Preventing Miscarriages of Justice: The Reliability of Forensic Evidence and the Role of the Trial Judge as Gatekeeper

Chris Maxwell*

Eliminating unreliable forensic evidence is crucial to ensuring that criminal trials are fair. Recent experience, both in Australia and abroad, has demonstrated that flawed forensic evidence can lead to grave miscarriages of justice. Alarmingly, however, there has been no rigorous scientific validation of any forensic method other than DNA analysis. In a jury trial, the trial judge needs to assess the reliability of scientific evidence before it is led, as that task is beyond the capability of most jurors. The role of the trial judge as "gatekeeper" on questions of reliability is firmly established in other jurisdictions – the United States, Canada and the United Kingdom – but the current interpretation of the Uniform Evidence Act prevents judges (in jurisdictions governed by that Act) from performing that function. The purpose of this article is to underline the critical importance of the gatekeeper role, and to highlight the urgent need for legislative intervention to reinstate it.

> The obvious risk in a criminal trial when expert evidence is led from a forensic scientist is that a jury will give the evidence more weight than it deserves. To prevent unfair prejudice of that kind, it is essential that the reliability of the expert evidence be established to the Court's satisfaction before it is led.[1]

In 1993, the US Supreme Court in *Daubert v Merrell Dow Pharmaceuticals Inc*[2] held that it was the responsibility of the trial judge to ensure that "all scientific testimony … is not only relevant but reliable".[3] Since *Daubert*, the role of the judge as guardian of the reliability of scientific evidence has come to be characterised as that of "gatekeeper".

In 2015, in *Tuite v The Queen*,[4] the Victorian Court of Appeal applied *Daubert*, holding that in a criminal trial the reliability of forensic evidence should be assessed by the trial judge as part of the determination of probative value under s 137 of the *Evidence Act*.[5] The touchstone of reliability, the Court said, should be proof of scientific validation.

In 2016, however, the High Court in *IMM v The Queen*[6] held that the assessment of probative value for the purposes of s 137 did not involve any consideration of the reliability of the proposed evidence. The judge was required to assess probative value on the assumption that the jury would accept the evidence. Accordingly, no question of reliability (or credibility) could arise.[7]

* President, Victorian Court of Appeal. This article is a revised version of a paper presented in July 2017 to the International Society for the Reform of Criminal Law, and in August 2017 to the Australian Academy of Forensic Sciences (Victorian Chapter).

[1] *Tuite v The Queen* (2015) 49 VR 196, 200 [11].

[2] *Daubert v Merrell Dow Pharmaceuticals Inc*, 509 US 579 (1993).

[3] *Daubert v Merrell Dow Pharmaceuticals Inc*, 509 US 579, 589 (1993).

[4] *Tuite v The Queen* (2015) 49 VR 196.

[5] Under *Evidence Act*, s 137, the court must refuse to admit evidence adduced by the prosecutor "if its probative value is outweighed by the danger of unfair prejudice to the defendant". The Court had earlier held, in *Dupas v The Queen* (2012) 40 VR 182, that an assessment of reliability was a necessary part of the determination of probative value.

[6] *IMM v The Queen* (2016) 257 CLR 300.

[7] *IMM v The Queen* (2016) 257 CLR 300, 315 [52]. See G Edmond, "Icarus and the Evidence Act: Section 137, Probative Value and Taking Forensic Science Evidence 'At Its Highest'" (2018) 41 MULR 106; D Hamer, "The Unstable Province of Jury Fact-Finding: Evidence Exclusion, Probative Value and Judicial Restraint after IMM v The Queen" (2018) 41 MULR 689.

On the present state of the law,[8] therefore, the judge in a criminal trial is unable to perform the "gatekeeper" role as defined in *Daubert* and adopted in *Tuite*. The purpose of this article is to underline the critical importance of that role, and to highlight the urgent need for legislative intervention to reinstate it.

The case for reinstating the gatekeeper role rests on three propositions. First, eliminating unreliable forensic evidence is crucial to ensuring that criminal trials are fair. As the Supreme Court of Canada said in 2007:

> Evidence that is not sufficiently reliable is likely to undermine the fundamental fairness of the criminal process.[9]

The danger, as that Court had pointed out previously, is that:

> expert evidence will be misused and will distort the fact-finding process. Dressed up in scientific language which the jury does not easily understand and submitted through a witness of impressive antecedents, this evidence is apt to be accepted by the jury as being virtually infallible and as having more weight than it deserves.[10]

Second, the reliability of most branches of forensic science is unproven. This lack of proven reliability was exposed by two landmark reports in the United States: the 2009 Report of the National Academy of Sciences[11] and the 2016 Report of the President's Council of Advisors on Science and Technology.[12] The NAS Report found that:

> With the exception of nuclear DNA analysis, no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source.[13]

And further:

> Much forensic evidence … is introduced in criminal trials without any meaningful scientific validation, determination of error rates, or reliability testing to explain the limits of the discipline.[14]

The multi-disciplinary committee which produced the NAS Report was chaired by Judge Harry Edwards, Senior Circuit Judge and Chief Judge Emeritus for the US Court of Appeals for the District of Columbia Circuit. He wrote about his experience in these terms:

> I started this project with no preconceived views about the forensic science community. Rather, I simply assumed, as I suspect many of my judicial colleagues do, that forensic science disciplines typically are well-grounded in scientific methodology and that crime laboratories are forensic science practitioners follow proven practices that ensure the validity and reliability of forensic evidence offered in court. I was surprisingly mistaken.[15]

Third, and self-evidently, jurors are in no position to assess the reliability either of the scientific foundation on which an expert's opinion rests or of the methodology adopted to arrive at the opinion. The PCAST Report made the point emphatically:

> The vast majority of jurors have no independent ability to interpret the probative value of results based on the detection, comparison, and frequency of scientific evidence. … The potential prejudicial impact is unusually high, because jurors are likely to overestimate the probative value of a 'match' between samples.

---

[8] In jurisdictions governed by the *Uniform Evidence Act*.

[9] *R v Trochym* [2007] 1 SCR 239, 260 [27].

[10] *R v Mohan* [1994] 2 SCR 9, 21; see also *R v Duke* (1979) 22 SASR 46, 48.

[11] *Committee on Identifying the Needs of the Forensic Science Community, National Research Council, Strengthening Forensic Science in the United States: A Path Forward* (National Academies Press, 2009) (NAS Report).

[12] *President's Council of Advisors on Science and Technology, Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* (2016) (PCAST Report).

[13] NAS Report, n 11, 7.

[14] NAS Report, n 11, 107–108.

[15] Harry T Edwards, "Solving the Problems that Plague the Forensic Science Community" (2009) 50 *Jurimetrics Journal* 5, 7. See also Harry T Edwards, "The National Academy of Sciences Report on Forensic Sciences: What It Means for the Bench and Bar" (2015) 51 *Jurimetrics Journal* 1.

In the United Kingdom and in Canada, courts have ruled that reliability is a condition of admissibility at common law.[16] In the United Kingdom, concerns over unreliable forensic evidence led the Law Commission in 2011 to recommend a statutory test for the admissibility of expert evidence in criminal proceedings. In 2014, in default of such legislation, the Lord Chief Justice issued a Practice Note adopting the criteria which the Commission had proposed for the assessment of reliability.

In the United States, advances in DNA technology have resulted in the exoneration of large numbers of individuals wrongfully convicted on the basis of faulty forensic science analysis. Growing disquiet about the shortcomings of forensic evidence prompted the US Congress in 2005 to commission the NAS to conduct the study which resulted in the NAS Report.

In 2015, President Obama asked PCAST to consider – in the light of the NAS Report – whether there were additional steps that could usefully be taken to strengthen the forensic science disciplines and ensure the validity of forensic evidence used in the legal system. PCAST reiterated the need to evaluate specific forensic methods to determine whether they had been scientifically established to be valid and reliable.

Focusing on "feature-comparison" methods, the PCAST Report concluded that the only way to determine whether such a method was reliable was to test it empirically, under circumstances reasonably similar to its intended use. Without actual empirical evidence of its ability to produce conclusions at a level of accuracy appropriate to its intended use, an examiner's conclusion that two samples were likely to have come from the same source was "scientifically meaningless".[17]

In Australia, by contrast, these issues have received comparatively little attention, despite sustained advocacy over many years by Professor Gary Edmond of UNSW and other scholars.[18] There has been no systematic, independent review of the state of forensic science, and questions of reliability have only occasionally arisen for appellate consideration.

As the wrongful conviction of David Eastman (see below) demonstrates, however, the risk of a miscarriage of justice resulting from flawed expert testimony is as real in this country as elsewhere. In considering how that risk might be addressed, there is much to be learnt from developments in other jurisdictions, and from the work of scholars both here and abroad.

One obvious solution would be for s 79 of the *Evidence Act* to be amended, along the lines recommended by the UK Law Commission, to make reliability a condition of the admissibility of expert evidence in criminal proceedings and to authorise the trial judge to make the threshold assessment of reliability.

Reinstatement of the gatekeeper role would, however, be only the first step. As other jurisdictions have recognised, separate questions arise about how best to equip judges to perform that function. The final section of this article describes recent procedural changes in Victoria, designed to enhance the scrutiny of expert evidence by trial judge and counsel.

MISCARRIAGES OF JUSTICE

In 1995, David Eastman was convicted by a jury of the murder of Colin Winchester, an Assistant Commissioner of the Australian Federal Police. A central element of the circumstantial case against him was the finding of gunshot residue in his car. The prosecution relied on expert opinion evidence to the effect that the residue found in the car matched that found at the crime scene.

In 2014, a Board of Inquiry, conducted by former Chief Justice Martin, concluded that a substantial miscarriage of justice had occurred, as "the issue of guilt was determined on the basis of deeply flawed

[16] *R v Duglosz* [2013] EWCA Crim 2, [11]; *R v Trochym* [2007] 1 SCR 239.

[17] PCAST Report, n 12, 27.

[18] See, eg, G Edmond, "The Admissibility of Forensic Science and Medicine Evidence under the Uniform Evidence Law" (2014) 38 Crim LJ 136; G Edmond, "What Lawyers Should Know about the Forensic 'Sciences'" (2015) 37 Adel L Rev 173; G Edmond, "Forensic Science Evidence and the Conditions for Rational (Jury) Evaluation" (2015) 39 MULR 75; T Ward et al, "Forensic Science, Scientific Validity and Reliability: Advice from America" [2017] Crim L R 357.

forensic evidence".[19] Evidence presented to the Board had a "devastating impact upon the reliability and the veracity of the [expert's] trial evidence".[20]

According to the Inquiry report, the forensic evidence had been "crucial" to the prosecution case. In his closing address, the prosecutor had repeatedly emphasised the reliability and importance of the evidence, and the judge had directed the jury that the expert's methodology had not been criticised and was supported by overseas experts.

Mr Martin said:

> Perhaps the best indication of how the jury is likely to have viewed the evidence of [the expert] concerning the gunshot residue is found in the view of the forensic investigation expressed by the trial Judge when sentencing [Eastman]:
>
> > This investigation must surely number as one of the most skilled, sophisticated and determined forensic investigations in the history of criminal investigations in Australia.
>
> This Inquiry has proved otherwise. It must be said that the inadequacies were not apparent at trial and the trial Judge had no reason for doubting the reliability of the forensic evidence, but his Honour's view highlights the danger of taking contentious forensic evidence at face value without properly investigating the records and the basis upon which opinions are expressed.

Among the deficiencies in the evidence which the report identified were the following:

- the expert had regularly failed to comply with accepted forensic practice and frequently failed to have his work peer reviewed;
- there were anomalies in the expert's database, of which he was aware but which were not disclosed; and
- the expert could not have undertaken some of the organic analysis on which he claimed to have based his opinions.

As a result, the Inquiry found, the purported expert evidence "lacked a proper scientific foundation".[21] The Inquiry recommended that Eastman's conviction be quashed. By the time that occurred, Eastman had spent 19 years in jail.

At his retrial in 2018, Eastman was acquitted. The expert evidence was not led at the retrial. Instead, it was agreed between prosecution and defence that:

> Scientific testing cannot determine whether individual GSR particles come from the same firearm or the same cartridge case. Accordingly, it cannot be determined whether the GSR particles found at the crime scene and in Mr Eastman's car came from the same firearm and the same cartridges or from different firearms and different cartridges.[22]

On 19 December 2018, the Texas Court of Criminal Appeals declared Steven Chaney an innocent man, after he had spent more than 25 years in jail for murder. The centrepiece of the prosecution case against him had been the testimony of two forensic odontologists, who said that bite marks found on the victim's forearm matched Chaney's teeth.

One of the witnesses had testified that there was only a "one to a million" chance that someone other than Chaney had bitten the victim, as the mark was a "perfect match", with "no discrepancies" and "no inconsistencies". The other testified that he was certain to a "reasonable degree of dental certainty" that Chaney was the "biter".

In its 2018 judgment, the Court of Criminal Appeals accepted that the bite mark evidence had been wholly discredited by subsequent peer-reviewed studies showing that:

---

[19] Martin J (Acting), "Inquiry into the Conviction of David Harold Eastman for the Murder of Colin Stanley Winchester", *Report of the Board of Inquiry*, (29 May 2014) [1832].

[20] Martin J (Acting), "Inquiry into the Conviction of David Harold Eastman for the Murder of Colin Stanley Winchester", *Report of the Board of Inquiry*, (29 May 2014) [1103].

[21] Martin J (Acting), "Inquiry into the Conviction of David Harold Eastman for the Murder of Colin Stanley Winchester", *Report of the Board of Inquiry*, (29 May 2014) [1115]–[1124].

[22] The author is indebted to George Georgiou SC (now a judge on the County Court of Victoria), counsel for Eastman in the 2018 retrial, for providing a copy of the Agreement as to Facts dated 18 June 2018.

> the uniqueness of human dentition can never be established within measurement error, and even if a person's dentition was unique, the skin cannot faithfully preserve that uniqueness such that a particular dentition can be associated with a particular putative bite mark.[23]

Bite mark analysis had never been scientifically validated.[24] As Fabricant and Carrington explain, it came to be viewed as "generally accepted" science[25] in American courts without any research having been undertaken to validate the technique's two underlying hypotheses, namely, that a bite mark expert could make an association between a bite mark and a suspect's "dentition", and could provide "a scientifically valid estimate" of the rarity or frequency of that association.[26]

Between 2000 and 2013, DNA testing led to the exoneration of at least 24 individuals convicted or indicted through the use of bite mark evidence.[27] On 14 January 2019, the Texas Court of Criminal Appeals stayed the execution of Blaine Milam, who in 2010 was convicted of murder on the basis of bite mark evidence. The stay – ordered only a day before the execution was scheduled – was granted in part "because of recent changes in the science pertaining to bite mark comparisons".[28]

Assessing Evidentiary Reliability: from "General Acceptance" to "Proven Validity"

Until *Daubert*, the "dominant standard" for determining the admissibility of novel scientific evidence was the so-called general acceptance test, laid down in 1923 by the then Court of Appeals for the District of Columbia in *Frye v United States*.[29] The *Frye* test was formulated in these terms:

> Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, *the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs*.[30]

In 1984, in *R v Bonython*,[31] the South Australian Full Court adopted a similar approach. King CJ (with whom Matheson and Bollen JJ agreed) said:

> Before admitting the opinion of a witness into evidence as expert testimony, the judge must consider and decide two questions. The first is whether the subject matter of the opinion falls within the class of subjects upon which expert testimony is permissible. This first question may be divided into two parts:
>
> (a) whether the subject matter of the opinion is such that a person without instruction or experience in the area of knowledge or human experience would be able to form a sound judgment on the matter without the assistance of witnesses possessing special knowledge or experience in the area[;] and
>
> (b) whether the subject matter of the opinion forms part of *a body of knowledge or experience which is sufficiently organized or recognized to be accepted as a reliable body of knowledge or experience,* a special acquaintance with which by the witness would render his opinion of assistance to the court.

In *Daubert,* the Supreme Court concluded that the *Frye* test had been superseded by r 702 of the *Federal Rules of Evidence*, which at that time provided as follows:

---

[23] *Ex Parte Steven Mark Chaney* (Texas Court of Criminal Appeal, General List No WR-84 09-091, 19 December 2018) slip op 31.

[24] NAS Report, n 11, 174, 176.

[25] The phrase comes from *Frye v United States*, 293 F 1013, 1014 (DC Cir, 1923).

[26] MC Fabricant and T Carrington, "The Shifted Paradigm: Forensic Science's Overdue Evolution from Magic to Law" (2016) 4(1) *Virginia Journal of Criminal Law* 38–39.

[27] MC Fabricant and T Carrington, "The Shifted Paradigm: Forensic Science's Overdue Evolution from Magic to Law" (2016) 4(1) *Virginia Journal of Criminal Law* 22.

[28] J McCullough, "Texas Court Stops First Execution of 2019, Citing Changes in Intellectual Disability Law and Bite Mark Science", *The Texas Tribune*, 14 January 2019.

[29] *Frye v United States*, 293 F 1013 (1923).

[30] *Frye v United States*, 293 F 1013, 1014 (1923) (emphasis added).

[31] *R v Bonython* (1984) 38 SASR 45.

> If scientific, technical, or other specialized 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.[32]

Rejection of the *Frye* test did not mean, however, that there were no limits on the admissibility of "purportedly scientific evidence".[33] On the contrary, the Court said:

> The trial judge must ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable.
>
> The primary locus of this obligation is Rule 702, which clearly contemplates some degree of regulation of the subjects and theories about which an expert may testify.
>
> …
>
> The subject of an expert's testimony must be "scientific … knowledge". The adjective "scientific" implies a grounding in the methods and procedures of science. Similarly, the word "knowledge" connotes more than subjective belief or unsupported speculation. The term "applies to any body of known facts or to any body of ideas inferred from such facts or accepted as truths on good grounds" … Of course, it would be unreasonable to conclude that the subject of scientific testimony must be "known" to a certainty; arguably, there are no certainties in science.
>
> But, in order to qualify as "scientific knowledge", an inference or assertion must be derived by the scientific method. Proposed testimony must be supported by appropriate validation – ie, "good grounds", based on what is known. In short, the requirement that an expert's testimony pertain to "scientific knowledge" establishes a standard of evidentiary reliability.[34]

The "overarching subject" of the admissibility inquiry, the Court said, should be:

> the scientific validity – and thus the evidentiary relevance and reliability – of the principles that underlie a proposed submission. The focus, of course, must be solely on principles and methodology, not on the conclusions that they generate.[35]

The inquiry must be a "flexible one",[36] the Court said, to which the following considerations would ordinarily be relevant:

(a) whether the theory or technique can be — and has been — tested;
(b) whether the theory or technique has been subjected to peer review and publication; and
(c) the known or potential rate of error, and the existence and maintenance of standards controlling the technique's operation.[37]

And there might still be a role for a test of general acceptance:

> A "reliability assessment does not require, although it does permit, explicit identification of a relevant scientific community and an express determination of a particular degree of acceptance within that community". – Widespread acceptance can be an important factor in ruling particular evidence admissible, and "a known technique which has been able to attract only minimal support within the community" … may properly be viewed with scepticism.[38]

Although *Daubert* was a decision under the *Federal Rules of Evidence*, the Supreme Court of Canada has treated the approach as equally applicable to the question of admissibility at common law. The Court has emphasised that:

---

[32] The rule was amended in 2000 to include a requirement that the testimony be "the product of reliable principles and methods": Andrew Ligertwood and Gary Edmond, *Australian Evidence: A Principled Approach to the Common Law and the Uniform Evidence Acts* (LexisNexis Butterworths, 5th ed, 2010) 624 [7.52].

[33] *Daubert v Merrell Dow Pharmaceuticals Inc*, 509 US 579, 589 (1993).

[34] *Daubert v Merrell Dow Pharmaceuticals Inc*, 509 US 579, 589–590 (1993) (emphasis added) (citations omitted).

[35] *Daubert v Merrell Dow Pharmaceuticals Inc*, 509 US 579, 594–595 (1993).

[36] *Daubert v Merrell Dow Pharmaceuticals Inc*, 509 US 579, 594 (1993).

[37] *Daubert v Merrell Dow Pharmaceuticals Inc*, 509 US 579, 593–594 (1993).

[38] *Daubert v Merrell Dow Pharmaceuticals Inc*, 509 US 579, 594 (1993) (citations omitted).

> The trial judge should take seriously the role of "gatekeeper". The admissibility of the expert evidence should be scrutinised at the time it is proffered, and not allowed too easy an entry on the basis that all of the frailties could go at the end of the evidence to weight rather than admissibility.[39]

In *J-L J*, the Court adopted the "reliable foundation" test laid down in *Daubert*, and applied the considerations there set out. The Court also endorsed the statement from *Mohan* that novel science should be subject to special scrutiny:

> In summary therefore, it appears from the foregoing that expert evidence which advances a novel scientific theory or technique is subjected to special scrutiny to determine whether it needs a basic threshold of reliability.[40]

In 2011, the author of the Supreme Court's opinion in *J-L J*, Ian Binnie J, addressed the Judicial Commission of New South Wales and the Judicial College of Victoria. He contrasted two decisions – one Canadian, one Australian – dealing with the "so-called science of barefoot morphology". In *R v Dimitrov*,[41] the Ontario Court of Appeal held that evidence about the imprints left by a person inside shoes or boots should have been stopped by the trial judge at "the gateway". The purported expert evidence failed to meet any of the *Daubert* criteria.

In *Rose*, by contrast, the South Australian Court of Criminal Appeal upheld the judge's decision to admit evidence of that kind, holding that it met the *Bonython* general acceptance test:

> Characteristics of the feet and points of comparison between wear marks in the shoes … forms part of the body of knowledge or experience which is sufficiently organised or recognised to be accepted as a reliable body of knowledge or experience.[42]

Binnie J commented:

> If the Daubert criteria were accepted in Australian law, the absence of published research methodology would have been very much to the point in Rose. So too, would the absence of any test of the podiatrists' methodologies or effort to determine the error rate in matching shoes to feet. The Ontario Court of Appeal set aside Dimitrov's original conviction on the basis that the trial judge had failed to perform a proper gatekeeper role. At the retrial in 2006, Dimitrov was readily acquitted. As between Rose and Dimitrov, I believe the critical analysis of the methodology at the admissibility stage advocated by Dimitrov is the preferable course of action.[43]

Just such a critical analysis was undertaken by Mullighan J in the South Australian Supreme Court in *R v Kargher*.[44] The case concerned the reliability of Profiler Plus, a methodology for interpreting DNA profiles. His Honour first considered whether the methodology had been "accepted by the relevant scientific community as reliable and accurate for DNA analysis".[45] His Honour then turned to the question whether the system had been "validated by men and women of science, found to be accurate and reliable and … accepted by them for use in the forensic context".[46]

His Honour quoted the following United States guidelines in relation to validation:[47]

> Validation is the process used by the scientific community to acquire the necessary information to assess the ability of a procedure to reliably obtain a desired result, determine the conditions under which such results can be obtained and determine the limitations of the procedure. The validation process identifies

[39] *R v J-L J* [2000] 2 SCR 600, 613.

[40] *R v J-L J* [2000] 2 SCR 600, 616 [35].

[41] *R v Dimitrov* (2003) 68 OR (3d) 641.

[42] *R v Rose* (1993) 69 A Crim R 1, 9.

[43] I Binnie, "Wrongful Convictions and the Magical Aura of Science in the Court Room" (2011) 10 TJR 141, 158 (emphasis in original). See also I Binnie and V Park-Thompson, "The Perils of Law Office Science" (2015) 36 Adel L Rev 125.

[44] *R v Karger* (2001) 83 SASR 1.

[45] *R v Karger* (2001) 83 SASR 1, [181], see also [188], [189], [228].

[46] *R v Karger* (2001) 83 SASR 1, [465].

[47] The guidelines were published by the body now known as the Scientific Working Group on DNA Analysis Methods, which has published current, equivalent guidelines: Scientific Working Group on DNA Analysis Methods, *Validation Guidelines for DNA Analysis Methods* (2012).

> the critical aspects of a procedure which must be carefully controlled and monitored. Validation studies must have been conducted by the DNA laboratory or scientific community prior to the adoption of a procedure by the DNA laboratory.[48]

Mullighan J concluded that, on the evidence before him, the methodology relied on had been adequately validated. The validation studies had been

> conducted without error and produced results which enabled the Forensic Science Centre to establish appropriate standards and protocols, including threshold levels which permit *accurate and reliable analysis and interpretation of results*, including for example with low levels of DNA.[49]

Tuite also involved close scrutiny by a trial judge of the scientific validity of forensic evidence. In that case, the DNA evidence was to be presented in the usual form of a "likelihood ratio", but calculated using a recently developed software package, known as "STRmix". At a pre-trial hearing, the accused challenged the admissibility of the DNA evidence on the ground that the new methodology was not – or had not been shown to be – sufficiently reliable for use in criminal trials. The methodology was largely untested, it was said, and had not been generally accepted by the forensic science community.

The pre-trial hearing extended over some 22 days, in the course of which the judge heard evidence from the three prosecution experts and one defence expert. Her Honour rejected the application to exclude the evidence, and her decision was upheld on an interlocutory appeal.[50]

The Court of Appeal concluded that the touchstone of reliability for scientific evidence must be trustworthiness, which in turn depended on validation.[51] For this purpose, the Court adopted the following statement from *Daubert*:

> We note that scientists typically distinguish between "validity" (does the principle support what it purports to show?) and "reliability" (does application of the principle produce consistent results?). Although "the difference between accuracy, validity, and reliability may be such that each is distinct from the other by no more than a hen's kick," … our reference here is to evidentiary reliability – that is, trustworthiness. In a case involving scientific evidence, evidentiary reliability will be based upon scientific validity.[52]

On this view, the focus of attention for the purposes of assessing the reliability of scientific evidence should be on proof of validation. Ideally, there should be proof of both in-house validation and independent external validation. But, as the Court noted, the commercialisation of forensic science makes this latter requirement increasingly difficult to satisfy.

The Court of Appeal suggested that the focus on proven validation had a number of advantages:

> First, and most importantly, it means that the scrutiny of scientific evidence in the judicial process will apply the rigour which the discipline of science itself demands. As it was put in Daubert, evidentiary reliability will be based on scientific validity. Secondly, the trial judge considering scientific evidence will ordinarily be able to assess the sufficiency of validation – based on the published results of validation tests – without needing to acquire particular expertise in the relevant field of science.
>
> Thirdly, validation studies provide a framework which assists the judge – and, ultimately, the jury – to evaluate the evidence. Fourthly, this approach avoids what we consider to be the unworkable imprecision of a "general acceptance" test, and would ensure that new developments and novel techniques are not excluded, provided always that their scientific validity is established to the satisfaction of the court.[53]

The United Kingdom Practice Direction

In the United Kingdom, as mentioned earlier, a 2011 report of the Law Commission recommended enactment of a provision to the effect that an expert's opinion evidence would be admissible only if

[48] The wording of the corresponding National Association of Testing Agencies standard was identical: *R v Karger* (2001) 83 SASR 1, 99 [140], [461].

[49] *R v Karger* (2001) 83 SASR 1, [543] (emphasis added).

[50] *Criminal Procedure Act 2009* (Vic) s 295.

[51] In *IMM v The Queen* (2016) 257 CLR 300, Gageler J (in dissent) equated reliability with trustworthiness: [82].

[52] *Daubert v Merrell Dow Pharmaceuticals Inc*, 509 US 579, 590 fn 9 (1993).

[53] *Tuite v The Queen* (2015) 49 VR 196, 225–226 [103]–[104].

it were "sufficiently reliable".[54] When the Government declined to legislate, the Lord Chief Justice in 2014 published a new Practice Direction on Expert Evidence in Criminal Proceedings, which relevantly states:

> The common law, therefore, remains the source of the criteria by reference to which the court must assess the admissibility and weight of such evidence.
>
> …
>
> Nothing at common law precludes assessment by the court of the reliability of an expert opinion by reference to substantially similar factors to those the Law Commission recommended as conditions of admissibility, and courts are encouraged actively to enquire into such factors.
>
> Therefore factors which the court may take into account in determining the reliability of expert opinion, and especially of expert scientific opinion, include:
>
> (a) the extent and quality of the data on which the expert's opinion is based, and the validity of the methods by which they were obtained;
> (b) if the expert's opinion relies on an inference from any findings, whether the opinion properly explains how safe or unsafe the inference is (whether by reference to statistical significance or in other appropriate terms);
> (c) if the expert's opinion relies on the results of the use of any method (for instance, a test, measurement or survey), whether the opinion takes proper account of matters, such as the degree of precision or margin of uncertainty, affecting the accuracy or reliability of those results;
> (d) the extent to which any material upon which the expert's opinion is based has been reviewed by others with relevant expertise (for instance, in peer reviewed publications), and the views of those others on that material;
> (e) the extent to which the expert's opinion is based on material falling outside the expert's own field of expertise;
> (f) the completeness of the information which was available to the expert, and whether the expert took account of all relevant information in arriving at the opinion (including information has to the context of any facts to which the opinion relates);
> (g) if there is a range of expert opinion on the matter in question, where in the range the expert's own opinion lies and whether the expert's preference has been properly explained; and
> (h) whether the expert's methods followed established practice in the field and, if they did not, whether the reason for the divergence has been properly explained.
>
> In addition, in considering reliability, and especially the reliability of expert opinion, the court should be astute to identify potential flaws in such opinion which detract from its reliability, such as:
>
> (a) being based on a hypothesis which has not been subjected to sufficient scrutiny (including, where appropriate, experimental or other testing), or which has failed to stand up to scrutiny;
> (b) being based on an unjustifiable assumption;
> (c) being based on flawed data;
> (d) being on an examination, technique, method or process which was not properly carried out or applied, or was not appropriate for use in the particular case; and
> (e) relying on an inference or conclusion which has not been properly reached.[55]

In *H v The Queen*,[56] the English Court of Appeal noted that these reforms had been prompted by "real concern about the use of unreliable or inappropriate expert evidence". Their introduction meant, the Court said, that:

> a new and more rigorous approach on the part of advocates and the courts to the handling of expert evidence must be adopted.[57]

[54] Law Commission, *Expert Evidence in Criminal Proceedings in England and Wales*, Report No 325 (2011).

[55] Lord Chief Justice of England and Wales, *Criminal Practice Direction 19A – Expert Evidence* (emphasis added). For an assessment of the likely impact of this reform, see M Stockdale and A Jackson, "Expert Evidence in Criminal Proceedings: Current Challenges and Opportunities" (2016) 80 J Crim L 344.

[56] *H v The Queen* [2014] EWCA Crim 1555.

[57] *H v The Queen* [2014] EWCA Crim 1555, [44].

THE ROLE OF LAWYERS AND JUDGES

The Supreme Court in *Daubert* defined the judge's "gatekeeper role" as entailing:

> a preliminary assessment of whether the reasoning or methodology underlining the testimony is scientifically valid and of whether that reasoning or methodology properly can be applied to the facts in issue.[58]

As American experience shows, however, having a clear set of criteria for testing evidentiary reliability is no guarantee that unreliable scientific evidence will be kept out of criminal trials. Edmond and Roberts have concluded that admissibility outcomes are not discernibly different when jurisdictions that have adopted the *Daubert* approach are compared to those that have not.[59]

In the adversary system, the responsibility for ensuring that forensic evidence is reliable lies with judges and practitioners, in particular prosecutors. The performance to date has been mixed at best. In its 2005 Report, the Science and Technology Committee of the House of Commons highlighted the responsibility of judges and lawyers for the "systems failures" in cases where unreliable forensic evidence had resulted in a miscarriage of justice.[60]

The NAS Report concluded that *Daubert* had done little to improve the use of forensic science evidence in criminal cases. The report adopted the following assessment:

> For years in the forensic science community, the dominant argument against regulating experts was that every time a forensic scientist steps into a court room, his work is vigorously peer reviewed and scrutinised by opposing counsel. A forensic scientist might occasionally make an error in the crime laboratory, but the crucible of court room cross-examination would expose it at trial. This "crucible", however, turned out to be utterly ineffective. … Unlike the extremely well-litigated civil challenges, the criminal defendant's challenge is usually perfunctory.[61]

In their 2016 article, Fabricant and Carrington delivered a powerful critique of what they describe as the "abdication" by US courts of their gatekeeper role.[62] They criticise the:

> failure of courts to distinguish between magic and science in the first instance and the judicial system's continuing reflexive reliance on deeply flawed, scientifically invalid precedents to support the admissibility of false and misleading evidence.[63]

In 2011, US Federal Judge Nancy Gertner said:

> Until courts address the deficiencies in the forensic sciences – until courts do what *Daubert* requires that they do – there will be no meaningful change here.[64]

Earlier, in *United States v Green*,[65] Judge Gertner had reluctantly admitted evidence of toolmark identification, while acknowledging that such evidence should not be considered admissible under the *Daubert* test. The judge pointed out that:

> The problem for the defence is that every single court post-Daubert has admitted this testimony, sometimes without any searching review, much less a hearing.

She added:

> [T]he more courts admit this type of toolmark evidence without requiring documentation, proficiency testing, or evidence of reliability, the more sloppy practices will endure; we should require more.[66]

---

[58] *Daubert v Merrell Dow Pharmaceuticals Inc*, 509 US 579, 592–593 (1993).

[59] G Edmond and A Roberts, "Procedural Fairness, the Criminal Trial and Forensic Science and Medicine" (2011) 33 Syd LR 359, 377.

[60] *Forensic Science on Trial* 2004–2005, HC 96–1, [169]–[170].

[61] PJ Neufeld, "The (Near) Irrelevance of Daubert to Criminal Justice: and Some Suggestions for Reform" (2005) 95 *American Journal of Public Health* S107; NAS Report, n 11, 95–96, 106–107.

[62] MC Fabricant and T Carrington, "The Shifted Paradigm: Forensic Science's Overdue Evolution from Magic to Law" (2016) 4 Va J Crim L 1.

[63] Fabricant and Carrington, n 62, 7. See also R Balko and T Carrington, "Bad Science Puts Innocent People in Gaol – And Keeps Them There", The Washington Post, 21 March 2018.

[64] N Gertner, "Commentary on the Need for a Research Culture in the Forensic Sciences" (2011) 58 UCLA L Rev 789, 790.

[65] *United States v Green*, 405 F Supp 2d 104 (D Mass, 2005).

[66] *United States v Green*, 405 F Supp 2d 104, 108–109 (D Mass, 2005).

If judges are to be equipped to discharge this function, they will need to become more scientifically literate and, to that end, will need expert assistance. The NAS Report said:

> The judicial system is encumbered by judges and lawyers who generally lack the scientific expertise necessary to comprehend and evaluate forensic science in an informed manner.

The House of Commons Science and Technology Committee recommended in 2005 that training and continuing professional development in forensic evidence be made compulsory for lawyers, and that judges should "have more scientific input when making decisions about whether to admit expert evidence".[67] The Law Commission likewise recommended training for both new and experienced practitioners.[68]

In Australia, the current system of judicial training does not extend to preparing judges to perform the role of scientific evaluator. Moreover, as Edmond and Martire have argued:

> tradition and adjectival rules prevent judges from unilateral engagement with materials, however authoritative, that are not adduced by the parties. Consequently, there are relatively few ways for Australian judges to find out about mainstream scientific perspectives. Australian judges tend to be insulated from, and practically incapable of engaging with, scientific knowledge and the best scientific advice.[69]

Performing the Gatekeeper Role

How would a trial judge go about assessing reliability in practice? A number of possibilities suggest themselves. The first is for the judge to undertake her own *Daubert* hearing. As mentioned earlier, that is what the trial judge did in *Tuite*, a process which took some 22 days.[70] In the "ephedra" case, Judge Rakoff of the US District Court undertook a two-week evidentiary hearing, in which he heard from the competing experts and made up his own mind as to which parts of the evidence could be relied on for which purpose.[71]

In the United States at least, a second possibility is for the judge to seek the assistance of a court-appointed expert. As Judge Domitrovich of Pennsylvania has pointed out, Federal Rule of Evidence 706 provides for court-appointed experts to assist judges in their gatekeeping role.[72]

Judge Domitrovich argues:

> Somehow trial judges need to acquire the requisite tools and knowledge to assess the reliability of the methods used by forensic scientists and other experts. To do so, they need assistance.[73]

And again:

> [Judges] must become sophisticated consumers of science who are capable of understanding the core issues relating to disputed evidence.[74]

Third, consideration could be given to the proposal, first formulated by Edmond and Roberts in 2011,[75] for a multidisciplinary advisory panel to assess the reliability of forensic science techniques and to provide authoritative advice to judges at the admissibility stage. They proposed that such a panel:

---

[67] *Forensic Science on Trial* 2004–2005, HC 96–1, [177]–[182].

[68] As to the limited progress since, see G Davies and E Piasecki, "No More Laissez Faire? Expert Evidence, Rule Changes and Reliability: Can More Effective Training for the Bar and Judiciary Prevent Miscarriages of Justice?" (2016) 80 J Crim L 327.

[69] G Edmond and K Martire, "Forensic Science in Criminal Courts: The Latest Scientific Insights" (2016) 42 ABR 1, 15.

[70] See also *Thompson v Johnson & Johnson Pty Ltd* (1989) Aust Torts Reports 80-278.

[71] J Rakoff, "Science and the Law: Uncomfortable Bedfellows" (2008) 38 Seton Hall L Rev 1379, 1391.

[72] S Domitrovich, "Fulfilling Daubert's Gatekeeping Mandate through Court-Appointed Experts" (2016) 106 *Journal of Criminal Law and Criminology* 35.

[73] Domitrovich, n 72, 41.

[74] Domitrovich, n 72, 48.

[75] G Edmond and A Roberts, "Procedural Fairness, The Criminal Trial and Forensic Science and Medicine" (2011) 33 SLR 359, 389–392. See also E Cunliffe and G Edmond, "What Have We Learned? Lessons from Wrongful Convictions in Canada" in B Berger et al (eds), *To Ensure that Justice is Done: Essays in Memory of Marc Rosenberg* (Carswell, 2017).

> be composed of a range of members with different scientific and technical training, though representation from statistics, biomedicine, engineering, chemistry and experimental psychology would seem to be highly desirable. One or two representatives of the forensic sciences should be included along with an experienced trial judge and a legal scholar. There ought, in addition, to be scope to extend membership should particular forms of expertise be required.[76]

Finally, and more generally, the forensic sciences themselves need to be encouraged to adopt a "research culture". This entails:

> a commitment to conducting, participating in, and relying upon high quality empirical research.[77]

This idea was first propounded in 2011 by Jennifer Mnookin and others.[78]

More recently, Koehler and Meixner from North Western Law School have recommended:

> a series of scientific studies that may provide guidance to legal decision-makers about the reliability and validity of forensic science conclusions.[79]

In *Tuite* the Court of Appeal suggested that consideration be given to the establishment – preferably on a national basis – of something equivalent to the UK's Forensic Science Regulator and Forensic Science Advisory Council (which provides advice to the Regulator). Their remit encompasses the establishment and maintenance of quality standards relating to forensic science organisations, processes and techniques, and specifically includes the development of "procedures for validating and approving new technologies".

In the Court's view, a regulatory body of that kind could provide independent, expert assessments of the reliability of particular scientific theories and techniques, which would in turn greatly assist the trial court in determining the reliability of evidence to be given by a particular expert.[80]

Improving Procedures for Scrutinising Expert Evidence

In 2008, in *R v Klamo*,[81] the Victorian Court of Appeal quashed a conviction for "baby-shaking" manslaughter. The Court concluded that, on a proper analysis, the expert evidence could not have established that any act of the accused man had caused the baby's death.[82] By the time the conviction was quashed, however, the accused man had spent a year in jail. Plainly enough, something needed to be done.

Professor Stephen Cordner, then the Director of the Victorian Institute of Forensic Medicine (VIFM), had been the Crown's expert in that trial. In an article published subsequently, Professor Cordner said that he had not been spoken to by the prosecutor, or by defence counsel, at any time before Mr Klamo's trial.

In his view, the problems in the case might well have been avoided if he had:

- been invited to a pre-committal or pre-trial conference with the prosecutor;
- been interviewed by defence counsel; or
- given evidence on a voir dire.[83]

---

[76] Edmond and Roberts, n 75, 390.

[77] J Koehler and J Meixner, "An Empirical Research Agenda for the Forensic Sciences", (2016) *Journal of Criminal Law and Criminology* 1, 31.

[78] J Mnookin et al, "The Need for a Research Culture in the Forensic Sciences" (2011) 58 UCLA L Rev 725.

[79] Koehler and Meixner, n 77, 5. See also the procedures reviewed in the PCAST report and recent initiatives by the Royal Society of London and Royal Society of Edinburgh, namely *Forensic Gait Analysis: A Primer for Courts* (2017) and *Forensic DNA Analysis: A Primer for Courts* (2017).

[80] *Tuite v The Queen* (2015) 49 VR 196, 227 [109]. See *R v Reed* [2009] EWCA Crim 2698 [111].

[81] *R v Klamo* (2008) 18 VR 644.

[82] *R v Klamo* (2008) 18 VR 644, 657 [53]–[55].

[83] S Cordner, "R v Klamo: An Example of Miscommunication and Misunderstanding of Expert Evidence Where the Conviction Was Overturned" (2012) 44 AJFS 323, 329.

In late 2012, Professor Cordner and the author convened what came to be known as the Forensic Evidence Working Group. It comprised judges (County Court, Supreme Court and Court of Appeal), forensic scientists (both from VIFM and Victoria Police Forensic Sciences Centre) and legal practitioners, including representatives of the Office of Public Prosecutions, Victoria Legal Aid and Criminal Bar Association. The then Director of Public Prosecutions, John Champion,[84] was a founding member.

A consensus quickly emerged that there was a need for a set of rules which would govern both the *content* of expert reports and the *procedures* for ensuring adequate pre-trial scrutiny of expert reports, both by prosecution and defence and – where necessary – by the trial judge.

The Group proceeded to develop a new Practice Note entitled "Expert Evidence in Criminal Trials", which came into force on 1 July 2014.[85] The Practice Note:

(1) contains detailed specifications of what an expert report must contain;
(2) establishes procedures to enable defence counsel to confer with a prosecution expert before trial;
(3) enables the trial judge to direct experts to confer and prepare a joint report; and
(4) makes provision for "concurrent evidence" in a case where both prosecution and defence rely on expert evidence.[86]

The latter procedure, known as the "hot tub", has been used for many years – and with considerable success – in civil litigation in Australia. The basic procedure is for the experts to be seated side by side in the courtroom and for them to be questioned by counsel on both sides, by the judge and, where appropriate, by each other. Experience shows that this procedure greatly reduces the time taken in the presentation of expert evidence and, most importantly, makes it easier for the judge to assess the competing views and to be clear about the areas of disagreement.

For criminal trials, however, this is uncharted territory. The new procedure has not yet been invoked in a criminal trial, although it has been used on several occasions on the voir dire.[87]

Trial judges have provided positive feedback since the Practice Note came into force. A County Court judge described how the new procedures had dramatically changed the course of events in a drug importation prosecution and, ultimately, led to its discontinuance. According to the judge:

> It was the most diabolically complicated expert evidence case I have ever seen. I was despairing as to how the jury would possibly understand one word of the evidence. [Applying the new Practice Note] I ordered the experts to confer and advise what areas they agreed on and what they disagreed on. As a consequence of what came out of the expert conference the DPP reconsidered the case and decided to discontinue it.

In 2016, the Court of Appeal handed down a decision which illustrates the advantages which these new procedures offer.[88] The offender was charged with culpable driving, following a road accident in which two people were killed. There was a significant issue about the offender's cognitive impairment and expert reports were prepared on both sides. The offender indicated early on that he would not plead to the culpable driving charges but would plead to charges of dangerous driving causing death.

The judge who was due to preside over the trial invoked the procedures under the Practice Note. He directed the experts to confer, with a view to ascertaining whether there were areas on which they could reach agreement.

In the event, the experts produced a joint report, the full text of which is set out in the Court of Appeal's judgment, in which they confirmed that the cognitive impairment materially affected the offender's state of mind at the time of the accident. Once that statement was provided, the prosecution accepted the plea. The time and expense and trauma of a trial was thus avoided.[89]

---

[84] Now a judge of the Supreme Court of Victoria.

[85] Supreme Court of Victoria, *Practice Note No 2 of 2014* — Expert Evidence in Criminal Trials, 25 June 2014 (Reissued on 30 January 2017 and replaced by Supreme Court of Victoria, Practice Note SC CR 3: Expert Evidence in Criminal Trials, 30 January 2017).

[86] The availability of this procedure is now underpinned by the *Criminal Procedure Act 2009* (Vic) s 232A.

[87] See, eg, *Gilham v R* [2012] NSWCCA 131, [271].

[88] *Director of Public Prosecutions (Vic) v Borg* (2016) 258 A Crim R 172; [2016] VSCA 53.

[89] *Director of Public Prosecutions (Vic) v Borg* (2016) 258 A Crim R 172, [67]–[70]; [2016] VSCA 53.