

Keelah E G Williams and Michael J Saks***

WHY DON'T THE GATEKEEPERS GUARD THE GATES? COMMENTS PROMPTED BY EDMOND

ABSTRACT

Despite clear rules and procedures directing judges to critically scrutinise proffers of expert testimony – and giving them the tools to do so – judges continue to allow unreliable expert evidence into the courtroom. What accounts for the gatekeepers' failure to guard the gates? Drawing from psychological literature, we offer possible explanations for neglectful judicial gatekeeping, and discuss potential solutions for the path forward.

I INTRODUCTION

At bottom, the touchstone of admission of expert evidence has always been the soundness of the expertise upon which proffered expert testimony claims to stand. In past centuries, courts looked for such assurance in the credentials of the proffered witness, then to the commercial success of that witness in his or her field, then to the general acceptance of the claimed knowledge and techniques in the eyes of the relevant fields themselves. Most recently, in the federal and many state courts of the United States, the test has become – for asserted scientific expert testimony – an assessment of the validity of the expert's field's claims based on empirical testing. In the case of *Daubert v Merrell Dow Pharmaceuticals Inc*,¹ the US Supreme Court stated: '[t]he overarching subject is the scientific validity – and thus the evidentiary relevance and reliability – of the principles that underlie the proposed submission.'² That is, the law in many US jurisdictions has directed judges to stop looking at indirect, superficial indicia of soundness, and instead evaluate more directly the actual (empirical, scientific) evidence upon which the (trial) expert evidence stands.

Despite such clear commands from on high, trial judges (the principal gatekeepers of expert evidence) have, for the most part, been nothing less than inspired in evading the task of applying the law's filters to so-called forensic science in criminal cases. When presented with support for the claims of forensic science expert testimony

* JD/PhD candidate, Arizona State University.

** Regents Professor, Sandra Day O'Connor College of Law and Department of Psychology, Arizona State University.

¹ 509 US 579 (1993) ('*Daubert*').

² *Ibid* 594–95.

that was far too weak (and for some fields, non-existent) to cross the legal threshold, the courts engaged in remarkable judicial gymnastics: relegate all criticisms to weight, not admissibility; lower the standard to ensure that the proffer can pass over it; emphasise flexibility of criteria (without stating what criteria were being used); look to a past century of admission (under different or no admissibility rules); appeal to the seemingly intuitively obvious; deny hearings so that the proponent's expert evidence will not have to be put to the test; shift the burden of persuasion from the proponent to the opponent; and, if necessary, extract support from thin air.

Upon analysing a large number of modern cases in which the admissibility of fingerprint expert testimony had been challenged and (almost always) admitted, a major treatise on scientific evidence in American courts described what was found:

Ironically, the failure of judges to write a coherent defense of asserted fingerprint expertise under *Daubert*, but only to seek ways to shelter it from serious scrutiny, suggests that fingerprint expert evidence actually does not meet the requirements of *Daubert*.

If the claims and assumptions of fingerprint identification expertise had been empirically tested, if these empirical tests were sufficiently well designed so as to survive peer review leading to publication in scientifically respectable journals and had [fared well] in the intellectual marketplace following publication, and the data convincingly showed low error rates for the relevant task-at-hand, and if these findings had come to be generally accepted among relevant scientific and professional communities beyond the circle of police technicians who practice the art – then the proponents no doubt would have eagerly offered such information to the courts and the judges would have had ample material with which to write cogent opinions. That such material appears in none of the opinions suggests that it does not exist. If the grounds for admitting fingerprint examiners' testimony were as strong and as sound as the judges assert that it is, then it should not be so difficult to write an opinion actually presenting those grounds.³

The question we tackle in this response to the Edmond essay is why courts behave in this manner. Why do they not simply follow the law, rule as required by the law when applied to the facts, trust the legal process, and embrace their role within that process?

Gary Edmond, in his essay 'What Lawyers Should Know About the Forensic "Sciences"',⁴ provides a review of inquiries into the forensic sciences across the globe, and identifies extensive and persistent problems with forensic science evidence.⁵ Edmond is careful to refer to the forensic sciences in plural, emphasising

³ David L Faigman et al, *Modern Scientific Evidence: The Law and Science of Expert Testimony* (Thomson Reuters, first published 1997, 2010 ed) vol 4, 381-382 § 33:17.

⁴ Gary Edmond, 'What Lawyers Should Know About the Forensic "Sciences"' (2015) 36 *Adelaide Law Review* 33.

⁵ *Ibid.*

the variability in history, reliability, methodology and scope of the many techniques included under the forensic science umbrella.⁶

Ultimately, Edmond concludes that many forensic techniques are insufficiently reliable, due largely to a lack of necessary scientific foundations and having no scientific infrastructure to carry out the research needed to fill the gap.⁷ After a thorough examination of the flaws highlighted by various independent reports, Edmond considers the question of where to go from here. Edmond suggests that continuing on with the now century-old practice of inaction ‘threatens the legitimacy of traditional criminal justice institutions.’⁸ However, Edmond acknowledges that the path forward is unclear. As noted by a comprehensive report on the state of forensic science reviewed in Edmond’s essay, the legal system has shown itself to be ‘ill-equipped to correct the problems of the forensic science community.’⁹ Edmond implicates forensic scientists for their failure to bring the inadequacies of their endeavours to the attention of forensic science consumers – lawyers, judges, and jurors.¹⁰ Such silence in the face of valid criticism is, Edmond suggests, ‘nothing short of scandalous.’¹¹

Although Edmond’s concerns are centred around the shortcomings of the forensic sciences, they imply complementary deficiencies in our courts. Were courts capably performing the job of evidence filtering that the law has entrusted to them, much of the flawed, unvalidated, exaggerated, and speculative expert opinion that has come to be routinely admitted would long ago have been cabined in appropriate ways, and some of it excluded altogether. Instead, those scientifically questionable types of expert opinion have been barely scrutinised and carelessly welcomed into the courthouse. Why do the gatekeepers fail to guard the gates?

That the gatekeepers fail where the forensic ‘sciences’ are concerned is made clear by the comments of a gatekeeper himself: Judge Harry T Edwards, Senior Circuit Judge and Chief Judge Emeritus of the United States Court of Appeals for the DC Circuit. Judge Edwards acted as Co-Chair of a committee tasked by the US Congress in 2006 with reviewing the forensic sciences.¹² Judge Edwards notes that he began the project

⁶ Ibid 34. See also National Research Council, *Strengthening Forensic Science in the United States: A Path Forward* (Washington, DC: National Academies Press, 2009) 38 (*‘NRC Report’*).

⁷ Edmond, above n 4, 81. This centrally important conclusion is shared by the National Research Council. See *NRC Report*, above n 6. It might be more precise to say: Proffered forensic science expert evidence is not known to be sound, valid, or dependable. In the larger world as well as in court, the burden is properly on the proponent to establish that it is sound. Some of it doubtless will be confirmed to be sound, while some of it will be found to be prone to error or be misleading.

⁸ Edmond, above n 4, 35.

⁹ *NRC Report*, above n 6, 53.

¹⁰ Edmond, above n 4, 69.

¹¹ Ibid.

¹² Ibid 37.

'with no preconceived views about the forensic science community,' assuming that 'forensic science disciplines ... are well-grounded in scientific methodology.'¹³ Judge Edwards was forced, however, to conclude that

[t]he bottom line is simple: In a number of forensic science disciplines, forensic science professionals have yet to establish either the validity of their approach or the accuracy of their conclusions, and the courts have been utterly ineffective in addressing this problem.¹⁴

In this Comment we focus our attention on the indirect light that Edmond's essay casts upon the courts. Despite rules and procedures designed to filter expert evidence with greater care than any other kind of proffered evidence, it is clear that courts nevertheless continue to promiscuously admit unreliable forensic science expert evidence. What accounts for the failings of courts in their evaluations of crime laboratory offerings?

Part II of this Comment briefly addresses the appropriate metric for evaluating forensic expert evidence in the United States court system,¹⁵ and provides support for the claim that the courts are failing their assigned role as gatekeeper. Part III addresses possible explanations for neglectful judicial gatekeeping. Potential solutions are discussed in Part IV, with attention paid to the successful exclusion of inadequate forensic expert evidence in a recent court opinion. We then offer general conclusions and comments on the path forward.

II *DAUBERT*: THE PINHOLE THAT BECAME A FLOODGATE

To understand how the courts are failing, we first briefly review the appropriate legal standard for admissibility of forensic expert evidence. We then examine the courts' persistently poor performance in meeting this standard when forensic science has been proffered (typically by the government).

A The Admissibility of Scientific Evidence Under Daubert

The US Supreme Court, for the first time in its history, considered the standard for evaluating the admissibility of scientific expert testimony in *Daubert*.¹⁶ Prior to *Daubert*, American courts had applied several tests, the most prominent of which

¹³ Harry T Edwards, 'Solving the Problems That Plague the Forensic Science Community' (2009) 50 *Jurimetrics Journal* 5, 8.

¹⁴ *NRC Report*, above n 6, 53; Edmond, above n 4, 33 (emphasis added).

¹⁵ The system with which the authors of this Comment are most conversant.

¹⁶ For the purposes of this article, only a concise summary is provided here. Please see David L Faigman et al, *Modern Scientific Evidence: The Law and Science of Expert Testimony* (Thomson Reuters, first published 1997, 2008 ed) vol 1, ch 1, 1-120, ('*Modern Scientific Evidence*') for a thorough review of scientific evidence admissibility in the United States.

(and the greatest contrast to *Daubert*) was the ‘general acceptance’ test¹⁷ announced in *Frye v United States*.¹⁸ In such a test, judges need not make independent assessments of validity, but can defer to what asserted scientists say about their respective fields’ views of their techniques and theories.¹⁹

In *Daubert*, however, the Court unanimously held that trial court judges must embrace their role as ‘gatekeepers’ by evaluating the validity of the basis for proffered scientific expertise before permitting the expert to testify.²⁰ No longer passive bystanders watching expert evidence flow to fact-finders, *Daubert* requires judges to actively assess the purported science underlying the asserted expert testimony. The Court concluded that the evidentiary relevance and reliability of such testimony could be determined only by an evaluation of scientific validity.²¹ Thus, under *Daubert* and the Federal Rules of Evidence, trial court judges are responsible for determining the admissibility of forensic expert evidence,²² and must do so by evaluating the soundness of the science supporting the proffered testimony.

Application of the *Daubert* standard to forensic science requires an understanding of scientific research. This presents judges with a challenge: as non-experts themselves, how can judges evaluate the scientific validity of forensic expert evidence?²³ Chief Justice Rhenquist anticipated the difficulty judges would have in assuming the role of ‘amateur scientists.’²⁴ To assist trial court judges with this demanding task, the Court in *Daubert* articulated several factors for consideration when determining the validity

¹⁷ Roselle L Wissler, Keelah E G Williams and Michael J Saks, ‘Dual-Processing Models of Admissibility: How Legal Tests for the Admissibility of Scientific Evidence Resemble Cognitive Science’s System 1 and System 2’ (2013) 17 *Virginia Journal of Law & Technology* 356, 363.

¹⁸ 293 F 1013 (DC Cir, 1923) (*Frye*). For a discussion of *Frye*’s predecessor, a ‘market-place’ test in which experts were judged by their success in the commercial market, see Faigman et al, above n 16, vi.

¹⁹ Faigman et al, above n 16, 15.

²⁰ 509 US 579, 597 (1993).

²¹ Ibid 594–95 (the ‘overarching subject [of Rule 702] is the scientific validity – and thus the evidentiary relevance and reliability – of the principles that underlie the proposed submission.’).

²² Federal Rules of Evidence, Rule 702, is the rule that governs admissibility, and it states:

a witness who is qualified as an expert by knowledge, skill, experience, training, or education may testify in the form of an opinion or otherwise if:

- (a) the expert’s scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;
- (b) the testimony is based on sufficient facts or data;
- (c) the testimony is a product of reliable principles and methods; and
- (d) the expert has reliably applied the principles and methods to the facts of the case.

²³ Roselle L Wissler, Keelah E G Williams and Michael J Saks, above n 17, 357.

²⁴ *Daubert* 509 US 579, 601 (1993).

of scientific expert opinion: testability ('falsifiability'),²⁵ quality of research design and methods ('peer review and publication'),²⁶ findings of the research ('error rates'),²⁷ 'maintenance of standards'²⁸ in the application of the research knowledge to the trial issue, and, in a greatly reduced role, 'general acceptance.'²⁹ The Court emphasised that these factors are not exclusive,³⁰ and each need not be met in all cases.³¹ Some scholars, however, have neatly summarised the task into a single directive: *Daubert* obliges judges to ask, 'Where are the data?'³²

As conceptually straightforward as the US Supreme Court's directive might seem to a conventional scientist, carrying it out requires sufficient scientific proficiency to not only inspect the scientific principles and findings underlying experts' opinions, but also to understand and evaluate the methodologies behind particular results. 'Data' are produced by both rigorous research and parodies of science. In order to evaluate the validity of scientific evidence, judges must be able to distinguish one from the other.³³ What two decades of cases and research have made clear, however, is that judges have not successfully embraced the *Daubert* directive where forensic science has been concerned.

B *How Are Judges Failing?*

The failure of judges to subject scientific expert evidence to the level of scrutiny required by *Daubert* is systemic. It is not the occasional judge who errs, but the great majority erring the great majority of the time when called upon to evaluate proffers of forensic science. Continued admission of unvalidated scientific evidence into the courtroom suggests that judges have a quasi-religious faith in the validity of the non-science forensic sciences, the soundness of their techniques, and the accuracy of resulting expert opinions.³⁴ These erroneous beliefs persist notwithstanding the

²⁵ Ibid 593.

²⁶ Ibid.

²⁷ Ibid 594.

²⁸ Ibid.

²⁹ Ibid. See also Faigman et al, above n 16, 30.

³⁰ See Mark McCormick, 'Scientific Evidence: Defining a New Approach to Admissibility' (1982) 67 *Iowa Law Review* 879, 911–12, for a former state supreme court justice's suggested list of additional factors for determining admissibility (including, for example, the care with which the technique was employed in the case).

³¹ 509 US 579, 593 (1993) (factors not 'a definitive checklist or test').

³² Faigman et al, above n 16, 47.

³³ Ibid 37.

³⁴ See the observations of Judge Thorne, concurring in *State v Quintana*, 103 P 3d 168, 170, 171 (Utah Ct App, 2004): 'fingerprint [identification] evidence has never truly been put to the test in either the courtroom or the scientific community,' that '[i]n essence, we have adopted a cultural assumption that a government representative's assertion that a defendant's fingerprint was found at a crime scene is an infallible

inadequate research base,³⁵ the failure to develop and disclose meaningful error rates,³⁶ the failure to develop and employ assessments of random match probabilities (or similar probabilistic methods) based on data,³⁷ exaggerated claims that exceed what is known or can be known,³⁸ and the unwillingness to shield analysts from domain-irrelevant information.³⁹

If, at the very least, *Daubert* requires that testable evidence be tested (and to require that the proponent of the expert opinion make the case for admission with those data), it is clear that several forensic sciences fail to meet this basic threshold.⁴⁰ For example, researchers have noted that handwriting identification, bite marks, and hair comparisons lack sufficient empirical evidence for admission under a serious application of *Daubert*.⁴¹ Those forensic sciences that have disappeared from the courts owing to findings of their lack of validity (eg, voiceprints, numerous arson indicators, comparative bullet lead analysis) were not generally (or at all) perceived as flawed by the courts, and their departure was not hastened by any judicial push out the door.

The National Research Council Report details many of the systemic problems with forensic evidence.⁴² The report's concluding call for the creation of a National Institute of Forensic Science was an acknowledgement of the failure of all relevant extant institutions to have built a sound and scientific forensic science and to keep pseudo-science from entering the courts. Those institutions that failed in this regard

fact, and not merely the examiner's opinion', and therefore courts should take steps to 'remove the near mystical awe that fingerprints evoke, and replace it with a more cautious regard for forensic evidence and its overall lack of certainty.'

³⁵ Faigman et al, above n 16, 63; *NRC Report*, above n 6, 189; Edmond, above n 4, 38–9.

³⁶ *NRC Report*, above n 6, 122.

³⁷ See, eg, Michael J Saks and Jonathan J Koehler, 'The Individualization Fallacy in Forensic Science Evidence' (2008) 61 *Vanderbilt Law Review* 199 (2008) ('*Individualization Fallacy*'); Michael J Saks and Jonathan J Koehler, 'The Coming Paradigm Shift in Forensic Identification Science' (2005) 309 *Science* 892 ('*Paradigm Shift*').

³⁸ See, eg, Mark Page, Jane Taylor and Matt Blenkin, 'Uniqueness in the forensic identification sciences – Fact or fiction?' (2011) 206 *Forensic Science International* 12; Edmond, above n 4, 46; Saks and Koehler, *Individualization Fallacy*, above n 37.

³⁹ Edmond, above n 4, 53. See also D Michael Risinger et al, 'The *Daubert/Kumho* Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion' (2002) 90 *California Law Review* 1.

⁴⁰ Faigman et al, above n 16, 63.

⁴¹ *Ibid.* See also Michael J Saks, 'Implications of the *Daubert* Test for Forensic Identification Science' (1994) 1 *Shepard's Expert & Scientific Evidence Quarterly* 427; D Michael Risinger and Michael J Saks, 'Science and Nonscience in the Courts: *Daubert* Meets Handwriting Identification Expertise' (1996) 82 *Iowa Law Review* 21; Randolph N Jonakait, 'Real Science and Forensic Science' (1994) 1 *Shepard's Expert & Scientific Evidence Quarterly* 435, 441; Saks and Koehler, *Paradigm Shift*, above n 37.

⁴² *NRC Report*, above n 6.

include the courts. When it comes to forensic evidence, what might account for the courts' persistently neglectful gatekeeping?

III POSSIBLE JUDICIAL WORLD VIEW THAT MIGHT EXPLAIN POROUS GATEKEEPING

Courts routinely admit forensic science expert testimony that has not been validated using any sort of conventional empirical testing. Why? A multitude of explanations have been offered for this uncritical admission of unvalidated techniques into the courtroom, so long as the proffers are flying under the banner of forensic science.⁴³ Although – as suggested by the *NRC Report*⁴⁴ – the loci of this failure are numerous, the final institutional quality control is the judicial admissibility decision. If courts screened conscientiously and competently, excluding that which had not been validated, they would create a powerful incentive for governments to invest in validity testing.⁴⁵ Were judges to demand the data, the forensic science community would be obliged to begin producing the data. Thereafter, courts would be in a better position to admit or exclude based on the results of such testing. At present, however, a variety of impediments keep the courts from embracing the role required of them by their rules of evidence, the common law, and certainly by *Daubert*.

A Conventionality

In an attempt to maintain the authority of their decisions and to shelter them from criticism, trial judges are likely to look to prevailing sentiment in their jurisdictions – not only formal precedent but popular beliefs, held by other judges and by the general public.⁴⁶ Judges feel they are on safe ground as long as they adhere to conventional beliefs.⁴⁷ This might be thought of as a companion – and sometimes a competitor – to precedent: a tendency to follow convention rather than deference to controlling law in accord with *stare decisis*. Consistency with law might require doing exactly what most judges are disinclined to do: evaluate old forensic techniques under unfamiliar

⁴³ See, eg, Edmond, above n 4, 81.

⁴⁴ *NRC Report*, above n 6, 85.

⁴⁵ Though sound research is necessary and welcome, at the same time the findings need to be absorbed by the law with a cautious and sophisticated appreciation that research can be designed – intentionally or inadvertently – in ways that produce illusory support as well as illusory disconfirmation. See, eg, D Michael Risinger and Michael J Saks, 'A House with No Foundation: Forensic science needs to build a base of rigorous research to establish its reliability' (2003) 20 *Issues In Science And Technology* 35.

⁴⁶ Faigman et al, above n 16, 51.

⁴⁷ One of the authors was present in a California courtroom when a judge stated during an evidentiary hearing: I don't want to be the first judge in a hundred years to conclude that [a particular forensic science] does not meet the test for admission. That judge was under no legal obligation to admit the proffered expert testimony without limitation or, indeed, to admit it at all, yet felt a compulsion not to depart from conventional practice regardless of what the applicable legal rules might say.

‘new’ standards.⁴⁸ What we have described in this subsection is an illustration of the psychological phenomenon of ‘social proof.’⁴⁹ Social proof leads individuals to conform to actions or ideas perceived as normative, predominantly when the ‘correct’ behaviour is ambiguous. Particularly in the case of forensic sciences with a long history of courtroom presence, it might be difficult for judges to question what has for so long been accepted without question. Deference to convention inevitably preserves the so-called forensic sciences in whatever state they have been operating.

B *Inertia*

A related notion is that of inertia. But whereas conventionality implies following popular beliefs set by current members of the bar, the bench, and the public, inertia implies continuing to do whatever one (or one’s peers) have done before. Despite numerous high level reports exposing serious shortcomings in forensic expert evidence, courts proceed as if nothing has changed in the understanding of the attainments and limits of forensic science knowledge.

Perhaps much of this judicial inertia can be explained by a wayward reliance on history. Courts have relied on certain techniques for so long that they are reluctant to rethink the role of those techniques in the trial process. Similar to the concept of ‘grandfathering’, courts continue to admit familiar species of evidence, regardless of new rules that impose new tests, or old rules that courts are urged to start applying properly. Thus, the sheer fact of a long history of admission carries enormous weight. However much sense this makes for dealing with normative principles in law, it is precisely backwards when dealing with science. Past generations of judges had less data available to them than today’s judges do. Judges today are in a far better position to assess the ability of forensic science expert testimony to do what its practitioners claim or to meet the law’s standards for admission. And tomorrow’s judges will be in a better position than today’s. But the judicial preference for repeating whatever helps to lock forensic science in place. The forensic scientists, too, know that if they keep doing what they have been doing they are likely to remain safe in the courts’ eyes.

C *Error Management in Admission of Expert Evidence*

The purpose of evidence screening by judges is to assist fact-finders in reaching correct conclusions by providing more probative evidence and screening out

⁴⁸ As one judge put it: under the *Daubert* line of cases, ‘[e]verything old is new again.’ (*United States v Horn*, 185 F Supp 2d 530, 554 (D Md, 2002)). Many other judges, of course, refused to read *Daubert* as commanding them to ‘depart from the well-traveled path,’ despite their inability to find justification for admission under the criteria provided by the *Daubert* trilogy. *US v Cline*, 188 F Supp 2d 1287 (D Kan, 2002), affd 349 F 3d 1276 (10th Cir, 2003). Such recalcitrance would be viewed, at least by Supreme Court Justice Antonin Scalia, as reversible error. See *Kumho Tire Co Ltd v Carmichael*, 526 US 137 (Scalia J concurring) (1999).

⁴⁹ Robert B Cialdini, *Influence: The Psychology of Persuasion* (Harper Business, revised ed, 2006) 114.

misleading evidence. As in many other settings, courtroom verdicts can produce one of two errors: a false negative (in the trial context, a guilty individual pronounced not guilty), or a false positive (the alternative error, in which an innocent individual is found guilty). Because errors are unavoidable, procedures must be calibrated to balance the risk of false positives and false negatives. Once that threshold is set – in criminal cases, it is the familiar ‘proof beyond a reasonable doubt’ – judges might implicitly adjust their evidence screening thresholds to help or hinder the government’s efforts to carry a case over the threshold.

The US Supreme Court in *Daubert* recognised that ‘a gatekeeping role ... inevitably on occasion will prevent the jury from learning of authentic insights’.⁵⁰ Excluding forensic expert evidence might thus increase the risk of a false negative, and so it is not surprising that judges who believe a conviction is the preferred outcome in a particular criminal case – or in criminal cases generally – might exhibit trepidation about ‘defanging prosecutors.’⁵¹ The result would be that, when evaluating expert evidence in a criminal case, anything that tends to inculpate a defendant tends to be admissible.

As a consequence, tacit trial verdict error management practices might be affecting the evaluation of forensic science in a way that unintentionally increases the risk of admitting misleading evidence. Some indication that this is occurring comes from studies of DNA exoneration cases, which have found erroneous and misleading forensic science testimony to be a leading cause of erroneous convictions, second only to eyewitness errors.⁵²

D Failure to Recognise Pseudo-Science

When testifying before the Goudge Inquiry, LeSage J stated:

I must say it came as somewhat of a shock to me, having spent forty (40) years plus in the justice system, to hear some of the scientific experts speaking about the uncertainty and the lack of clarity in areas of science which I had always thought were far more certain than they really are. And I felt very guilty that I had not better educated myself on these areas long before.⁵³

Justice LeSage is surely in crowded company. Few judges or lawyers understand how science comprehends empirical reality and, as a result, few are able to recognise the pre-science nature of certain forensic ‘sciences’. Judges’ most typical evidentiary experience involves individual, narrative storytelling by ordinary fact witnesses,

⁵⁰ 509 US 579, 597 (1993).

⁵¹ Faigman et al, above n 16, 64.

⁵² Saks and Koehler, *Paradigm Shift*, above n 37, 892. See also Brandon L Garrett and Peter J Neufeld, ‘Invalid Forensic Science Testimony and Wrongful Convictions’ (2009) 95 *Virginia Law Review* 1.

⁵³ Stephen T Goudge, *Inquiry into Pediatric Forensic Pathology in Ontario* (Queen’s Printer, 2008) 501.

rather than an assessment of the scientific validity underlying empirical data. Not only are judges untrained in evaluating scientific, empirical issues, they tend to be unaware of the extent of their limitations.⁵⁴

Because few judges understand how science builds knowledge, they appear to assume that once scientific expert evidence has been admitted as valid, it cannot become invalid. In consequence, what has been admissible cannot become inadmissible. But it is not a rare experience in any scientific field to see scientific phenomena – be they empirical relationships or theoretical understandings – that have had to be revised or revoked as new and better studies exposed flaws in the prior studies.⁵⁵

Ironically, progress in science that contradicts old science is the unrecognised enemy of past judicial action. Judges, like almost everyone in Western society, glibly declare that they welcome scientific progress. But scientific progress is not welcome when it advances by subtraction – that is, by finding that something once believed to be true is instead false. One of the authors observed an exchange that ensued following a continuing education program presented by university scientists to the state's judges. A medical school gynaecology professor had explained to the judicial audience how she and her colleagues once believed that certain markings on a child's vagina signified abuse, but more recent and better research revealed that entirely innocent and common causes for those markings existed. Following her talk, a judge approached and with apparent annoyance made essentially the following statement: 'Do you remember testifying in my courtroom five years ago, and saying that marks such as those did indicate sexual abuse? That child's father is now in prison because of your testimony.' Even more interesting was the doctor's reply: 'What would you have us do, remain ignorant forever?'

This anecdote illustrates the tension between judges and scientific experts when science marches onward by replacing earlier erroneous beliefs with updated findings. The law is not prepared to handle such tensions. It has unrealistic expectations of what scientific knowledge should do for trials. And, much like forensic scientists themselves, judges prefer to protect old beliefs they and their brethren have acted upon from refutation.

⁵⁴ See, eg, Sophia Gatowski et al, 'Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-*Daubert* World' (2001) 25 *Law and Human Behavior* 433; Margaret Bull Kovera et al, 'Assessment of the Commonsense Psychology Underlying *Daubert*: Legal Decision Makers' Abilities to Evaluate Expert Evidence in Hostile Work Environment Cases' (2002) 8 *Psychology, Public Policy, and Law* 180. This is not surprising; it is difficult to know what one does not know and easy to think a matter simpler than it actually is. To turn the microscope around, think of the numerous mistaken things that doctors and scientists and other legally untrained persons think they understand about the law.

⁵⁵ See, eg, Mario Livio, *Brilliant Blunders* (Simon & Schuster, 2013).

E Over-Reliance on the Adversarial Process

Judges and scientists do not share the same concept of ‘reliability,’ and thus possess a different metric for achieving it. As demonstrated by Blackmun J in the US Supreme Court’s opinion in *Daubert*, many judges seem to believe that ‘vigorous cross-examination, presentation of contrary evidence, and careful instruction on the burden of proof are the traditional and appropriate means of attacking shaky but admissible evidence.’⁵⁶ Put more simply, this signifies a belief that trial procedure is capable of separating good scientific evidence from bad scientific evidence. ‘Shaky’ forensic science evidence will be exposed during trial through the safeguards of the adversarial process. Thus, judges may safely play a passive role in the filtering of evidence, trusting that the ‘truth will out’.

Such confidence is misplaced for several reasons. Defence attorneys often do not have the resources to challenge prosecutors’ forensic experts. Even with adequate resources, a defence attorney’s best challenge against forensic science testimony would be empirical data indicating the science is flawed – and yet, as we know, empirical inquiry into many forensic science techniques remains limited, methodologically inadequate, or non-existent. Finally, forensic science evidence can be overly persuasive to juries, regardless of the shaky scientific basis for such belief – and cross-examination itself is an ineffective neutraliser of unreliable expert testimony.⁵⁷

Before the trial begins, trial judges (sitting alone) must decide evidentiary issues with few resources and little time for extensive research and reflection. After conclusion of the trial, there exists very limited appellate review of trial court rulings admitting disputed evidence. Moreover, because those are evidentiary rulings, appellate courts traditionally review them deferentially.⁵⁸ Thus, from before the trial begins to well after the trial concludes, admission of weak forensic evidence eludes the ability of the adversarial system to effectively evaluate the proffered (or admitted) evidence. When faced with the unpleasant consequences produced by poor quality forensic science – notably, wrongful convictions – courts have a tendency to blame individual experts rather than recognise systemic issues. By doing so, courts ‘sidestep engaging with the possibility that legal processes might create systematic vulnerabilities to unreliable and speculative forms of expert opinion evidence.’⁵⁹

⁵⁶ 509 US 579, 596 (1993).

⁵⁷ See, eg, Joseph Sanders, ‘The Merits of the Paternalistic Justification for Restrictions’ (2003) 33 *Seton Hall Law Review* 881, 931–935. See also Shari Seidman Diamond et al, ‘Juror Reactions to Attorneys at Trial’ (1996) 87 *Journal of Criminal Law & Criminology* 17, 41 (results indicating that once a forensic psychiatrist has made a prediction of dangerousness, no amount of high quality challenge can undo the damage).

⁵⁸ See, eg, *General Electric Co v Joiner*, 522 US 136 (1997).

⁵⁹ Gary Edmond et al, ‘Admissibility Compared: The Reception of Incriminating Expert Evidence (i.e., Forensic Science) in Four Adversarial Jurisdictions’ (2013) 3 *University of Denver Criminal Law Review* 31, 93.

We have suggested a number of cognitive habits of Anglo-American judges that together, or separately, might explain the continuing admission by courts of expert testimony from witnesses representing subfields of the non-science forensic sciences – fields that have yet to conduct rigorous research to evaluate the validity of their techniques, despite calls from several high-level commissions of inquiry that have looked into these fields and found them to be surprisingly short on science.⁶⁰ Can the future be expected to offer much improvement over the past?

IV LOOKING TO THE FUTURE

Numerous voices have detailed serious deficiencies of forensic science expert testimony, yet the courts' response thus far has been, shall we say, unresponsive. However, significant improvements are within reach. In Part IV we briefly mention potential solutions, and call attention to a recent court opinion excluding inadequate forensic expert evidence as an indication that judicial progress is attainable, even while scientific progress is slow.

A Helpful Half-Measures

As scholars such as Edmond have indicated,⁶¹ the forensic science community has much to answer for in their failure to improve the science behind forensic expert evidence. Peter Neufeld has argued that the courts are not the place where improvements will be made, and that the focus of attention must be on other institutions that can compel or encourage improvements in the forensic science fields.⁶² Certainly, courts cannot conduct the scientific experimentation required to validate forensic science techniques. However, courts have great power to catalyse the necessary changes. By not holding forensic scientists accountable, the courts allow these troubles to continue. And by contrast, if the gatekeepers were to more assiduously 'guard the gates,' that would force the forensic sciences (or others) to do the science and produce the evidence needed to permit the gatekeepers to conduct serious substantive evaluations. Thus, courts could rapidly improve the quality of forensic science by refusing to admit poorly scientifically supported, highly exaggerated testimony. The courts cannot be responsible for improving the science, per se, but they can, and should, be responsible for following the requirements of the law. This

⁶⁰ Edmond, above n 4, 34. See also Paul C Giannelli, 'Forensic Science: Why No Research?' (2010) 38(2) *Fordham Urban Law Journal* 503, 517 ('the reason for the lack of empirical research [in the past decade] was simply a stubborn refusal to reconsider beliefs in light of credible challenges. This is the antithesis of the scientific method.') Lack of research is only one of the major criticisms from the various commissions. But, validation research is the indispensable starting point – without it, little else can improve.

⁶¹ Edmond, above n 4, 69.

⁶² See, eg, Peter Neufeld and Barry Scheck, 'Making Forensic Science More Scientific' (2010) 464 *Nature* 351.

in turn would impel improvements by those whose interest it is to keep forensic science expert evidence flowing into the courts.

Beyond more careful scrutiny of proffered forensic science expert evidence – or instead of more careful scrutiny – courts could take other steps to rein in unvalidated forensic science. In this Comment, we merely list them, but further explication can be found in the margin.⁶³ None of these is a substitute for sound validation, rigorous quality control oversight, and more innovative institutional fixes,⁶⁴ but for courts and justice systems that for whatever reason can do no more, here are some possible courses of action:

- As a condition of admission, require examiners to be certified and their laboratories to be accredited.
- As a condition of admission, require examiners to have participated in regular proficiency testing, and review their performance results (aware that sometimes the tests are so easy that it would be hard not to do well).
- As a condition of admission, require labs to have submitted to routine scientific audits.
- Make use of court appointed experts as one route for reducing biased or weak forensic evidence in the courtroom.⁶⁵
- Require blind examinations, evidence lineups, sequential unmasking, or other recognised procedures to minimise unintended bias in examinations.⁶⁶
- Employ partial admission – allowing pattern comparison examiners to describe similarities and differences between questioned and known samples, but not to opine on ultimate conclusions of identity.⁶⁷
- Require experts to stay well within the bounds of their asserted expertise. That is, to not go beyond what their field claims to be able to do.⁶⁸
- Prohibit assertions of unique individualisation (for fields that perform pattern comparison) or assertions of perfect or near-perfect accuracy (for all fields). Prohibit use of unreasonably overpowering terminology generally, such as ‘indeed and without doubt’, ‘match’, ‘share a common source’, ‘identification to the exclusion of all others in the world’.

⁶³ See, eg, Michael J Saks, ‘Protecting Factfinders from Being Overly Misled, While Still Admitting Weakly Supported Forensic Science into Evidence’ (2008) 43 *Tulsa Law Review* 609.

⁶⁴ See Roger Koppl, *CSI For Real: How to Improve Forensics Science* (1 December 2007) Reason Foundation <<http://reason.org/files/d834fab5860d5cf4b3949fecf86d3328.pdf>> 19–30.

⁶⁵ *General Electric Co v Joiner*, 522 US 136, 147 (Breyer J concurring) (1997).

⁶⁶ Risinger et al, above n 39; Dan E Krane et al, ‘Sequential Unmasking: A Means of Minimizing Observer Effects in Forensic DNA Interpretation’ (2008) 53 *Journal of Forensic Sciences* 1006.

⁶⁷ See, eg, *United States v Hines*, 55 F Supp 2d 62 (D Mass, 1999).

⁶⁸ Even when the field’s claims have not been validated, one can insist that the expert not become a field unto himself of unvalidated claims of ability.

- Instruct the jury on the limitations of accuracy of particular types of expertise, even though the court has chosen to admit such testimony.⁶⁹

B *An Exception to the Rule*

A recent court opinion demonstrates that it is possible for judges to undertake more searching evaluations of proffered forensic science, and might serve as a model, or at least inspiration, for proper consideration of scientific evidence.⁷⁰ The trial judge in *United States v Johnsted* evaluated a motion to exclude the expert testimony of a forensic document examiner on handwriting identification, ultimately concluding that ‘the science or art underlying handwriting analysis falls well short of a reliability threshold when applied to hand printing analysis,’⁷¹ having found that the government had not sufficiently demonstrated that the expert’s analysis was ‘supported by principles and methodology that are scientifically valid.’⁷²

The Court noted several troubling aspects of the proffered testimony. First, little research existed ‘demonstrating the ability of certified experts to distinguish between [an] individual’s handwriting or identity forgeries to any reliable degree of certainty.’⁷³ Of the limited research available, results indicated that forensic experts ‘made correct identifications less frequently than laypersons in hand printing analyses.’⁷⁴

Second, the court commented on the discretionary nature of the expert’s analysis. The expert conceded that, in contrast to objective criteria, ‘analysts must rely entirely on their experiences and individual training to determine when a case warrants a particular conclusion.’⁷⁵

Third, the judge was unimpressed by the proffered published peer reviewed studies, stating, ‘a mere list of journals does not convince the court that the specific techniques at issue in this case have been peer reviewed.’⁷⁶ A consistency-based argument was equally unpersuasive, representing a rare and thoughtful departure from customary judicial treatment of handwriting evidence.⁷⁷

⁶⁹ See, eg, *United States v Starzeczyzel*, 880 F Supp 1027 (SD NY, 1995).

⁷⁰ *United States v Johnsted*, 30 F Supp 3d 814 (WD Wis, 2013) (*Johnsted*) (order excluding expert testimony).

⁷¹ *Ibid.*

⁷² *Ibid.*

⁷³ *Ibid* 818.

⁷⁴ *Ibid* 819.

⁷⁵ *Ibid.*

⁷⁶ *Ibid* 820.

⁷⁷ The full opinion is well worth reading for the thoroughness and acuity of its consideration of the factual issues and its analysis.

The conclusions of the *Johnsted* opinion indicate that change is possible in how judges screen proffers of expert evidence for validity. However, such change requires judges to exert effort, exercise independence, and perhaps summon courage as well.

V CONCLUSIONS

Evaluating scientific and pseudo-scientific evidence is a challenging task for the courts, and perhaps one for which they are ill-suited. Although we may acknowledge the difficulty of the task, the ‘appropriate response to complexity should not be to call in the witch doctor for a magic spell, but rather to demand the best science available and remain aware of its limitations.’⁷⁸ In *General Electric Co v Joiner*,⁷⁹ Breyer J anticipated the trouble judges would face, noting that the gatekeeping requirement ‘will sometimes ask judges to make subtle and sophisticated determinations about scientific methodology and its relation to the conclusions an expert witness seeks to offer.’⁸⁰ Acknowledging that ‘judges are not scientists,’ he nevertheless counseled that ‘neither the difficulty of the task nor any comparative lack of expertise can excuse the judge from exercising the “gatekeeper” duties that the Federal Rules impose.’⁸¹ Justice Breyer’s sentiments were echoed by Scalia J’s concurrence in *Kumho Co Tire Ltd. v Carmichael*,⁸² in which he concluded that a failure to apply the *Daubert* factors where they are applicable would constitute an abuse of discretion and therefore be a reversible error.⁸³

Perhaps the duty to evaluate forensic science evidence confronts courts with a perfect storm, encompassing judicial incapacity (ignorance of how science builds knowledge), lack of will (bias), and pressure to move cases expeditiously (managerial judging). The situation likely is not solvable from within the judicial process. As Peter Neufeld has argued,⁸⁴ much of the change will have to come from outside: by testing and improving the forensic sciences, by engaging with mainstream research scientists, and by legislative and administrative processes. The US National Research Council’s suggestion to create a National Institute of Forensic Science reflects this realisation.

Yet, as the US Supreme Court has observed, judges are not to be absolved of responsibility. And when judges embrace the role required of them by the law, the gates can be better guarded against invasion by the non-science forensic ‘sciences’.

⁷⁸ Faigman et al, above n 16, 62.

⁷⁹ 522 US 136 (1997).

⁸⁰ *General Electric Co v Joiner*, 522 US 136, 146 (Breyer J) (1997).

⁸¹ *Ibid.*

⁸² 526 US 137 (Scalia J) (1999).

⁸³ *Ibid.* 159.

⁸⁴ Neufeld and Scheck, above n 62, 351.