predating the \textit{Racial Discrimination Act} are valid,\textsuperscript{65} as are acts validated by the operation of the \textit{Native Title Act}. The past acts referred to in Division 2B are generally those done between the time of commencement of the \textit{Racial Discrimination Act} and 31 December 1993 and are validated by Division 2. A previous act under Division 2B will be valid if it is not invalidated by the operation of the \textit{Racial Discrimination Act}, or if the \textit{Native Title Act} renders the act valid.\textsuperscript{66} Intermediate acts fall between 1 January 1994 and 23 December 1996 and are validated by Div 2A.\textsuperscript{67} Certain past and intermediate period acts that would otherwise be invalid due to the operation of the \textit{Racial Discrimination Act} are also expressly validated by Div 2B.

\begin{itemize}
\item \textsuperscript{57} Id at [32].
\item \textsuperscript{58} Id at [64], [465].
\item \textsuperscript{59} Id at [20].
\item \textsuperscript{60} Id at [568] (Kirby J).
\item \textsuperscript{61} Id at [7] (Gleeson CJ, Gaudron, Gummow & Hayne JJ).
\item \textsuperscript{62} Id at [24].
\item \textsuperscript{63} Id at [26].
\item \textsuperscript{64} Id at [8], [10]. Acts attributable to the States and Territories are picked up under the various State Validation Acts that operate in conjunction with the \textit{Native Title Act}.
\item \textsuperscript{65} The \textit{Racial Discrimination Act} commenced on 31 October 1975.
\item \textsuperscript{66} \textit{Ward}, above n1 at [8].
\item \textsuperscript{67} See id at [5–6], [11].
\end{itemize}
Sections 23C and 23G of Division 2B mandate full and partial extinguishment respectively. Section 23C deals with extinguishment by 'previous exclusive possession acts', as defined in s23B, which completely extinguish native title. Section 23G deals with extinguishment by 'previous non-exclusive possession acts', as defined in s23F, which extinguish native title to the extent of any inconsistency.68

(iii) The Test for Extinguishment in the Event of Inconsistency

The regime established by the Native Title Act distinguishes between previous exclusive and previous non-exclusive possession acts. As seen above, previous exclusive possession acts will completely extinguish native title, and must be determined on a case-by-case basis.69 Where there has been a previous non-exclusive possession act, native title will be partially extinguished to the extent of the inconsistency.

The High Court held that the appropriate test to determine the extent of extinguishment is the ‘inconsistency of incidents’ test, whereby the legal nature and extent of the rights of the non-indigenous interest are identified and objectively compared against an established native title claim.70 The High Court preferred this test because of its objectivity, whereas examining acts for a ‘clear and plain intention’ necessarily involves looking at the subjective intention of the legislature. In agreeing with the majority on this issue, Justice Kirby added that it would be useful to supplement this test with the ‘reasonable user’ test propounded by Justices Beaumont and von Doussa in the Full Federal Court.71

‘Operational inconsistency’ was ruled out as a basis for extinguishment, but is nevertheless useful as an analogy.72 It is the legal rights of both parties that must be compared, not the practical effect of their actions. For example, when examining possible inconsistencies with the Ord River Irrigation Project, the court was not to view the Project as a holistic economic or geographic entity, but as a creature of statute. The legal effect of each statutory grant that established the Project, rather than the effects of its daily operations, needed to be compared with native title rights.73

The High Court affirmed the rejection by the Full Federal Court of the ‘adverse dominion test’.74 This was the test adopted by the primary judge. The test’s three elements required:

68 Id at [9].
69 Id at [588] (Kirby J).
70 Id at [78], [149] (Gleeson CJ, Gaudron, Gummow & Hayne JJ).
71 Id at [590] (Kirby J). See Western Australia v Ward above n15 at [329], [641] (Beaumont & von Doussa JJ).
72 Ward, above n1 at [149] (Gleeson CJ, Gaudron, Gummow & Hayne JJ).
73 Id at [147], [151]. This comparison was not undertaken by the High Court due to the lack of evidence on this issue before the Full Federal Court: at [156].
74 This test was proposed, but not adopted, in the dissenting judgment of Lambert JA in the Canadian case of Delgamuukw v British Columbia (1993) 104 DLR (4th) 470 at 669–672.
First, that there be a clear and plain expression of intention by parliament to bring about extinguishment in that manner; secondly, that there be an act authorised by the legislation which demonstrates the exercise of permanent adverse dominion as contemplated by the legislation; and thirdly, unless the legislation provides the extinguishment arises on the creation of the tenure inconsistent with an aboriginal right, there must be actual use made of the land by the holder of the tenure which is permanently inconsistent with the continued existence of aboriginal title or right and not merely a temporary suspension thereof.75

The High Court reasoned that the first step of the adverse dominion test, namely identification of a ‘clear and plain expression of intention by parliament to extinguish’ is again ‘apt to mislead’ by creating irrelevant inquiries into subjective intentions.76 An objective comparison between indigenous and non-indigenous legal rights is preferred.

(c) Practical Application of Extinguishment

(i) Pastoral Leases

Wik determined that native title could coexist with pastoral leases. Further questions of extinguishment, such as whether it is whole or partial, were addressed in the present case. If the pastoral leases in this case conferred a right of exclusive possession, they would extinguish native title completely under s23C of Division 2B. Under the ‘inconsistency of incidents’ test, the court must first look at what rights are granted under the pastoral lease, and then classify the grant.77 The High Court concluded that the rights under a pastoral lease are limited due to the extensive conditions and reservations imposed.78 Reaffirming Wik, the majority held that even with all its common law connotations, the nomenclature of a ‘lease’ does not of itself grant exclusive possession.79

Without the right of exclusive possession, the Western Australian pastoral leases fell within the s248B definition of a ‘non-exclusive pastoral lease’.80 As any lease is valid if granted before the commencement of the Racial Discrimination Act in 1975, the WA pastoral leases in this instance were ‘previous non-exclusive possession acts’ as defined in s23F of the Native Title Act. Native title was, therefore, open to partial extinguishment, under s23G, if inconsistent with the rights granted under the lease.81 Where the ‘inconsistency of incidents’ test found

75 Ward v State of Western Australia, above n11 at 508 (Lee J) (original emphasis).
76 Ward, above n1 at [78] (Gleeson CJ, Gaudron, Gummow & Hayne JJ).
77 Id at [186].
78 For example, the broadest condition reserved a right to any person ‘to enter, pass over, through, and out of any [unenclosed or enclosed but otherwise unimproved part of the land] while passing from one part of the country to another…on all necessary occasions’: Id at [178].
79 Ward, above n1 at [180] (Gleeson CJ, Gaudron, Gummow & Hayne JJ). Note, however, that McHugh J strongly dissented on this point: see [484]–[500].
80 Id at [188] (Gleeson CJ, Gaudron, Gummow & Hayne JJ).
81 This is under s23G of the Native Title Act for Commonwealth acts, and under its parallel (s12M) in the WA Titles (Validation) and Native Title (Effect of Past Acts) Act 1995 for previous non-exclusive possession acts attributable to the State.
that the continuation of native title was not inconsistent with the rights under the pastoral lease, the latter were held to prevail over the native title rights and interests but not to extinguish them.\footnote{\textit{Ward, above n1} at [190] (Gleeson CJ, Gaudron, Gummow & Hayne JJ).}

The main native title right held by the court to be inconsistent with the pastoral leases was the right to control access to the land, which is also reflected in the right to ‘speak for country’:

\begin{quote}
The pastoral leases … denied to the native title holders the continuation of a traditional right to say who could or who could not come onto the land in question.\footnote{\textit{Id} at [192].} The ability of the statutory scheme to recognise partial extinguishment led to the finding that the \textit{particular} Aboriginal right to control access to the land was extinguished, but that other rights and interests not inconsistent remained, and ‘probably continue unaffected’.\footnote{\textit{Id} at [193]–[194]. For the same reasons, the High Court held that excepting the right to control access, the leases in the Northern Territory were also not inconsistent and aspects of native title remained: at [415]–[417].} Specific determinations of extinguishment were remitted back to the Federal Court for reconsideration following further fact-finding.
\end{quote}

\textbf{(ii) Mining Leases}

Whether native title is extinguished either wholly or partially under mining leases is determined using a similar analysis to that discussed above in relation to pastoral leases. Again, the statutory framework is complicated, and careful attention to the time when the non-indigenous interest was granted must be appreciated. As native title finds its source in Aboriginal law and not common law, it cannot benefit from any protection granted by the common law in relation to priority disputes. Just as the statutory interests of pastoralists prevail over native title rights, the rights conferred under mining leases also prevail over native title rights and interests where the competing interests come into conflict.

Mining leases are statutory grants, with their meaning found in the statutes creating them. Rights granted under a mining lease include a right to exclude others but for mining purposes only.\footnote{\textit{Id} at [308].} In this case, the Full Federal Court had held that the Argyle Diamond mining lease granted exclusive possession to the company, and thus warranted the complete extinguishment of native title.\footnote{\textit{Id} at [333].} The large scale and permanence of the venture swayed the Full Federal Court towards this finding. The High Court overturned this view, determining that exclusive possession existed \textit{only} for mining purposes. Consequently, the court found that ‘it does not follow that all native title rights and interests have been extinguished.’\footnote{\textit{Id} at [308] [emphasis added].} For example, native title holders were still allowed access to and from the land.
There was too little evidence before the High Court for the judges to determine whether the native title rights asserted under the mining leases were extinguished wholly, or whether only particular rights and interests were extinguished. Again, the one exception was that the native title right to control access to the land was held to be inconsistent with mining access rights, and was certainly extinguished. As all the mining areas were previously held under pastoral leases, the right to control access was extinguished when the first pastoral grant was effected. Native title remains extinguished under the mining lease due to the operation of the non-revival principle.  

(d) Ownership and Control of Resources

(i) Minerals and Petroleum

Although the primary judge found that the native title holders in the claim area held a concurrent right to resources, including minerals and petroleum, the Full Federal Court held that any indigenous right to these resources was necessarily extinguished when the Crown appropriated ‘full and beneficial ownership’ of the resources to itself. The only surviving right found by the Full Federal Court was the right to ochre, as the traditional use of ochre had been established on the evidence.

The majority of the High Court held that, on the evidence before it, there was:

no evidence of any traditional Aboriginal law, custom or use relating to petroleum … [or] to any of the substances dealt with in the [relevant] Mining Acts. In these circumstances, no question of extinguishment arises. No relevant native title right or interest was established.

Even if a native title right to minerals or petroleum existed, the Court held that it would necessarily be extinguished because of an inconsistency with the legislation vesting all interests in minerals and petroleum in the Crown.

Justice Kirby disagreed with the majority on this issue, although he did agree that the legislation extinguished the right to resources in this case. According to his view, the common law has the capacity to recognise and reflect changes and developments in traditional Aboriginal laws and customs. It should not propound an approach that is frozen in time:

It would be a mistake to ignore the possibility of new aspects of traditional rights and interests developing as part of Aboriginal customs not envisaged, or even imagined, in the times preceding settlement.

88 Id at [309].
89 Ward v State of Western Australia, above n11 at 639 (Lee J).
90 Western Australia v Ward, above n15 at 290 (Beaumont & van Doussa JJ).
91 Id at 291-292.
92 Ward, above n1 at [382] (Gleeson CJ, Gaudron, Gummow & Hayne JJ).
93 Id at [383].
94 There would potentially be a very different outcome in another case if the mining legislation was different: Ward, above n1 at [575] (Kirby J).
95 Id at [574].
By not recognising such rights and interests in minerals, the majority of the Court has effectively held that those of Anglo-Australian descent can innovate and mine with modern methods as the level of technology and knowledge develops along with society’s corresponding needs for such resources, but that Aboriginal peoples must retain the traditions they held and a level of knowledge prevailing in pre-settlement times — that is, two centuries ago.

However, it can be argued that this approach conflicts with the principle underlying the majority finding, namely that continued physical occupation of the land is not a necessary requirement for a determination of native title.96 This principle is a clear demonstration that the Court does recognise that the law has and can move on to accommodate both modern conditions and the distinctiveness of indigenous Australian history. In allowing itself to be guided by an awareness of the changing circumstances of Aboriginal communities, the Court recognised that the dispossession, dwindling population numbers and discrimination faced by the indigenous Australians after 1788 may have made practical physical occupation impossible. Yet this lack of possession did not necessarily destroy their connection to the land.97 The Court could extrapolate this reasoning to allow the law to progress and consequently afford indigenous communities a native title right to resources, whether or not indigenous people used the resources pre-settlement.

(ii) Fishing Rights

The High Court affirmed that the public rights of navigation and fishing over the inter-tidal waters within the determination area demonstrated a fundamental inconsistency with the native title right to exclusively fish in those waters.98 The public interest falls under the wide definition of ‘any other right … in connection with … the land or waters’ in s253 of the Native Title Act, determined as required under s225(c) that all interests be considered.99

The denial of rights and interests to minerals and petroleum, and the denial of the exclusive right to fish in inter-tidal waters are examples of how the law is required to bend in order to apportion interests in a limited resource. The problem is not a new one, but it is amplified in cases involving aboriginal interests, as indigenous claimants are faced with the disadvantage of being compelled to argue for their rights within a legal system that does not reflect their traditional laws and customs. It would be interesting to see how Anglo-Australian land law would fare if the situation were reversed, and mining companies had to make their claim for minerals before a meeting of Aboriginal elders.

96 See Ward, above n1 at [63]–[64] (Gleeson CJ, Gaudron, Gummow & Hayne JJ).
97 See Boge, above n48 at 24 where he notes that ‘a spiritual connection, and a performance of responsibility for the land could be maintained even where physical presence had ceased, either because the indigenous people had been hunted off the land, or because their numbers had become so thinned that it was impracticable to visit the area.’
98 Ward, above n1 at [388] (Gleeson CJ, Gaudron, Gummow & Hayne JJ).
99 Id at [387].
(e) Compensation and s10 of the Racial Discrimination Act

In *Mabo*, the High Court stated that ‘native title may be protected by such legal or equitable remedies as are appropriate to the particular rights and interests established by the evidence’.  

The Court, in the present case, commented that principles of what may constitute an ‘appropriate remedy’ are yet to be developed.  

This includes the notion of compensation.

In the absence of a common law basis for compensation due to the extinguishment of native title, claims for compensation are determined by considering s23J of Division 2B of the *Native Title Act*, sections 17 and 20 of Division 2, and the interaction of the *Native Title Act* with s10(1) of the *Racial Discrimination Act*.  

The engagement of the *Racial Discrimination Act* is possible because:

Native title characteristically is held by members of a particular race so that interference with the enjoyment of native title is capable of amounting to discrimination on the basis of race, colour, or national or ethnic origin.

Section 10(1) of the *Racial Discrimination Act* is directed to the discriminatory operation of laws. It provides:

If, by reason of, or of a provision of, a law of the Commonwealth or of a State or Territory, persons of a particular race, colour or national or ethnic origin do not enjoy a right that is enjoyed by persons of another race, colour or origin, or enjoy a right to a more limited extent … persons of the first-mentioned race, colour or national or ethnic origin shall, by force of this section, enjoy that right to the same extent as persons of that other race, colour or national or ethnic origin.

The possible engagement of s10 is examined on a case-by-case basis because it depends on the effects of the particular legislation. As it is, the practical effect of the law is examined for discrimination, even statutes not specifically aimed at native title can fall within its scope if they are shown to have a discriminatory effect.  

Section 10 may be triggered either where the law ‘omits to make the enjoyment of the right universal’, or where it ‘imposes a discriminatory burden or prohibition’.  

---

100 *Mabo*, above n3 at 70.


102 Additionally, s7 of the *Native Title Act* says that the ‘*Native Title Act* is to be read and construed subject to the provisions of the *Racial Discrimination Act*’, so that any validation effected by the *Native Title Act* displaces invalidity resulting from operation of the *Racial Discrimination Act*: *Ward* at [99].

103 *Ward*, above n1 at [117]; *Mabo (No 1)* (1988) 166 CLR 186 at 218 (Brennan, Toohey, & Gaudron JJ), and at 230 (Deane J); *Native Title Act Case*, above n45 at 437 (Mason CJ, Brennan, Deane, Toohey, Gaudron & McHugh JJ).

104 *Ward*, above n1 at [126].

105 Id at [115].

Regarding compensation, s10(1) could be activated if there is extinguishment of native title without compensation (or if a lesser amount is offered) in an instance where extinguishment of a non-indigenous interest would be compensated.\textsuperscript{107} If compensation were not offered to \textit{any} holder for an interference with their interest in land, then a failure to compensate for interference with native title would not amount to discrimination.\textsuperscript{108} In the present case, the High Court found that, having been asked to rule upon principles of extinguishment, issues regarding compensation under the \textit{Native Title Act} did not directly arise.\textsuperscript{109} A definitive statement on the issue of compensation was again delayed.

### 3.2 The Minority Judgments: Justices McHugh and Callinan

Justice Callinan provided a long dissenting judgment in this case. Justice McHugh concurred with Justice Callinan, with one exception.\textsuperscript{110} Both judges held that pastoral and mining leases do confer exclusive possession, and so native title was fully, rather than partially, extinguished over the area.\textsuperscript{111} However, Justice Callinan does provide for the possibility of partial extinguishment in other circumstances.\textsuperscript{112} Despite the majority determining that the nomenclature of an interest as a ‘lease’ is not indicative of the intention to grant the lessee exclusive possession, Justice McHugh discusses the meaning of possession under the expressions ‘lease’ and ‘demise’ in some depth, arriving at the contrary conclusion.\textsuperscript{113}

Both dissenting judges comment that the \textit{Wik} judgment should be treated with caution. Even though \textit{Wik} must be followed as precedent, the judges argue that less significance should be attributed to it. Justice Callinan states that this is because there is no \textit{ratio decidendi} in that case;\textsuperscript{114} Justice McHugh adds that even if this analysis is incorrect, \textit{Wik}’s value as a useful precedent is necessarily limited as it deals with legislation arising in a different jurisdiction.\textsuperscript{115}

Notably, the dissenting view that native title was fully extinguished under pastoral and mining leases led both judges to call for significant reforms in native title law. Justice McHugh stated that:

\begin{quote}
The dispossession of the Aboriginal peoples from their lands was a great wrong ... Redress cannot be achieved by a system that depends on evaluating the competing legal rights of landholders and native title holders. The deck is stacked against the native title holders whose fragile rights must give way to the superior
\end{quote}

\textsuperscript{107} \textit{Ward}, above n1 at [124].
\textsuperscript{108} Id at [126].
\textsuperscript{109} Id at [12].
\textsuperscript{110} \textit{Ward}, above n1 at [472] (McHugh J). McHugh J disagrees with Callinan J’s finding (\textit{Ward}, above n1 at [854]) that native title holders fall within the definition of ‘occupier’ for the purposes of the \textit{Mining Act 1978} (WA): at [559].
\textsuperscript{111} \textit{Ward}, above n1 at [699]-[701] (Callinan J).
\textsuperscript{112} Id at [616]-[618].
\textsuperscript{113} Id at [482]-[500], and again relating to pastoral and mining leases at [549]-[558].
\textsuperscript{114} Id at [696].
\textsuperscript{115} Id at [480].
rights of the landholders whenever the two classes of rights conflict. And it is a
system that is costly and time-consuming. … A better system may be an arbitral
system that declares what the rights of the parties \textit{ought to be} according to the
justice and the circumstances of the individual case.\footnote{116}

Similarly, Justice Callinan called for a ‘true and unqualified settlement of lands or
money’ in order to redress the wrongs of dispossession, instead of the present
attempts to battle through the ‘labyrinth of Minos’\footnote{117} that he sees the current body
of law as reflecting. The only resolution of this labyrinth in its current state,
according to his view, is the ‘futile or unsatisfactory … attempt to fold native title
rights into the common law.’\footnote{118}

### 3.3 The Orders

Both appeals were allowed in this case, although neither side could be said to ‘win
or lose’.\footnote{119} The High Court was unable to make a final determination on the
outcome given the present state of fact-finding. The Full Federal Court did not
apply the correct law, and in particular did not take Div2B of Pt 2 of the Native
Title Act into account.\footnote{120} Thus, the matter was remitted to the intermediate
appellate court to complete the hearing, taking into consideration the remaining
questions of extinguishment arising from the High Court decision.\footnote{121}

### 1. The Irrelevancy of International Law?

International law has the potential to affect native title claims in two ways: firstly
through international treaties and covenants that Australia has ratified;\footnote{122} and
secondly, through the persuasive influence of comparable cases in similar
jurisdictions. Given both the proven ability of these avenues to influence domestic
law, and the submissions by the Human Rights and Equal Opportunities
Commission as an intervener in this case, the scarcity of reference to international
law is surprising. Only two judges make explicit reference to it at all. Justice Kirby
is in favour of using international law to expand upon the principles governing the
recognition and extinguishment of native title by using international human rights
law.\footnote{123} Justice Callinan, on the other hand, makes a forceful statement that
international precedents should not be used in an Australian context.\footnote{124} In

\footnote{116} Id at [561] (emphasis added).
\footnote{117} \textit{Ward}, above n1 at [969] (Callinan J).
\footnote{118} Id at [970].
\footnote{119} \textit{Ward}, above n1 at [471] (Gleeson CJ, Gaudron, Gummow & Hayne JJ).
\footnote{120} The Federal Court is bound (on appeal) to ‘decide the rights of the parties upon the facts and in
accordance with the law as it exists at the time of hearing the appeal’. Id at [70].
\footnote{121} Id at [72], [467].
\footnote{122} Although treaties have no effect in domestic law without implementing legislation, the High
Court held that there was a ‘legitimate expectation’ that administrative decisions would be made
in accordance with ratified, but as yet un-implemented treaties: \textit{Minister for Immigration and
\footnote{123} See \textit{Ward}, above n1 at [566]–[567] (Kirby J).
\footnote{124} See Part IV of Callinan J’s judgment at [954]–[963].
response to a submission by HREOC that ‘lawfulness is supplied not only by national law, but also by international law’, Callinan J states that:

The provisions of the Constitution are not to be read in conformity with international law. It is an anachronistic error to believe that the Constitution, which was drafted and adopted by the people of the colonies well before international bodies such as the United Nations came into existence, should be regarded as speaking to the international community. The Constitution is our fundamental law, not a collection of principles amounting to rights of man, to be read and approved by people and institutions elsewhere. The approbation of nations does not give our Constitution any force, nor does its absence deny it effect. Such a consideration should, therefore, have no part to play in interpreting our basic law.

However, foreign developments in aboriginal law are frequently used for comparative analysis with Australia. Due to the legal and historical similarities, comparisons with Canada are often utilised. One area in which Canadian jurisprudence offers guidance is in the evolution of native title rights, which could influence the recognition of a native title right to minerals and petroleum in Australia. The recent case of *Delgamuukw v British Columbia* held that imposing limits that effectively force aboriginal rights holders to engage only in traditional activities upon their land would ‘amount to a legal straitjacket on aboriginal peoples who have a legitimate claim to the land’. However, the manner of usage cannot be ‘irreconcilable with the nature of the attachment to the land which forms the basis of the group’s claim to aboriginal title’. For example, as aboriginal title stems from a connection with the land, strip mining that causes massive devastation to the land would not be allowed. Canadian precedents have also demonstrated a willingness to allow indigenous communities to ‘move with the times’, for example by allowing fishing using modern methods.

A second area where Canadian jurisprudence could assist native title claimants is in the area of priority disputes. Lambert JA in *Delgamuukw v British Columbia* questioned why aboriginal rights necessarily give way first when there is an inconsistency with a non-indigenous grant. It is possible for this approach to operate effectively under the ‘adverse dominion’ test proposed by Lambert JA in that case. However, given the High Court’s rejection of that test and their affirmation of the ‘inconsistency of incidents’ test, it is unlikely that this issue will be successfully raised in the near future.

---

125 Id at [960].
126 Id at [961].
127 *Delgamuukw v British Columbia* [1997] 3 SCR 1010 at 1091 (Lamer CJC).
128 Id at 1088.
129 See for example *R v Sparrow* [1990] 1 SCR 1075 at 1093.
2. Conclusions and Implications

Native title is governed by the *Native Title Act* and related legislation. The High Court, by majority, held that partial extinguishment and suspension of native title rights is possible under this statutory framework. This view of native title seemingly supports a ‘bundle of rights’ characterisation, which may lead to additional gradual erosion of native title rights and interests over time. Yet this is the very difficulty that the *Native Title Act* was supposed to guard against.

Further, the Court confirmed that native title coexists to the extent that it is not inconsistent with other interests over the same land or waters. Where there is conflict, the non-indigenous rights prevail. Where there is inconsistency between the interest and a question of partial extinguishment arises, the correct test to use is the ‘inconsistency of incidents’ test, and not the ‘operational inconsistency’ nor ‘adverse dominion’ tests.

By majority, the Court held that indigenous Australians do not have native title rights to minerals or cultural property. This approach has frozen Aboriginal communities in the pre-settlement era, denying them the ability to redefine their relationship with the land as challenges and changes arise. To rethink this view in future cases, the courts need to consider, among other things, how Aboriginal laws and customs traditionally dealt with development and modernisation. However, there are substantial institutional and evidentiary impediments to taking such an approach.

If nothing else, wading through the immense complexity of the statutory scheme now governing native title claims highlights the pressing need for a simplified yet comprehensive land claims process. When both the difficulties in giving evidence and the procedural requirements under the present system are added to this complexity, it can be seen that native title disputes demand a lot of patience and capability, not only on behalf of the parties, but of the courts too. A thorough overhaul of the system is needed. Possible approaches could be in the form of Justice McHugh’s arbitral system, Justice Callinan’s definitive settlement of lands or money, or in a combination of proposals. Determining which one of these approaches is best will be difficult, resource intensive and will take years. In the meantime, courts and parties have no choice but to continue to limp clumsily through the ‘labyrinth of Minos’. It will be a long time before the native title dispute resolution system is overhauled, unless a political priority is given to this issue.