NOVEL SCIENTIFIC EVIDENCE AND JUDICIAL GATEKEEPING: *R v CALDER* AND *DAUBERT v MERRELL DOW PHARMACEUTICALS* COMPARED

Karen Belt*

This article examines the approach of the High Court to the admissibility of novel scientific expert evidence in R v Calder. *In* Calder, *Tipping J establishes a "gatekeeping" role for judges which requires them to test novel scientific evidence for relevance and reliability. The article compares that approach with the approach taken by the United States Supreme Court in* Daubert v Merrell Dow Pharmaceuticals. *The implications of such a test are considered. Although the Court of Appeal has not considered the issues raised in* Calder, *the article concludes that the approach is the most suitable one for New Zealand.*

I THE APPROACH IN CALDER

A Facts and Previous Case Law

In the case of *R v Calder*,[1] Dr Vicky Calder was charged with attempting to murder her ex-partner, Professor David Lloyd, by poisoning him with acrylamide. The Court was asked by the defence, in a pre-trial application, to rule on the admissibility of the results of a scientific technique which analysed Professor Lloyd's hair for traces of CEC, which is a by-product of acrylamide. The defence objected to the admission of the hair analysis, claiming that the analysis amounted to a scientific experiment and should be excluded as unreliable novel scientific evidence. The defence submitted that the "area of

* This paper was submitted as part of the LLB (Hons) programme at Victoria University of Wellington, 1997.

1 Unreported, 12 April 1995, High Court, Christchurch Registry, T 154/94.

expertise" rule applied to the evidence and that the evidence should be excluded on the basis that its prejudicial effect exceeded its probative value.

Prior to the 1980s, expert evidence in New Zealand (as in Australia and England) was generally allowed to go before the trier of fact as long as it was relevant, offered by a qualified expert and did not offend one of the traditionally recognised rules of evidence.[2] A rule excluding novel scientific evidence was not recognised. However, since the 1980s, a number of admissibility decisions, in both New Zealand and Australia, have addressed the potentially prejudicial nature of expert scientific evidence.[3]

Freckelton states that these admissibility decisions show the development of a rule which would exclude scientific evidence which was not based on a recognised "area of expertise".[4] Scientific theories and techniques which had not achieved sufficient acceptance within the relevant professional community have been excluded in these decisions. In New Zealand, a version of the rule has been applied to novel psychological evidence. Experts basing their opinions on such evidence have been required to demonstrate that the subject matter of their opinion is "a sufficiently recognised branch of science".[5]

In reaching his decision, however, Tipping J stated that there was no direct authority in New Zealand dealing with the admissibility of novel scientific evidence.[6] He had been referred to the cases[7] which dealt with the admissibility of novel scientific evidence offered by psychologists and psychiatrists, but did not consider them to be of any assistance due to their different subject matter. He therefore considered academic writing and the approach of other common law jurisdictions to determine what principles should apply when scientific evidence is challenged.

Tipping J found the approach of the North American courts useful and discussed the United States Supreme Court case of *Daubert v Merrell Dow Pharmaceuticals*[8] at some length.[9] In *Daubert* the Supreme Court ruled that the *Frye*[10] test, which required a novel

---

2 I Freckelton and H Selby *Expert Evidence* (Loose-leaf ed, Law Book Company, Australia, 1993) 1-2221.

3 See, for example, *R v B* [1987] 1 NZLR 362.

4 The "area of expertise" rule is discussed in Freckelton and Selby above n 2, 1-2221.

5 Above n 3, 367.

6 Above n 1, 5.

7 *R v B*, above n 3; *R v Accused* [1989] 1 NZLR 714; *R v CS* (1993) 11 CRNZ 45; *R v R* (1994) CRNZ 402.

8 (1993) 125 L Ed 2d 469.

9 Above n 1, 4.

scientific technique or theory to have gained general acceptance in the field to which it belonged, had been replaced by a new test which required the judge to play a "gatekeeping" role with regard to expert scientific testimony. Judges should no longer rely on general acceptance to determine reliability but must assess the reliability of expert scientific testimony for themselves. Criteria for assessing the reliability of scientific testimony were discussed in *Daubert*.

The Judge also discussed the Victorian case of *Lucas*.[11] This decision emphasised the need for caution in admitting scientific evidence, as it may appear to the jury to have an overwhelming appearance of validity. However, Tipping J focused on a quote in the judgment from *US v Baller*,[12] which suggested that the danger of a scientific opinion having undue weight with the jury could be countered by requiring a demonstrable, objective procedure for reaching the opinion and qualified witnesses who could duplicate the result or criticise the means by which it was reached. Scientific evidence need only be excluded where there was an exaggerated popular opinion of the accuracy of a particular technique. Otherwise, relevant scientific evidence should be admitted in the same manner as other expert testimony and its weight tested by cross-examination and refutation.

Finally, the Judge considered the Law Commission's discussion paper on opinion and expert evidence.[13] This paper had discussed the need to assess scientific evidence for scientific reliability, including the validity of the underlying theory and the reliability of the procedures and techniques used in the particular case.[14] The Law Commission concluded that "a theory need not be accepted by all or most scientists working in the relevant area".[15] While idiosyncratic and unsatisfactory theories must be guarded against, it considered that theories which were newly developed or which represented the views of a minority might still be reliable and helpful.[16]

---

10 *Frye v United States* (DC Cir. 1923) 293 F 1013.

11 [1992] 2 VR 109.

12 (1975) 519 Fed 2d 463.

13 New Zealand Law Commission *Evidence Law: Expert Evidence and Opinion Evidence - Preliminary Paper No 18* (Law Commission, Wellington, 1991).

14 Above n 1, 6.

15 As cited in *Calder*, above n 1, 6.

16 Above n 1, 6.

B The Calder *Test*

Drawing on the case law and reform options, Tipping J developed a test for the admissibility of novel scientific evidence.[17] First, the proponent of such evidence must be a suitably qualified person. Second, the evidence must be shown to be both relevant and helpful. To be relevant, the evidence must logically tend to show that a fact in issue is more or less likely. To be helpful, the evidence must pass a threshold test which Tipping J called the "minimum threshold of reliability".[18] This means that the party offering the expert evidence (the "proponent") must demonstrate that it has a sufficient claim to reliability to be admitted. If the judge is satisfied that the evidence is sufficiently reliable, the evidence may be admitted at trial. Its probative value can then be tested by cross-examination and counter evidence and is a matter for the jury to determine. Under this approach, evidence which has been shown to be relevant and sufficiently reliable may still, however, be excluded if its prejudicial effect outweighs its probative value.[19]

The proponent of the evidence must establish expertise, relevancy and helpfulness, and also demonstrate that the evidence is more probative than prejudicial. The Judge did not explicitly refer to the standard of proof, but said that any real doubt should be resolved in favour of exclusion. At first sight, this seems to suggest that the standard is beyond reasonable doubt. However, the standard of proof with regard to preliminary facts is usually the balance of probabilities.[20] It is therefore more likely that the judge is merely required to be satisfied that the evidence is relevant and helpful. This does not require proof beyond a reasonable doubt, simply that the judge makes up his or her mind on the issue.[21]

Tipping J anticipated criticism that a test requiring "a sufficient claim to reliability" may be too general to be of assistance.[22] However, he justified his test by pointing to its flexibility which would enable it to be applied to different types of novel scientific evidence. The test could also be supplemented by appropriate factors drawn from the case law. He suggested that it may be useful to consider the factors outlined in the

---

17 Above n 1, 7.

18 Above n 1, 7.

19 Above n 1, 7.

20 DL Mathieson (ed) *Cross on Evidence* (NZ Loose-leaf ed, Butterworths, Wellington, 1996) 119-1; *Police v Anderson* [1972] NZLR 233 (CA).

21 In *R v White (David)*, [1988] 1 NZLR 264, the Court of Appeal held that s 75(2) of the Criminal Justice Act, which required that a judge be satisfied that preventive detention is expedient for the protection of the public (a serious matter indeed), did not require proof beyond a reasonable doubt, simply that the judge makes up his or her mind on the issue.

22 Above n 1, 7.

Canadian cases of *Johnston*[23] and *Melaragni*[24] (both decisions of the Ontario Court, General Division) when deciding whether the threshold had been crossed.

II THE APPROACH IN DAUBERT

A The Daubert *Test*

The decision in *Calder* has much in common with the decision of the United States Supreme Court in *Daubert*. The *Daubert* decision involved the rejection of the general acceptance test which Tipping J was invited, but declined, to apply in *Calder*. A comparison between the two tests highlights both the strengths and weakness of the *Calder* decision.

The *Daubert* decision is based upon the Federal Rules of Evidence (FRE) Rule 702 which states:

> If scientific, technical, or other specialised knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise.

---

23 In *Johnston* (1992) 69 CCC (3d) 395, 415, these were "(1) The potential rate of error. (2) The existence and maintenance of standards. (3) The care with which the scientific technique has been employed and whether it is susceptible to abuse. (4) Whether there are analogous relationships with other types of scientific techniques that are routinely admitted into evidence. (5) The presence of fail safe characteristics. (6) The expert's qualifications and stature. (7) The existence of specialised literature. (8) The novelty of the technique in its relationship to more established areas of scientific analysis. (9) Whether the technique has been generally accepted by experts in the field. (10) The nature and breadth of the inference adduced. (11) The clarity with which the technique may be explained. (12) The extent to which basic data may be verified by the court and jury. (13) The availability of other experts to evaluate the technique. (14) The probative significance of the evidence".

24 In *Melaragni* (1992) 73 CCC (3d) 348, 353, these were "(1) Is the evidence likely to assist the jury in its fact finding mission, or is it likely to confuse and confound the jury? (2) Is the jury likely to be overwhelmed by the "mystic infallibility" of the evidence, or will the jury be able to keep an open mind and objectively assess the worth of the evidence? (3) Will the evidence, if accepted, conclusively prove an essential element of the crime which the defence is contesting, or is it simply a piece of evidence to be incorporated into a larger puzzle? (4) What degree of reliability has the proposed scientific technique or body of knowledge achieved? (5) Are there a sufficient number of experts available so that the defence can retain its own expert if desired? (6) Is the scientific technique or body of knowledge such that it can be independently tested by the defence? (7) Has the scientific technique destroyed the evidence upon which the conclusions have been based, or has the evidence been preserved for defence analysis if requested? (8) Are there clear policy or legal grounds which would render the evidence inadmissible despite its probative value? (9) Will the evidence cause undue delay or result in the needless presentation of cumulative evidence?".

The Supreme Court ruled that under the FRE a federal trial judge must ensure that all scientific evidence is both relevant and reliable, and that, with regard to scientific evidence, evidentiary reliability is based on scientific validity.[25]

The Supreme Court inferred the reliability requirement from the interpretation of the term "scientific … knowledge" in rule 702. The adjective "scientific" implies a grounding in the methods and procedure of science. Similarly, the word "knowledge" connotes more than subjective belief or unsupported speculation.[26] In order to determine whether the expert's testimony pertains to "scientific knowledge", the trial judge must assess "whether the reasoning or methodology underlying the testimony is scientifically valid".[27] The focus must be on the principles and methodology, not on the conclusions they generate.[28]

The Supreme Court also held that there should be a heightened relevance requirement for expert scientific evidence. This was implicit in the rule 702 requirement that the expert evidence assist the trier of fact to understand the evidence or to determine a fact in issue. In order for scientific evidence to be "helpful", the expert's opinion must relate to an issue that is actually in dispute and must provide a "valid scientific connection to the pertinent enquiry".[29] This requires an assessment of whether the "reasoning or methodology properly can be applied to the facts in issue".[30]

The Supreme Court offered some guidance to the lower courts to help them determine whether the reasoning or methodology underlying the testimony was scientifically valid, and whether it could properly be applied to the facts in issue. This guidance consisted of four criteria: whether the theory or technique could (and had) been tested; whether the theory or technique had been subjected to peer review and publication; the known or potential rate of error; and acceptance within the relevant scientific community. However, the Court emphasised that this was not a definitive checklist or test.[31]

Finally, the Supreme Court pointed out that other rules may be applicable even if the expert evidence is admissible under rule 702. In particular, rule 403, which permits the

---

25 Above n 8, 470.

26 Above n 8, 481.

27 Above n 8, 482.

28 Above n 8, 484.

29 Above n 8, 482.

30 Above n 8, 481.

31 Above n 8, 482.

exclusion of relevant evidence "if its probative value is substantially outweighed by the danger of unfair prejudice, confusion of the issues, or misleading the jury".[32]

B Daubert *Compared to* Calder

Although *Daubert* was based on the interpretation of a relevant statute, Tipping J was strongly influenced by the decision. He adopted the "gatekeeping" analogy of the Supreme Court[33] and the analysis which rejects the general acceptance test for court determined reliability.[34] Both courts created very general tests requiring relevance and reliability. Both offered non-definitive criteria to determine reliability. There are also, however, differences between the two cases.

1 "*Fit*"

Daubert sets a more stringent test for relevancy than that advocated by Tipping J. *Calder* requires only the usual "logical" relevancy. To pass the relevancy test, the evidence need only show that a fact in issue is more or less likely.[35] In contrast, *Daubert* subjects scientific expert evidence to a heightened relevancy requirement. The judge must inquire whether the reasoning or methodology employed by the expert can properly be applied to the facts in issue.[36] For example, studies that show that a particular chemical causes cancer in animals may not be sufficiently "relevant" to be admissible to prove that the same chemical causes cancer in humans.[37] There must be good grounds to extrapolate from animals to humans before the tests can be said to be relevant under *Daubert*.[38] The Supreme Court referred to this heightened relevancy requirement as "fit".[39]

It is possible to argue that a heightened relevance test is implicit in the *Calder* test. Expert evidence that is not closely related to the factors present in the case could be said to not be sufficiently reliable. However, given that the Judge had considered *Daubert* in some detail, it is likely that he would have expressly referred to "fit" if he had intended it to be part of his test. Thus, although *Daubert* is considered a liberal admissibility

---

32 Above n 8, 484.

33 Above n 1, 7.

34 Above n 1, 7.

35 Above n 1, 7.

36 Above n 8, 481.

37 *In Re Paoli Railroad Yard PCB Litigation* (CA3 1994) 35 F 3d 717, 743. A post *Daubert* decision.

38 Above n 37, 743.

39 Above n 8, 481.

standard, its heightened relevancy requirement makes it a stricter test than that set in *Calder*.

This difference in standards can possibly be traced to the different legal environments of New Zealand and the United States. Jury trials are much more common in the United States, where they are often used in civil cases as well as criminal cases. Professional experts are now an integral part of the United States litigation process. Many see such experts as "hired guns" who will testify to whatever their client wishes. In civil cases, juries have given large awards to "deserving" plaintiffs on the basis of "junk" scientific evidence let in under a low admissibility standard.[40] This has led to greater concern in the United States over the admission of unreliable scientific evidence.

In New Zealand, however, the small number of trials requiring scientific evidence will not support a "professional expert witness" class. Expert witnesses have other jobs and are therefore perceived as being more independent. Juries are not required to cope with scientific evidence in civil trials where they might be tempted to find for "deserving" plaintiffs on the basis of poor science. Consequently, expert scientific witnesses are viewed with less suspicion in New Zealand and are not seen as needing regulation to the same extent as in the United States.

The Law Commission has rejected suggestions that relevance should be accorded any meaning going beyond logical relevance.[41] It argues that a relevance inquiry going beyond logical relevancy is either an inquiry as to probative value or a device for excluding evidence which is unfairly prejudicial, misleading or time-wasting by comparison with its true worth.[42] Compressing these issues into one can result in conceptual confusion. It is better to require the court to address them separately. In as much as "fit" requires anything further than logical relevancy, it should not be incorporated into the test in *Calder*.

2 Testability and application to the "soft" sciences

In *Daubert*, the Supreme Court elevated one particular criteria for establishing reliability, "testability", which it described as "ordinarily a key question".[43] This highlights the Supreme Court's conception of science as an empirical endeavour in which testing a theory in order to prove it false plays a key role. The *Calder* test, in contrast,

40 For examples of such cases see P Huber *Galileo's Revenge: Junk Science in the Courtroom* (Basic Books, USA, 1991) 42.

41 Above n 13, 26.

42 Above n 13, 26.

43 Above n 8, 482.

does not concentrate on testability, but refers to a much broader range of criteria[44] which focus on policy issues as well as scientific validity. This suggests the *Calder* test should have a wider application than the *Daubert* test. Interestingly, the opposite is the case.

In the United States, there was confusion after *Daubert* as to whether the case applied to expert evidence from the social and behavioural sciences. In the past, the federal courts had generally distinguished between "hard" science, such as DNA analysis and "soft" expertise, such as social science and mental health testimony.[45] However, Justice Blackmun's definition of science was so broad that it arguably included social and mental health science. Consequently, lower federal courts have applied the case to determine the admissibility of child sexual abuse accommodation syndrome evidence,[46] rape trauma syndrome evidence,[47] and psychological testimony about the unreliability of eyewitness identifications.[48]

The *Daubert* focus on "testability" makes it an unsuitable test for expert evidence based on the "soft" sciences. Because these disciplines deal with human beings, they must employ methodologies which do not harm their subjects. This may make falsifying a hypothesis impossible. While scientists in these disciplines also experiment on animals, allowing more rigorous methodologies to be employed, these experiments are likely to fall foul of the *Daubert* "fit" requirement. Consequently, the application of *Daubert* has usually resulted in the exclusion of expert testimony based on the "soft" sciences.

The *Calder* test, in contrast, does not focus on whether a theory may be falsified. The judge is directed to apply any criteria from the case law which is thought appropriate. This flexibility could allow the development of suitable criteria to test the methodology of "soft" science evidence. However, because of Court of Appeal precedents on novel psychological evidence,[49] the *Calder* test is limited to "hard" science.

---

44 Via the references to *Johnston* and *Melaragni*.

45 E Imwinkelreid "Evidence Law Visits Jurassic Park: The Far Reaching Implications of the *Daubert* Court's Recognition of the Uncertainty of the Scientific Enterprise" (1995) 81 Iowa Law Review 55, 69.

46 *State v Foret* (La 1993) 628 So 2d 1116; *Gier v Educational Service Unit No 16* (D Neb 1994) 845 F Supp 1342.

47 *State v Alberico* (NM 1993)861 P 2d 192.

48 *US v Amador-Galvan* (9th Cir 1993) 9 F 3d 1414.

49 Above n 7.

3 Should the Calder test also apply to "soft" science?

A single admissibility test for all forms of scientific evidence is supported by the Law Commission.[50] It proposes a "helpfulness" test (almost identical to that in *Calder*) applicable to all forms of expert evidence.[51] It is submitted that a reliability test should apply to all forms of expert evidence purported to be based on scientific research. The dangers of putting unreliable evidence with an aura of scientific infallibility before a jury exist for both "hard" and "soft" science. However, the criteria for assessing reliability should differ according to the type of science involved.

A possible criterion for a judge to apply in deciding if a theory based on "soft" science research is sufficiently reliable is whether there are any avoidable methodological weaknesses in the study. For example, in research into the effects of child sexual abuse it is not possible to test a theory by abusing one sample of children and comparing the results with a non-abused control group.[52] However, it is possible to compare studies of children known to have been abused with a group of children whom the researcher is reasonably sure have not been abused. Further, the study may have focused on a particular group of children or a particular type of abuser. Thus, the evidence would not be reliable if used to comment on persons outside of those groups. The size of the group studied and its method of selection could also have a bearing on reliability. Peer review and general acceptance would also provide an indication of whether the research was sound.

Reliability should also depend on the purpose for which the expert evidence is being offered. It may be shown to be sufficiently reliable for one purpose but not another. For example, rape trauma syndrome was developed from a study of the behaviour of women known to have been raped. Thus, it is sufficiently reliable to be helpful when the expert witness is giving evidence about the ways a woman who has been raped might behave. However, it may be judged insufficiently reliable to support an expert opinion that a woman had been raped. Such an assertion would have to be based on research that had eliminated other possible causes of the symptoms observed in the women who were the subject of the study.

Such admissibility decisions would depend on the judge weighing up the probative/prejudicial effects of the testimony. Some commentators argue that juries are

---

50 Above n 13, 21.

51 Above n 13, 21.

52 LR Askowitz and MH Graham "The Reliability of Expert Psychological Testimony in Child Sexual Abuse Prosecutions" (1994) 15 Cardozo Law Review 2027, 2040.

better able to evaluate "soft" science evidence as opposed to "hard" science evidence.[53] Consequently, there is less need to shield the jury from such evidence.

53 D McCord "Syndromes, Profiles and Other Mental Exotica: A New Approach to the Admissibility of Non-Traditional Psychological Evidence in Criminal Cases" (1987) 66 Oregon Law Review 19, 183.

4 Should all scientific evidence be screened?

Another possible difference between *Calder* and *Daubert* is that the *Daubert* test applies to all scientific evidence not just novel scientific evidence.[54] In *Calder*, the objection to the evidence was couched in terms of it being novel scientific evidence.[55] Tipping J stated that the test applied to "scientific evidence, such as that in issue in this case".[56] This could be read broadly or narrowly.

The *Calder* test should apply to all scientific evidence. The general acceptance test, by definition, only applied to novel scientific evidence. However, there is no logical reason for limiting a judicial assessment of reliability to evidence based on novel science. Scientific techniques which had been previously accepted in court and were no longer novel scientific evidence have been ruled inadmissible under *Daubert* because they have subsequently been demonstrated to be unreliable.[57] This result should also occur under *Calder*.

III CONSIDERATION OF THE CALDER *APPROACH*

A The General Acceptance and Judicial Assessment Tests Compared

In the past, New Zealand courts were content to admit relevant scientific evidence as long as the expert was qualified. In cases where concern has been voiced regarding novel scientific (psychological) evidence, judges have relied on a version of the general acceptance test to decide admissibility.[58] The decision in *Calder*, which marks a move toward judicial assessment, makes it timely to compare the two tests.

The main benefit of the general acceptance test is that it provides a relatively quick method of determining the admissibility of novel scientific evidence. Being a conservative standard, it is said to provide greater assurance of reliability than other tests.[59] It also ensures that there will be a pool of experts who could be called upon to

---

54 However, the court noted that once a scientific theory became firmly established and developed into what the court termed "scientific law" it would be subject to judicial notice and would not be subject to an admissibility test. See *Daubert*, above n 8, n 11.

55 Above n 1, 3.

56 Above n 1, 7.

57 *United States v Posado* (5th Cir. 1995) 57 F 3d 428.

58 The subject matter of the expert opinion was required to be a sufficiently recognised branch of science at the time the evidence was given. *R v B*, above n 3, 367.

59 PC Gianelli "The Admissibility of Novel Scientific Evidence: *Frye v United States*, a Half Century Later" (1980) 80 Columbia Law Review 1197, 1207.

testify with regard to the scientific evidence.[60] Supporters claim this test promotes consistency of decision making.[61]

However, the general acceptance test may not be easy to apply. There may not be a consensus in the relevant scientific community concerning a particular issue. There may also be difficulties in even determining who constitutes the relevant scientific community.[62] In applying the *Frye* rule in the United States, judges were tempted to manipulate the concepts of either scientific community or general acceptance to exclude or admit evidence.[63] This manipulation has prevented general acceptance from providing a consistent standard of admissibility in the United States. A requirement of general acceptance also prevents relevant and reliable evidence from going before the fact finder merely because it is new.[64] Scientific work performed for the purposes of litigation will therefore often be inadmissible.[65] Further, commentators have challenged the claim that the test provides an assurance of reliability, as general acceptance of a theory does not guarantee empirical validation.[66]

The benefit of the judicial assessment approach is that it permits a greater degree of relevant and reliable scientific evidence to go before the fact finder. Such evidence will not be excluded merely because it is new. However, being a less stringent standard than the general acceptance test, less reliable material may be presented to a jury. Critics claim that ignorant juries are unable to correctly assess the "junk science" that is allowed in under a lower reliability standard than the general acceptance test.[67] The judicial

---

60 Above n 59, 1207.

61 JL Brown "DNA and *Kelly-Frye*: Who will Survive in California?" (1988) 11 Criminal Justice Journal 1, 27.

62 Above n 59, 1208.

63 EJ Chan "The 'Brave New World' of *Daubert*: True Peer Review, Editorial Peer Review, and Scientific Validity" (1995) 70 New York University Law Review 100.

64 EJ Imwinkelreid "A New Era in the Evolution of Scientific Evidence: A Primer on Evaluating the Weight of Scientific Evidence" (1981) 23 William and Mary Law Review 261.

65 For instance, in the *Daubert* case the defendant wished to offer in evidence a statistical re-analysis of the published epidemiological studies offered as proof by the plaintiff. Although the re-analysis was performed by a highly qualified and reputable expert, it was inadmissible as it had not been published or subject to peer review and therefore could not be said to have been accepted by the scientific community. However, the fact that the re-analysis was unpublished did not necessarily mean it was unreliable, although clearly it is a factor to take into account in making such a determination.

66 L Etlinger "Social Science Research in Domestic Violence Law: A Proposal to Focus on Evidentiary Use" (1995) 58 Albany Law Review 1259, 1279.

67 PW Huber "Junk Science in the Courtroom" (1992) 29 Val U L Rev 723.

assessment test has also been criticised for having unrealistic expectations of judges' abilities to assess scientific evidence and for requiring them to become "amateur scientists".[68] Further, the test will consume more judicial time and resources as judges will have to come to grips with the evidence themselves. It will also be more costly for litigants.

1 Can juries cope with scientific evidence?

Freckelton has questioned the supposed inability of juries to cope with scientific evidence, and, therefore, the need to shield them with overly strict reliability tests.[69] There has, however, been little empirical testing of the ability of juries to understand expert evidence. Two studies conducted in the United States indicated that juries could not cope with scientific testimony and would take irrelevant factors, such as the experts' appearance, into account.[70] However, these studies have been criticised by Freckelton as being little more than anecdotal or arising out of a single and unusual case.[71] A later Hong Kong study indicated 22 out of 58 jurors experienced difficulties in a complex trial, but these difficulties were due to a number of factors not all of which related to expert evidence.[72]

Freckleton concludes that the empirical information available about the difficulties encountered by jurors in understanding expert evidence specifically is extremely limited in both extent and quality. His conclusions have been supported by later studies. Jacobs notes that recent surveys support the conclusion that jurors are capable of deciding complex scientific issues.[73] Indeed, jurors may give less weight to scientific expert evidence than to evidence obtained through personal understanding.[74] Vidmar and Schuller, after reviewing studies on the effect of expert "social framework" testimony on jurors, conclude that jurors are not overwhelmed by such testimony.[75] Thus, there is no

---

68 Above n 8, 487.

69 I Freckelton "The Egg Shell Jury and the Expert Evidence Conundrum", a paper delivered at the Australian Law Commission Conference on Expert Evidence, 1993. On file with the author.

70 Forkosch "The Lie Detector and the Courts" (1938) 16 New York University Law Quarterly 202; P Rosenthal "Nature of Jury Response to the Expert Witness" (1983) 28 Journal of Forensic Sciences 528, cited above in Freckleton, n 86, 145.

71 Above n 69, 145.

72 Above n 69, 146.

73 M Jacobs "Testing the Assumptions Underlying the Debate about Scientific Evidence: A Closer Look at Juror Incompetence and Scientific Objectivity" (1993) 25 Conn L Rev 1083.

74 Above n 73, 1098.

75 N Vidmar and R Schuller "Juries and Expert Evidence: Social Framework Testimony" (1989) 52 Law and Contemporary Problems 133, 173.

strong empirical support for the often expressed judicial fear that juries will be overwhelmed by "shaky" scientific evidence and will give it more credibility than it deserves.

Freckelton also contends that those rules of expert evidence which help the fact finder to evaluate expert evidence (area of expertise, qualification rule and factual basis rule) are being applied with increasing stringency by the courts.[76] In contrast, those rules which arose out of concern to protect jurors from being intimidated and overwhelmed by expert evidence (the common knowledge rule and the ultimate issue rule) are being relaxed. He sees this as a move toward greater confidence in the abilities of juries. Freckelton is an Australian and no doubt his comments reflect principally on developments within that country. Nevertheless, to a considerable extent, they also hold true for New Zealand.[77]

2 *Should New Zealand judges be "gatekeepers"?*

The "gatekeeping" role assigned to judges by *Calder* and *Daubert* puts a heavy burden on the judiciary. Chief Justice Rehnquist, dissenting in *Daubert*, said that the task imposed on trial judges by the majority judgment involved "matters far afield from the expertise of judges" and required them to become "amateur scientists".[78] Others have argued that to implement *Daubert*, judges will have to achieve at least a basic level of scientific literacy involving an understanding of the scientific method, the rudiments of statistics and probability theory, and an appreciation of error factors and of the limitations of frequently used methods of observation, measurement and detection.[79] This will require a considerable investment of time and resources as judges and lawyers will not develop an understanding of the nature of science without a major educational and training effort.

The response to *Daubert* in the United States has been to try to educate judges to enable them to cope with the task of assessing the reliability of scientific evidence. A reference manual of scientific evidence was published by the Federal Judicial Centre in 1994 to assist judges in implementing effective management of expert evidence involving scientific issues.[80] Some scientific organisations and universities are conducting

---

76 Above n 69, 166.

77 For example, in *A-G v Equitycorp Industries Group Ltd (in statutory management)* [1995] 2 NZLR 135, the Court of Appeal doubted that the ultimate issue rule was still an absolute rule. In *R v Decha-Iamsukun* [1993] 1 NZLR 141, the Court of Appeal said that expert evidence was admissible even if the matter in question was to a considerable extent within the experience of the fact finder.

78 Above n 8, 487.

79 P Miller, B Rein and E Bailey "*Daubert* and the Need for Judicial Scientific Literacy" (1994) 77 Judicature 254.

80 Work on the Reference Manual actually began prior to *Daubert*. However, the decision encouraged

education programmes for judges on scientific issues and mathematics. Special training seminars for applying *Daubert* were organised for New York trial judges.

Unfortunately, there has not been a similar response in New Zealand. Judicial education is organised by the Department for Courts. No training is offered as a matter of course with regard to expert scientific evidence. However, judges may attend externally organised courses.

One method of helping judges cope with assessing the reliability of scientific evidence is the appointment of court experts. This was suggested by the Supreme Court in *Daubert*[81] and has also been proposed by the Law Commission.[82] The use of court experts may appear to go against the requirement that judges determine reliability themselves. However, such an expert could be used to explain concepts to the judge rather than acting a surrogate decision maker.[83] It would be part of the judge's educational process rather than an abdication of responsibility. Parties would remain free to lead their own experts and thus to counter any perceived deficiencies or bias on the part of the court-appointed expert.

3 *The resource implications of* Calder

It seems undeniable that a judicial assessment test will consume more time and resources than a general acceptance test. The relatively quick investigation of the status of a particular area of science will be replaced by time consuming hearings as to the reliability of new theories and techniques. There may be disputes over scientific evidence that parties would never bother leading under a general acceptance test. Nevertheless, a judicial assessment test in conjunction with procedural reform may reduce areas of disagreement, and thus reduce some costs.

4 *The procedural implications of* Calder

Under a judicial assessment test, experts will require sufficient time to investigate and test the challenged evidence. American commentators have recognised the need for a more sophisticated disclosure regime in the wake of *Daubert*.[84] The Law Commission has

---

the continuation and rapid completion of the project.

81 Above n 8, 484.

82 Above n 13, 37.

83 In "Improving Judicial Gatekeeping: Technical Advisors and Scientific Evidence" (1997) 110 Harvard Law Review 941, 947, it was suggested that procedures should be developed by the courts to prevent the judges from delegating their "gatekeeping" role to the expert.

84 MA Berger "Procedural Paradigms for Applying the *Daubert* Test" (1994) 78 Minn L Rev 1345, 1372.

also recognised the need for procedural reform as a part of the changes it proposes for the admissibility of expert evidence.[85] The Law Commission's proposals were a strong influence on Tipping J and the test in *Calder* closely resembles its recommendations.

The Law Commission proposes that opposing parties should have sufficient notice of expert evidence to enable them to investigate and test the evidence fully. The proposed rule[86] would require parties to give notice to each other, in advance of the trial, of the expert evidence they intended to call. The notice would include the name, address and qualifications of the proposed witness and the substance of the proposed evidence. Notice would be mandatory with the court having power to exclude unnotified evidence.

Such procedural reforms are necessary to ensure a fair process under the *Calder* test. New rules must ensure opposing experts are adequately prepared to test the scientific evidence, both at the admissibility hearing and (if the evidence is admitted) at trial. Procedural reform would add to trial efficiency by facilitating quick identification of the real issues and encouraging opposing experts to narrow their differences prior to trial. This would save time and reduce costs for both the judicial system and for litigants.

IV CONCLUSION

Although *Calder* has not been considered by the Court of Appeal, it is submitted that it is the most suitable approach for establishing the admissibility of scientific evidence. It provides a flexible test for ensuring that relevant and reliable scientific evidence is admissible, taking into account both scientific validity and policy factors. The general acceptance test, in contrast, has been found wanting by the jurisdiction which developed it and replaced with a test similar to that of *Calder*. Use of the *Calder* test will, however, require concurrent procedural reforms because opposing parties will need adequate notice to enable them to adequately test the evidence. It will also require a willingness in the judicial system to invest time and resources in educating judges so as to enable them to competently assess complex scientific evidence.

Calder illustrates the growing significance of scientific evidence in the justice system. Science is a powerful tool which can greatly aid the fact finder. However, it can also powerfully mislead the fact finder. The decision to "gatekeep" demonstrates a desire to make full and appropriate use of such evidence without abdicating judicial power and responsibility to the expert scientific witness.

---

85 Above n 13, 40.

86 Above n 13, 52, rule 5.