

2010

“UTTERLY INEFFECTIVE”: DO COURTS HAVE A ROLE IN IMPROVING THE QUALITY OF FORENSIC EXPERT TESTIMONY?

Joseph Sanders

Follow this and additional works at: <https://ir.lawnet.fordham.edu/ulj>



Part of the [Science and Technology Law Commons](#)

Recommended Citation

Joseph Sanders, “*UTTERLY INEFFECTIVE*”: *DO COURTS HAVE A ROLE IN IMPROVING THE QUALITY OF FORENSIC EXPERT TESTIMONY?*, 38 Fordham Urb. L.J. 547 (2010).

Available at: <https://ir.lawnet.fordham.edu/ulj/vol38/iss2/7>

This Article is brought to you for free and open access by FLASH: The Fordham Law Archive of Scholarship and History. It has been accepted for inclusion in Fordham Urban Law Journal by an authorized editor of FLASH: The Fordham Law Archive of Scholarship and History. For more information, please contact tmelnick@law.fordham.edu.

“UTTERLY INEFFECTIVE”: DO COURTS HAVE A ROLE IN IMPROVING THE QUALITY OF FORENSIC EXPERT TESTIMONY?

*Joseph Sanders**

Introduction: Forensic Science and the National Research Council Report	547
I. The NRC’s Reasons for Giving the Courts Almost No Role in Improving Forensic Expertise.....	550
II. Why Are Courts So Lenient in Admitting Expert Forensic Evidence?.....	555
A. Saks’ Explanation.....	555
B. Contextual Approach to Knowledge	558
C. Information About the Particular Case	560
III. What Role Might Courts Play?.....	562
Conclusion	569

INTRODUCTION: FORENSIC SCIENCE AND THE NATIONAL RESEARCH COUNCIL REPORT

For at least three decades we have witnessed an academic critique of the quality of forensics evidence as it is employed in criminal prosecutions.¹ Although the critique has proceeded at many levels, at its core is the fact that many forensic science fields have failed to conduct the research necessary to test the reliability and validity of the methods and techniques forensic witnesses employ to support their courtroom testimony in criminal cases. The critique became more vocal and more persistent with the United States Supreme Court admissibility decisions in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,² *General Electric v. Joiner*,³ and *Kumho Tire Co. v. Carmichael*.⁴

* A.A. White Professor of Law, University of Houston Law Center.

1. The argument is as least as old as Professor Giannelli’s insightful critique of the Frye rule in 1980. Paul Giannelli, *The Admissibility of Novel Scientific Evidence: Frye v. United States, A Half Century Later*, 80 COLUM. L. REV. 1197 (1980).

2. 509 U.S. 579, 590 n.9 (1993).

3. 522 U.S. 136 (1997).

4. 526 U.S. 137 (1999).

In the early years following *Daubert*, it sometimes seemed that the forensic expert community simply didn't understand the academic critique.⁵ For example, D. Michael Risinger reports that during a 1996 discussion, document examiners made the following comments: "Only we who do it can know that what we say about it is true," and "[they] went after our weak point: no data."⁶ Later, the expert community developed more elaborate justifications for its practices and its lack of interest in conducting research designed to test the validity and reliability of its conclusions. Michael J. Saks and Jonathan J. Koehler have persuasively argued that at the heart of this set of justifications are the concepts of individualization and uniqueness.⁷ As Simon Cole notes, individualization is understood to mean that it is possible to narrow the potential sources of a forensic trace "to a single object in the universe," and this sort of individualization is itself supported by the assumption that each forensic object (a fingerprint, a spent bullet, a bite-mark, a signature) is unique.⁸ Combined, these assumptions lead to assertions such as the following: "And we profess as fingerprint examiners that the rate of error is zero. And the reason we make that bold statement is because we know based on 100 years of research that everybody's fingerprints are unique, and in nature it [sic] is never going to repeat itself again."⁹ Empirical investigations by outside experts on the ability of forensic experts to make such assertions are claimed to be unnecessary.¹⁰

It is against this backdrop that the National Research Council (NRC) published its Report, *Strengthening Forensic Science in the United States: A Path Forward*. The Report clearly sides with the critics. The Report's summary makes the following observation: "With the exception of nuclear DNA analysis . . . no forensic method has been rigorously shown to have

5. See Andre A. Moenssens, *Handwriting Identification Evidence in the Post-Daubert World*, 66 UMKC L. REV. 251 (1997).

6. D. Michael Risinger, Mark P. Denbeaux & Michael J. Saks, *Brave New "Post-Daubert World"—A Reply to Professor Moenssens*, 29 SETON HALL L. REV. 405, 441 n.108 (1998).

7. See Michael J. Saks & Jonathan J. Koehler, *The Individualization Fallacy in Forensic Science Evidence*, 61 VAND. L. REV. 199 (2008).

8. See Simon Cole, *Forensics Without Uniqueness, Conclusions Without Individualization: The New Epistemology of Forensic Identification*, 8 LAW, PROBABILITY & RISK 233, 235 (2009); D. Michael Risinger & Mark P. Denbeaux, *Kumho Tire and Expert Reliability: How the Question You Ask Gives the Answer You Get*, 34 SETON HALL L. REV. 15 (2003).

9. Cole, *supra* note 8, at 237 (quoting Transcript of Trial, *People v. Gomez*, No. 99CF0391 (Cal. Super. Ct. Orange Cty. 2002)).

10. See Moenssens, *supra* note 5, at 303-04.

the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source.”¹¹

The Report offers a series of recommendations designed to change this state of affairs. Its first recommendation is the establishment of a National Institute of Forensic Science that would be independent from the existing forensics community and would have an advisory board with expertise in multiple disciplines. Its third recommendation is a call for research to assess accuracy, reliability, and validity in the forensic science disciplines.¹²

What is interesting is that the recommendations are neither directed at the courts, nor do they call on the courts to use the admissibility standards developed in *Daubert* and state court analogs to tighten admissibility standards. Certainly, this is not because the courts are already doing a good job in this regard. The Report agrees with many commentators that the courts have not adequately policed forensic experts.¹³ In a rather blunt passage, the Report adopts Peter Neufeld’s assessment that “the courts have been ‘utterly ineffective’ in addressing the problem” of unreliable testimony.¹⁴

In Part I, I review the NRC’s stated reasons for giving the courts little or no role in improving forensic evidence and argue that these reasons cannot explain the fact that the same courts have played a significant role in policing expertise in civil cases. Why then have courts been so reluctant to exclude forensic expert evidence? I explore this question in Part II. I argue that two deep seated factors: (1) the courts’ contextual approach to knowledge, and (2) the limited ability of science to provide causal answers about the particular case, limit the courts’ willingness to raise admissibility standards in forensic cases. In Part III, I argue that courts can play a limited role in improving expert forensic evidence by excluding the worst evidence in each area. I believe that to date they have not done so consistently in

11. NAT’L RES. COUNCIL, NAT’L ACADEMY OF SCI., STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 7 (2009) [hereinafter STRENGTHENING FORENSIC SCIENCE].

12. *Id.* at 22.

13. See Jennifer L. Groscup et al., *The Effects of Daubert on the Admissibility of Expert Testimony in State and Federal Criminal Cases*, 8 PSYCHOL. PUB. POL’Y & L. 339, 364 (2002) (“[T]he *Daubert* decision did not impact on the admission rates of expert testimony at either the trial or appellate court levels.”); D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock*, 64 ALB. L. REV. 99, 103-12 (2000) (providing empirical evidence that judges are more likely to admit prosecution expert testimony than other types of expert testimony). For a review of the empirical findings, see David L. Faigman, *Admissibility Regimes: The “Opinion Rule” and Other Oddities and Exceptions to Scientific Evidence, The Scientific Revolution and Common Sense*, 36 SW. U. L. REV. 699, 717-18 (2008).

14. STRENGTHENING FORENSIC SCIENCE, *supra* note 11, at 109 (quoting Peter J. Neufeld, *The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform*, 95 AM. J. PUB. HEALTH 107, 108 (2005)).

part because the forensic community has made it difficult to easily distinguish between better and worse testimony. I propose some steps that could be taken to make such distinctions more transparent. I briefly review admissibility decisions in drunk driving cases to indicate how greater transparency in that area has led to more frequent exclusion of less well-supported conclusions. I end with a brief conclusion.

I. THE NRC'S REASONS FOR GIVING THE COURTS ALMOST NO ROLE IN IMPROVING FORENSIC EXPERTISE

Prior to the *Daubert* trilogy of cases, one might have argued that the courts could play no role in improving expert testimony because they do not have the conceptual tools to do so. This is no longer the case. Unlike the *Frye* test that asked judges when making admissibility determinations to defer to the judgment of the relevant expert community—too often interpreted as the community of individuals offering their testimony in court—the *Daubert* opinion invited judges to assess the evidence themselves, undermining the self-validating tendency of the *Frye* test.¹⁵ Moreover, the *Daubert* Court equated evidentiary reliability with scientific validity¹⁶ and declared that two important admissibility factors were whether theory or technique could be, and has been, tested,¹⁷ and the error rate associated with a procedure.¹⁸ Many forensic areas score poorly on these factors.¹⁹

Why, then, does the Report fail to call upon the courts to change their liberal admissibility policy? The answer, apparently, is that the Report has no faith in the ability of the courts to use these tools, or at least no faith in their ability to use them effectively. Here is one key passage explaining why the Report gives the courts a pass:

The judicial system is encumbered by, among other things, judges and lawyers who generally lack the scientific expertise necessary to comprehend and evaluate forensic evidence in an informed manner, trial judges (sitting alone) who must decide evidentiary issues without the benefit of judicial colleagues and often with little time for extensive research and reflection, and the highly deferential nature of the appellate review af-

15. Michael J. Saks, *Merlin and Solomon: Lessons From the Law's Formative Encounters With Forensic Identification Science*, 49 HASTINGS L.J. 1069, 1139 (1998) ("But perhaps the purpose of the rules is simply to hold up a target to the courts; call one the *Frye* target and the other the *Daubert* target. The *Frye* ideal says: do whatever the experts tell you to do. The *Daubert* ideal says: figure out the science yourself.")

16. *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 590 n.9 (1993).

17. *Id.* at 593.

18. *Id.* at 594.

19. See, e.g., STRENGTHENING FORENSIC SCIENCE, *supra* note 11, at 142-44 (discussing fingerprint friction analysis).

forded trial courts’ *Daubert* rulings. Furthermore, the judicial system embodies a case-by-case adjudicatory approach that is not well suited to address the systematic problems in many of the various forensic science disciplines.²⁰

The implication is that judges are simply not up to the challenge. A combination of lack of time, lack of general focus, and lack of expertise, places this task beyond the ability of the judiciary.

Undoubtedly, a great many opinions admitting forensic evidence leave much to be desired. The inanities sometimes advanced to admit forensic testimony are well documented. Consider the following examples.

In a microscopic hair comparison case, the Kentucky Supreme Court, purportedly applying the state’s *Daubert* test, was presented with no empirical support for the reliability of the test and ultimately fell back on general acceptance.²¹ But it did so in a very backhanded way. There were no Kentucky cases reviewing such evidence under *Frye*. The court took this absence of any prior assessment of the admissibility of microscopic hair analysis as evidence for its admissibility.

Although we have never specifically addressed the scientific reliability of this method of hair analysis, we must assume that it at least satisfied the *Frye* test of general acceptance; for otherwise, the evidence would never have been admitted in the first place. The absence in our previous opinions of any in-depth analysis under the “general acceptance” test was probably due to the overwhelming acceptance of this procedure as a reliable scientific method for the past fifty years.²²

Additionally, the United States District Court of the Southern District of Indiana in *United States v. Havvard*,²³ ruled on the issue of whether fingerprint analyses had been tested per *Daubert*’s falsification criterion.

They have been tested for roughly 100 years. They have been tested in adversarial proceedings with the highest possible stakes—liberty and sometimes life. The defense has offered no evidence in this case undermining the reliability of the methods in general. The government points

20. *Id.* at 110.

21. *Johnson v. Commonwealth*, 12 S.W.3d 258 (1999).

22. *Id.* at 262. In lieu of any Kentucky cases on point, the court points to opinions in a number of other states declaring such testimony to be admissible. Never mind that in none of those cases did the court perform even a perfunctory analysis of the reliability of this evidence. In most of these cases either there was no mention of the *Frye* test, or the evidence that the testimony was generally accepted came from the witness himself. *See, e.g.*, *McGrew v. State*, 682 N.E.2d 1289, 1291 (1997); *see also* Michael J. Saks, *Explaining the Tension Between the Supreme Court’s Embrace of Validity as the Touchstone of Admissibility of Expert Testimony and Lower Courts’ (Seeming) Rejection of Same*, 5 *EPISTEME* 329, 339 (2008) (discussing *Johnson*, 12 S.W.3d 258).

23. 117 F. Supp. 2d 848 (S.D. Ind. 2000).

out correctly that if anyone were to come across a case in which two different fingers had identical fingerprints, that news would flash around the legal world at the speed of light. It has not happened in 100 years.²⁴

In the fingerprint arena, *Havvard* hardly stands alone. With very few exceptions, courts have refused to conduct *Daubert* hearings,²⁵ have implicitly reversed the burden of persuasion to require the defendant to demonstrate that a fingerprint identification is not reliable,²⁶ have admitted expert testimony by relying on the fact that other courts have admitted the testimony,²⁷ have relegated any concerns about validity to weight, not admissibility,²⁸ and in general have lowered the bar to the level necessary to admit fingerprint identification.²⁹ Perhaps most importantly, most courts have not been sensitive to the importance of distinguishing among different identification situations and, therefore, have failed to follow *Kumho Tire's* admonition to focus on the task at hand.³⁰

These examples to the contrary notwithstanding, the reasons the Report gives as to why the judiciary has not done a better job policing the quality

24. *Id.* at 854. The most obvious problem with this passage is that the court substitutes the adversarial process for scientific investigation, a move all too often seen in forensic admissibility opinions. Moreover, the court simply confuses the relevant issue. The last sentence focuses on the issue of whether any two different fingerprints are identical. Even if we agree that they are not, this does not mean that there is good empirical research on the ability of an examiner to state with certainty that a blurred partial print could come from one and only one source.

Out of abundant caution, I want to be clear that no one thinks fingerprint identification and microscopic hair analysis are equally likely to produce errors. Clearly, there are better (fingerprints) and less reliable (voiceprint) methods available to forensic scientists. See Jennifer L. Mnookin, *The Validity of Latent Fingerprint Identification: Confessions of a Fingerprinting Moderate*, 7 LAW, PROBABILITY & RISK 127, 134 (2008). However, in all of these areas there are numerous opinions that simply do not address the quality of the evidence supporting expert testimony.

25. See, e.g., *United States v. Ambriz-Vasquez*, 34 F. App'x 356, 359 (9th Cir. 2002); *United States v. Joseph*, No. CR.A.99-238, 2001 WL 515213 (E.D. La. May 14, 2001).

26. See, e.g., *United States v. Rogers*, 26 F. App'x 171, 173 (4th Cir. 2001).

27. See, e.g., *Havvard*, 260 F.3d at 601; *United States v. Frias*, No. S3 01 Crim. 307 (AGS), 2003 WL 352502 (S.D.N.Y. Feb. 13, 2003).

28. See, e.g., *United States v. Cline*, 188 F. Supp. 2d 1287, 1294 (D. Kan. 2002), *aff'd*, 349 F.3d 1276 (10th Cir. 2003).

29. See *id.* A rare exception to this landslide of opinions admitting fingerprint evidence was *United States v. Llera-Plaza*, 57 Fed. R. Evid. Serv. 983 (E.D. Pa. 2002) but Judge Polak withdrew his initial opinion excluding fingerprint evidence and admitted the evidence. *United States v. Llera-Plaza*, 188 F. Supp. 2d 549 (E.D. Pa. 2002); see also *United States v. Crisp*, 324 F.3d 261, 278 (4th Cir. 2003) (Michael, J. dissenting) (suggesting that fingerprint evidence should have been excluded).

30. For a discussion of the post-*Daubert* case law on the admissibility of fingerprint evidence see 4 DAVID L. FAIGMAN ET AL., MODERN SCIENTIFIC EVIDENCE §§ 33:3-33:19 (2008-2009). Although I have focused on federal court opinions, the states generally replicate this pattern. See *id.*

of expert testimony seems to me to miss the mark. Throughout the *Daubert* era, I have played the role of an interested observer of the forensics wars. My own focus has largely been on experts in civil cases, most specifically in toxic tort litigation. In those cases, the judiciary has played an active role in policing the quality of expert testimony. Critics usually complain that, if anything, the courts have been too restrictive.³¹

In my opinion, on balance, the courts have performed their gate-keeping role with a reasonable degree of competence. Moreover, in the years since *Daubert*, they have become more sophisticated in their assessment of expert testimony, primarily because they have moved away from using the four *Daubert* factors to make global statements concerning the reliability of an area of knowledge and instead focused on the fit between the available scientific evidence and the testimony of the witness.³²

The fact that courts do police the reliability of expert testimony in toxic cases raises the question of why they cannot play a significant role in the forensic arena. Civil law jurisprudence suggests that the reason cannot be, as the Report suggests, that judges lack the scientific expertise necessary to comprehend and evaluate this evidence. There are many civil law cases in the toxic tort arena where the scientific issues are at least as complex as

31. Among the critical articles are: Michael F. Baumeister & Dorothea M. Capone, *Admissibility Standards as Politics—The Imperial Gate Closers Arrive!!!*, 33 SETON HALL L. REV. 1025 (2003); Margaret A. Berger & Aaron D. Twerski, *Uncertainty and Informed Choice: Unmasking Daubert*, 104 MICH. L. REV. 257 (2005); Margaret A. Berger, *Upsetting the Balance Between Adverse Interests: The Impact of the Supreme Court’s Trilogy on Expert Testimony in Toxic Tort Litigation*, 64 LAW & CONTEMP. PROBS. 289, 324 (2001); Carl F. Cranor & David A. Eastmond, *Scientific Ignorance and Reliable Patterns of Evidence in Toxic Tort Causation: Is There a Need for Liability Reform?*, 64 LAW & CONTEMP. PROBS. 5, 15 (2001); Lucinda M. Finley, *Guarding the Gate to the Courthouse: How Trial Judges are Using Their Evidentiary Screening Role to Remake Tort Causation Rules*, 49 DEPAUL L. REV. 335, 337 (1999); Allan Kanner & M. Ryan Casey, *Daubert and the Disappearing Jury Trial*, 69 U. PITT. L. REV. 281 (2007); Bobak Razavi, *Admissible Expert Testimony and Summary Judgment*, 29 J. LEGAL MED. 307 (2008).

32. See Edward J. Imwinkelried, A ‘Daubert’ Checklist, NAT’L L.J., Sept. 12, 2005, at 12. “Consequently, today many judges have moved beyond a mechanistic checklist approach. Rather, they address the broader, bottomline question of the reliability of the evidence They sift through the foundational testimony to determine whether there is an ‘empirically supported rational explanation’ for the expert’s opinion.” Bert Black, *Learn the Science in Your Cases*, 39 TRIAL 18, 19-20, 25 (2003).

Increasingly, judges feel comfortable addressing such considerations as the size of the researchers’ database, its composition, the test conditions, and the rates for false positives and negatives. These judges no longer mechanistically tally up the number of *Daubert* factors favoring or opposing admission. Instead, they quite correctly address the bottom line: Does the proponent’s empirical data establish that this expert can accurately draw the inference that he or she is prepared to testify to?

those in the forensic area.³³ It is true that trial courts, especially state trial courts, have limited time and resources to devote to admissibility issues, but federal district courts confront the same time and resource pressures in civil and criminal cases and still do a more thorough job on the civil side.

Another argument advanced by the Report as to why the judiciary has no substantial role to play in improving forensic science is that “the case-by-case adjudicatory approach [is] not well suited to address the systematic problems in many of the various forensic science disciplines.”³⁴ Once again it is important to note that the same constraint confronts admissibility decisions in civil cases and yet, over time, courts are able to sort out the merits of expert testimony with respect to different alleged injuries.³⁵ Indeed, the underlying issue in *Daubert* is a case in point. Over time, the courts concluded that human and animal studies on the relationship between the morning sickness drug Bendectin and birth defects in the children of mothers who ingested the drug was insufficient to support expert testimony alleging such a relationship.³⁶

Perhaps an even better example is the silicone breast implant litigation. In the earliest years of the litigation, before much scientific evidence had been accumulated, courts generally admitted plaintiffs’ expert testimony.³⁷ By the late 1990s, there was more epidemiological data indicating no causal relationship between implants and traditional autoimmune disease. Court decisions were more mixed on the admissibility of plaintiff experts.³⁸ In 1998, a Federal Rule of Evidence 706 expert panel appointed by the silicone breast implant multi-district litigation judge issued its report.³⁹ This was followed in 2000 by a report from the National Academy of Science (NAS).⁴⁰ Both concluded that there was no relationship between implants and autoimmune or connective tissue diseases.⁴¹

33. See 3 DAVID L. FAIGMAN, ET AL., MODERN SCIENTIFIC EVIDENCE § 22:23 (2008-2009) for insightful analyses of difficult toxicology and epidemiology questions.

34. STRENGTHENING FORENSIC SCIENCE, *supra* note 11, at 110.

35. See FAIGMAN ET AL., *supra* note 33, at §§ 22:23-22:42.

36. See MICHAEL D. GREEN, BENDECTIN AND BIRTH DEFECTS: THE CHALLENGES OF MASS TOXIC SUBSTANCES LITIGATION (1996); JOSEPH SANDERS, BENDECTIN ON TRIAL: A STUDY OF MASS TORT LITIGATION (1998).

37. *Hopkins v. Dow Corning Corp.*, 33 F.3d 1116 (9th Cir. 1994).

38. Compare *Vassallo v. Baxter Healthcare Corp.*, 696 N.E.2d 909 (Mass. 1998), and *Jennings v. Baxter Healthcare Corp.*, 14 P.3d 596 (Or. 2000) (experts admitted), with *In re Breast Implant Litig.*, 11 F. Supp. 2d 1217 (D. Colo. 1998) (expert excluded).

39. *In re Breast Implant Cases*, 942 F. Supp. 958, 960 (S.D.N.Y. 1996).

40. INST. OF MED., SAFETY OF SILICONE BREAST IMPLANTS 10-11 (Stuart Bondurant et al., eds., 2000).

41. The Multidistrict Panel findings are summarized in *Grant v. Bristol-Myers Squibb*, 97 F. Supp. 2d 986, 989-90 (D. Ariz. 2000) (alteration in original):

The results of these two reports were what one would expect. Plaintiff experts were routinely excluded and, as a result, silicone breast implant cases dealing with autoimmune and connective tissue diseases—but not cases dealing with localized injury—soon became a thing of the past.⁴²

Clearly there are differences between these two reports and the National Research Council’s Forensics Report. The former two were mostly backward looking, focusing on the science concerning implants and autoimmune disease. The Forensics Report is forward looking, searching for ways to improve the quality of forensic evidence. Nevertheless, the courts have responded positively to the increasing research on implants and corresponding expert reports. Why, then, should we anticipate they would not act similarly in response to the latter report?

II. WHY ARE COURTS SO LENIENT IN ADMITTING EXPERT FORENSIC EVIDENCE?

To answer that question one needs to have some purchase on the question of why courts have been so lenient up to this point. One effort to answer that question comes from Michael Saks.

A. Saks’ Explanation

Saks offers several possible explanations. One explanation is that the courts are grandfathering in the traditional forensic sciences. They would never permit a new area of forensic evidence to gain a foothold based on such little support. DNA is, in fact, a case in point. It took two separate

The executive summary of that Panel reported their conclusion that “[n]o association was evident between breast implants and any of the individual connective tissue diseases, all definite connective tissue diseases combined, or the other autoimmune/rheumatic conditions” The Panel also found no association between breast implants and atypical connective tissue diseases or any distinctive constellation of symptoms observed in women with breast implants. Panelists noted that their findings and conclusions were unanimous, and that “a large majority of scientists in our respective disciplines would find merit on our reviews and analysis.”

The NAS Committee concluded:

- There is no increase in primary or recurrent breast cancer in implanted women.
- [Autoimmune and connective tissue] diseases or conditions are no more common in women with breast implants than in women without implants.

See INST. OF MED., *supra* note 40 (alteration in original).

42. See *Norris v. Baxter Healthcare Corp.*, 397 F.3d 878 (10th Cir. 2005); *Allison v. McGhan Medical Corp.*, 184 F.3d 1300 (11th Cir. 1999); *Bristol-Myers Squibb*, 97 F. Supp. 2d at 986. For discussions of the life cycle of mass tort congregations, see Francis E. McGovern, *Toward a Functional Approach for Managing Complex Litigation*, 53 U. CHI. L. REV. 440, 488 (1986); Joseph Sanders, *The Bendectin Litigation: A Case Study in the Life Cycle of Mass Torts*, 43 HASTINGS L.J. 301 (1992).

National Academy reports on the validity of DNA typing before courts routinely admitted this evidence.⁴³ Working against this explanation is the fact that many courts recognize that *Daubert* invites re-examination of types of testimony that in the past had been routinely admitted.⁴⁴

A second explanation is that judges are willing to overlook the shortcomings of forensic evidence “once they are satisfied that they know it is being used to bring about the conviction of a person they know to be guilty.”⁴⁵ One way to view this would be through the lens of “harmless error.” Other evidence in the case persuades the court that the individual is guilty and admitting the forensic evidence therefore will not work an injustice. But if the forensic evidence is truly superfluous, why not exclude it? An even more damning explanation is that judges are prepared to permit what they believe to be evidence with no probative value simply because they believe most criminal defendants to be guilty.

Saks rejects all of these arguments in favor of the explanation that judges admit the testimony of forensic experts “because they substantially believe the claims of these fields.”⁴⁶ He then offers several explanations for why this is the case. One explanation is that judges use peripheral processing in deciding to admit this evidence. By peripheral processing, social psychologists mean the decision-maker does not focus on the substance of an argument—called “central processing”—but rather on other cues to its validity. One such cue is the general acceptance of an argument in the society. Saks notes that American popular culture generally accepts the forensic claims made in the legal context⁴⁷ and that by calling their fields “scientific

43. See NAT'L RES. COUNCIL, DNA TECHNOLOGY IN FORENSIC SCIENCE 8 (1992); NAT'L RES. COUNCIL, THE EVALUATION OF FORENSIC DNA EVIDENCE 88 (1996).

44. See, e.g., *United States v. Horn*, 185 F. Supp. 2d 530 (D. Md. 2002).

45. Saks, *supra* note 22, at 334.

46. *Id.* at 335.

47. For an illustrative example of this tendency, see the introductory paragraphs in Jennifer L. Mnookin's article, *The Courts, the NAS, and the Future of Forensic Science*, 75 BROOK. L. REV. 1209 (2010).

On a recent flight, the person next to me on the crowded airplane began to chat with me. When I told her about what I researched and studied, she looked at me with a big grin. “I LOVE forensic science,” she said. “I watch CSI whenever I can. They can do such amazing things. It's all so high tech—and incredibly accurate! It's almost like magic, isn't it?” She leaned in a bit closer and looked at me intently. “Tell me, is it like that in real life?”

I looked at her for a moment before answering. I felt a bit like the older child on the playground about to reveal to her younger friend that Santa Clause doesn't really exist. I shook my head. “No, I wouldn't say that CSI's depiction is entirely realistic. In the real world, forensic science isn't nearly so glossy. It isn't nearly so speedy. And most important, it isn't nearly so foolproof, either.”

ic,” forensic experts also tap into the general credulity with which we accept statements made by scientists. If, in fact, peripheral processing is a significant part of the judicial acceptance of forensic experts, then the Forensics Report should lead to a less lenient approach to admitting forensic expert testimony. The Report is an important voice expressing skepticism about many forensic expert assertions and it clearly undermines any effort by many forensic experts to wrap themselves in a scientist cloak. At the same time, it undermines the position taken by some judicial opinions claiming that forensic expertise has strong scientific underpinnings. There is some hope that this will tip the judicial scales in favor of less lenient admissibility decisions.⁴⁸

I, however, remain somewhat skeptical that the Report itself will produce substantial change. I take this view in part because I think that in another part of its Report, the National Research Council came closer to putting its finger on why courts adopt permissive admissibility rules when it quoted Joan Griffin and David J. LaMagna for the following proposition: “[S]ome courts appear to be loath to insist on [rigorous systematic] research as a condition of admitting forensic science evidence in criminal cases, perhaps because to do so would likely ‘demand more by way of validation than the disciplines can presently offer.’”⁴⁹

This quote captures a key insight into all judicial admissibility decisions: courts believe admissibility standards should not be set “too high.” This self-imposed judicial constraint is, in my opinion, the product of two other, more deeply seated issues: the judiciary’s contextual approach to know-

“Really? That’s too bad,” she told me. She looked at me directly for a brief moment, shook her head, and then looked away. “Well, to tell you the truth, I think I’d rather just keep believing in the television version.” Figuring that reality was not going to be any match for CSI, I shrugged, and went back to the book I was reading.

Id.

48. See William C. Thompson, *The National Research Council’s Plan to Strengthen Forensic Science: Does the Path Forward Run Through the Courts?*, 50 JURIMETRICS J. 35, 36-37 (2009).

As commentators have noted, there currently are two opposing “literatures” on these disciplines. One literature, consisting of court opinions and the writings of forensic practitioners, holds that these disciplines are valid, well-grounded in science, and easily meet the *Daubert* and *Frye* standards for admissibility of scientific testimony. The other literature, consisting of commentary by academics (most of whom are not forensic practitioners), holds that these disciplines rest on shaky scientific foundations and produce evidence of questionable quality. By adding the considerable weight of the NRC’s authority to the “critical” literature on forensic identification science, the NRC Report will make it much harder for courts to continue siding with forensic practitioners and ignoring the academic critics.

49. STRENGTHENING FORENSIC SCIENCE, *supra* note 11, at 109 (alteration in original) (quoting Joan Griffin & David J. LaMagna, *Daubert Challenges to Forensic Evidence: Ballistics Next on the Firing Line*, 26 CHAMPION 20, 21 (2002)).

ledge and the limited ability of science to provide causal information about a particular case.

B. Contextual Approach to Knowledge

If anything is clear from a comparison between forensic science and toxic tort cases it is that courts do not use a single standard when determining whether or not to admit expert testimony. A brief foray into epistemology helps us to understand and, perhaps, to justify this position.⁵⁰

The standard approach to the question of when it is proper to say someone knows something involves the interplay of three factors: belief, truth and justification.⁵¹ Belief is a person's subjective position concerning the truth of a proposition.⁵² Truth is the reality of the proposition independent of belief.⁵³ Justification involves the quality of the reasons for a belief. To count as knowledge, something must be believed to be true, it must be true, and a person's belief that it is true must be justified. In the absence of belief, we have ignorance.⁵⁴ In the absence of truth, we have error.⁵⁵ In the absence of appropriate justification, we have mere opinion.⁵⁶ What is most noteworthy about this standard approach is that its main concern is not knowledge per se but justification. Even correct beliefs without justification are not knowledge.

For courts, the relevant question is what level of justification should be required before experts are permitted to testify. The legal answer to this question has been to adopt a contextual approach to knowledge.⁵⁷ The central idea behind contextualism is that the standards for making knowledge

50. I have discussed the law's adoption of a contextual approach at greater length elsewhere. See Joseph Sanders, *Expert Witness Ethics*, 76 FORDHAM L. REV. 1539 (2007); Joseph Sanders, *Science, Law, and the Expert Witness*, 72 LAW & CONTEMP. PROBS. 63 (2009).

51. See D. Michael Risinger & Michael J. Saks, *Rationality, Research and Leviathan: Law Enforcement-Sponsored Research and the Criminal Process*, 2003 MICH. ST. L. REV. 1023, 1024.

52. *Id.*

53. *Id.*

54. *Id.*

55. *Id.*

56. See MICHAEL WILLIAMS, PROBLEMS OF KNOWLEDGE: A CRITICAL INTRODUCTION TO EPISTEMOLOGY 16-19 (2001).

57. In fact, there are several variations on the contextual approach. See David Lewis, *Elusive Knowledge*, in EPISTEMOLOGY: AN ANTHOLOGY 691 (Ernest Sosa et al. eds., 2d ed. 2008); see also Stewart Cohen, *Contextualist Solutions to Epistemological Problems: Scepticism, Gettier, and the Lottery*, in EPISTEMOLOGY: AN ANTHOLOGY, *supra*, at 706; Keith DeRose, *Solving the Skeptical Problem*, in EPISTEMOLOGY: AN ANTHOLOGY, *supra*, at 661; Michael Williams, *Epistemological Realism*, in EPISTEMOLOGY: AN ANTHOLOGY, *supra*, at 51. Cranor advocates a contextual approach in the area of regulation. See CARL F. CRANOR, REGULATING TOXIC SUBSTANCES: A PHILOSOPHY OF SCIENCE AND LAW 152, 152-78 (1993).

attributions vary depending on the context within which they are made. The rules governing the admissibility of expert testimony are applied in a way that is consistent with the contextualist's fundamental observation that the level of justification we require for something to count as knowledge varies according to the context within which the belief is held and expressed. Within the confines of the present discussion, the most important context is the quantity and quality of the available evidence.

The contextual nature of the legal approach to expertise permits the law to sidestep some difficult philosophical questions concerning knowledge and get on with the business of deciding cases.⁵⁸ However, this approach does not provide a clear-cut admissibility standard. One possible standard, and indeed the one often espoused by courts, is that an expert should use the same level of intellectual rigor that is employed in her field. The general acceptance test articulated in *Frye v. United States*⁵⁹ can be understood this way. In *Kumho Tire v. Carmichael*,⁶⁰ the United States Supreme Court explicitly adopts this test. The same intellectual rigor test has the virtue that it does not set the admissibility threshold too high. It does not create a standard that condemns a whole group of litigants to certain defeat. It does not ask for the impossible.⁶¹

Of course the problem in the forensics area is not that the bar may be set too high but that, in some cases, it is set too low. This leads to the second deeply seated issue, the often-limited ability of science to provide causal information about a particular case.

58. One is reminded of Sir Frederick Pollock's famous aphorism that “[t]he lawyer cannot afford to adventure himself with philosophers in the logical and metaphysical controversies that beset the idea of cause.” FREDERICK POLLOCK, *THE LAW OF TORTS* 36 (11th ed. 1920).

59. 293 F. 1013, 1014 (D.C. Cir. 1923); see also Samuel R. Gross & Jennifer L. Mnookin, *Expert Information and Expert Evidence: A Preliminary Taxonomy*, 34 *SETON HALL L. REV.* 141, 148 (2003).

60. 526 U.S. 137, 152 (1999) (“The objective of [the *Daubert*] requirement is to ensure the reliability and relevancy of expert testimony. It is to make certain that an expert, whether basing testimony upon professional studies or personal experience, employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field.”).

61. For example, in the area of toxic torts most agree that the best evidence as to whether some substance causes human injury is a well-conducted body of epidemiological research. What should we do, however, when there is no epidemiology with respect to some substance? Following the contextual approach, almost all courts would say that it simply is not required. See generally 3 DAVID L. FAIGMAN ET AL., *MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY* § 23 (2010-2011).

C. Information About the Particular Case

Science tends to focus on systematic and general knowledge. As David Faigman notes, “[s]cientists typically study variables at the population level, and most of their methodological and statistical tools are designed for this kind of work.”⁶² The heroes of science are those who are able to put forth explanations in terms of general laws that explain a myriad of particular observations. The structure of their investigation is to search for the effects of causes. The law often is concerned with such general questions, e.g. does drug X cause injury Y? Most trials, however, must deal with specific events. They are concerned with what happened to a specific person or at a particular crime scene at a specific point in time. For example, did drug X cause the plaintiff’s injury? Were the fingerprints left at the crime scene those of the defendant? That is, the law is searching not for the effects of cause but instead is searching for the causes of effects.⁶³

Sometimes experts can provide us with quite precise answers to these sorts of questions. Within the forensic science arena, DNA typing frequently offers substantial exactitude.⁶⁴ Other forensic areas may offer less precise, but still adequate understanding of the cause of some effect.⁶⁵ Unfortunately, in other areas because our understanding of the effects of some causes is quite imprecise there is substantial uncertainty surrounding any effort to pinpoint the cause of an effect.⁶⁶

It is important to note that this problem is not unique to forensic science. Consider once again the area of toxic torts. When substances lead to “signature diseases,” experts may be substantially certain that the substance caused the plaintiff’s injury. Asbestosis is caused by exposure to asbestos.⁶⁷ With respect to many other exposures, however, our general understanding of the effect of a cause, e.g., that Vioxx causes heart problems, does not translate into a clear understanding as to which Vioxx takers with heart problems can claim that the drug caused their injury.⁶⁸ Because

62. David L. Faigman, *Evidentiary Incommensurability: A Preliminary Exploration of the Problem of Reasoning From General Scientific Data to Individualized Legal Decision Making*, 75 BROOK. L. REV. 1115, 1115 (2010).

63. I borrowed this terminology and several of the points below from A. Philip Dawid, *The Role of Scientific and Statistical Evidence in Assessing Causality*, in PERSPECTIVES ON CAUSATION (Richard Goldberg ed., forthcoming 2011).

64. See STRENGTHENING FORENSIC SCIENCE, *supra* note 11, at 133.

65. Fingerprints are in this category. See generally Mnookin, *supra* note 24.

66. See STRENGTHENING FORENSIC SCIENCE, *supra* note 11, at 178-79 (discussing bloodstain pattern analyses).

67. See 3 FAIGMAN ET AL., *supra* note 61, § 26:22.

68. See *Merck & Co. v. Ernst*, 296 S.W.3d 81, 95-101 (Tex. App. 2009).

plaintiffs in toxic tort cases must prove specific causation⁶⁹ they turn to experts who introduce “differential diagnosis” evidence intended to prove that among the many possible causes of the plaintiff’s injury (including unknown causes) the substance in question is in fact the cause. The quality of this evidence is as mixed as the quality of forensic science evidence.⁷⁰

In sum, the combination of a contextual approach to knowledge and the limited ability to acquire knowledge about the causes of many effects combine to cause courts to establish low thresholds for the admission of individual evidence, be it forensic evidence about the source of a bite mark or the specific causation evidence as to whether a drug exposure caused the plaintiff’s decedent’s suicide.⁷¹ Certainly there is some reliability threshold below which even a court with a contextual approach will not go. But this does not influence the many areas where there is some evidence that expert opinion, however weak it may be, is superior to lay opinion.

I do not wish to argue that these two factors, a contextual approach to knowledge and expert’s limited ability to pinpoint the causes of effects, are the only determinants of admissibility standards. Other factors surely influence judges. However, I do believe that they are elements in the courts’ overall approach to expertise. Moreover, they are both foundational. They are not easily changed. A contextual approach commits courts to adjust admission standards so as to fit within the boundaries of available knowledge and the available knowledge on the causes of effects is often quite limited. If I am correct on these points, we should not expect that courts would be easily persuaded to exclude whole areas of forensic expertise solely because the available evidence is relatively weak. Undoubtedly, Saks is correct to conclude that some judges simply believe that forensic expert opinion is well supported by empirical evidence. Nevertheless, I am skeptical that all courts are admitting this evidence because they are fundamentally confused about the relevant scientific community whose general acceptance they should be gauging or because they are simply unable to distinguish between weak and strong evidence. They understand the evidence is weak but admit it anyway.

If I am correct, the NRC Report will not produce a radical change in admissibility standards simply because it will cause the judiciary finally to

69. That is, it must be proven by a preponderance of the evidence that the substance in question is capable of causing the type of injury suffered (general causation) and that the particular injury was caused by the substance (specific causation). See RESTATEMENT (THIRD) OF TORTS: LIABILITY FOR PHYSICAL AND EMOTIONAL HARM § 28 cmt. C (Proposed Final Draft No. 1, 2005).

70. See Joseph Sanders, *Applying Daubert Inconsistently? Proof of Individual Causation in Toxic Tort and Forensic Cases*, 75 BROOK. L. REV. 1367, 1369 (2010).

71. See *id.*

understand the weakness of much forensic data. To believe this demeans both the ability of a cadre of very competent academics to persuade others of the merit of their assessment of forensic science and the ability of judges to distinguish self-serving reliability assertions by individuals whose livelihood turns on the admission of this testimony from the reasoned critiques of their arguments.

III. WHAT ROLE MIGHT COURTS PLAY?

This does not mean, however, that courts have no role to play in improving forensic evidence. Contextualism and the difficulty of ascertaining the causes of effects help to explain why courts are unwilling to exclude whole areas of forensic expertise, but they do not explain why courts should be unwilling to exclude the worst evidence in a particular field. For example, courts might exclude fingerprint evidence when the trace evidence left at a crime scene is of particularly poor quality. Or they might permit a handwriting expert to testify that a signature is a forgery based on many exemplars of a known authentic signature while excluding testimony on the authorship of an attempted forgery.⁷²

Why don't courts more frequently exclude the worst forensic evidence and is there anything we can do to encourage them to do so?⁷³ One reason courts may be reluctant to take this step is that they cannot easily ascertain the relative quality of various proffers. Of course, much of the blame for this state of affairs lies at the feet of the forensic science community, its commitment to the ideas of individualization and uniqueness, its unwillingness to recognize in any area other than DNA testing that its conclusions are probabilistic and, therefore, its assertion that any errors that do occur are the result of examiner error, i.e., a more proficient examiner would not have made this mistake.⁷⁴ If courts cannot get behind such assertions, they are left with the impossible task of estimating whether a par-

72. See D. Michael Risinger, *Goodbye to All That, or a Fool's Errand, by One of the Fools: How I Stopped Worrying About Court Responses to Handwriting Identification (and "Forensic Science" in General) and Learned to Love Misinterpretations of Kumho Tire v. Carmichael*, 43 TULSA L. REV. 447, 448-50 (2007).

73. Altering the willingness of courts to exclude the worst forensic evidence would hopefully encourage forensics experts to improve the overall quality of forensic evidence. For several other suggestions for how courts might directly improve the quality of forensic evidence, see Michael J. Saks, *Protecting Factfinders From Being Overly Mislead, While Still Admitting Weakly Supported Forensic Science Into Evidence*, 43 TULSA L. REV. 609 (2007).

74. See Simon A. Cole, *More Than Zero: Accounting for Error in Latent Fingerprint Identification*, 95 J. CRIM. L. & CRIMINOLOGY 985, 1042-43 (2005).

ticular examiner has reached a particularly suspect conclusion in the case at hand.

The NRC Report will enjoy substantial success if it causes the forensics community to abandon this type of argument and do the type of research that will permit better calibration of the diagnostic quality of forensic identifications. Perhaps it is not possible to produce the type of base rate data that makes DNA estimates relatively exact,⁷⁵ but, as the Report notes, well-designed research could examine the variables that affect quality.⁷⁶ Most immediately, we could better calibrate reliability through serious proficiency testing. Such testing would provide estimates of the error rate (the frequency of false positives and false negatives) associated with a reported match using data of varying quality using various laboratory procedures. This assessment would be independent from the alleged skill of a particular examiner.

Efforts in this direction would benefit from the inclusion of computer algorithms to interpret results. A movement toward routine, systematic, quantified methods of interpretation further reduces the variability in reliability introduced by the use of individual examiners. More importantly, it provides a systematic benchmark against which to judge the reliability of a given forensic identification.

Would any of this actually change current court admissibility practices? There is, of course, reason to remain skeptical. Forensic experts certainly have not embraced the NRC Report or its research agenda.⁷⁷ However, if we can begin down this road, there is some evidence that courts may become more aggressive in policing the reliability of forensic expert opinion. Consider the case of alcohol testing in drunk driving cases.

Most comparisons of court admissibility standards pit criminal forensic standards against standards in civil cases.⁷⁸ While these comparisons are worthwhile, conclusions must be tempered because of the myriad ways civil and criminal cases differ. A closer comparison is between alcohol testing

75. Even if base rate information were available, it might be of little practical value in areas such as fingerprint identification. In most DNA cases, samples contain complete or nearly complete information about the sequences of base pairs at multiple loci. Metaphorically, they are nearly “perfect prints.” The problem in the real world of fingerprint identification is that many latent prints are far from perfect.

76. See STRENGTHENING FORENSIC SCIENCE, *supra* note 11, at 105-06.

77. See Jonathan J. Koehler, *Forensic Science Reform in the 21st Century: A Major Conference, a Blockbuster Report and Reasons to be Pessimistic*, 9 LAW, PROBABILITY & RISK 1, 3-6 (2010).

78. See Margaret A. Berger, *Expert Testimony in Criminal Proceedings: Questions Daubert Does Not Answer*, 33 SETON HALL L. REV. 1125, 1140 (2003); Paul C. Giannelli, *The Supreme Court’s “Criminal” Daubert Cases*, 33 SETON HALL L. REV. 1071, 1072-73 (2003).

in drunk driving cases and forensic evidence in other criminal cases. Both pit the state against individuals and both compel the judge to balance the welfare and safety of society with the rights of the individual.⁷⁹ Drunk driving admissibility cases indicate that if courts are able to distinguish cases based on the quality of the evidence, they will exclude evidence. This is not a place for a full discussion of admissibility issues in alcohol testing cases.⁸⁰ I restrict myself to a brief overview and a few examples.

The need for a method to test whether an individual is intoxicated must have arisen shortly after the first automobile departed a tavern. Statutes criminalizing drunk driving were enacted in the first decades of the twentieth century.⁸¹ Before the development of objective tests for intoxication, courts had to rely exclusively on the testimony of police officers and others who in turn relied on behavioral indicia of drunkenness and the presence of alcohol on the breath.⁸² Standards began to change with the development of objective chemical tests, e.g., breath and blood tests for the presence of alcohol in an individual's body.⁸³ As the sophistication and accuracy of tests improved, admissibility questions changed. Importantly, courts developed admissibility standards that exclude the least probative test results.⁸⁴

One should not overstate this tendency. It is still possible for someone to be convicted of driving while impaired based solely on the observations of a police officer. However, these situations most frequently arise when the defendant has refused to take a breath test or where for various reasons

79. One might be concerned that the stakes in drunk driving cases are not as high as they are in other criminal cases. Perhaps on balance this is true, but certainly the stakes in many drunk driving cases are far from trivial, especially when someone has been killed and the defendant is being prosecuted for manslaughter. *See, e.g.*, *State v. Mayl*, 833 N.E.2d 1216 (Ohio 2005). Cases with potentially more serious consequences for the defendant tend to be those in which admissibility issues are raised.

80. For a more complete review, see 5 DAVID L. FAIGMAN ET AL., *MODERN SCIENTIFIC EVIDENCE* § 41 (2008-2009). Parts of the following discussion rely on this source.

81. Pennsylvania passed its first DUI statute in 1909. It prohibited operation of a motor vehicle while intoxicated, without a precise definition of intoxication. *See* Robert J. Schefter, *Under the Influence of Alcohol Three Hours After Driving: The Constitutionality of the (A)(5) Amendment to Pennsylvania's DUI Statute*, 100 DICK. L. REV. 441, 444 (1996).

82. *See* H. LAURENCE ROSS, *CONFRONTING DRUNK DRIVING: SOCIAL POLICY FOR SAVING LIVES* 43 (1992).

83. *See* 5 DAVID L. FAIGMAN, *supra* note 80, §§ 41:30, 41:55.

84. *Bransford v. State Taxation & Revenue Dep't. Motor Vehicle Div.*, 960 P.2d 827, 831-32 (N.M. Ct. App. 1998); *State v. Korsakov*, 34 S.W.3d 534, 539-42 (Tenn. Crim. App. 2000).

blood test results have been excluded.⁸⁵ In most cases, convictions are based in part on quantified evidence of intoxication.⁸⁶ Courts routinely restrict the admissibility of expert testimony based on these procedures in a number of ways.

Many courts have established a variety of protocols that must be followed if results are to be admitted. These include: frequently testing equipment on a known sample,⁸⁷ conducting tests within a set period of time after the defendant was driving,⁸⁸ and observing the defendant for a period of time (usually fifteen minutes) prior to the administration of the test to be certain she did not burp, belch, vomit, smoke, or consume any alcohol, thus contaminating the results.⁸⁹

Courts are also sensitive to the timing of a test. What does a test conducted at some later time tell us about the level of intoxication of the defendant while she was driving?⁹⁰ One way to address this issue is to attempt to extrapolate backward to the time of driving, a procedure often called “retrograde extrapolation.” By statute, some states require an expert to extrapolate measured blood alcohol concentration (BAC) back to the time the defendant was driving; the absence of expert retrograde extrapolation testimony is fatal to the state’s case.⁹¹ However, accurate extrapola-

85. For example, this may occur when blood is tested in a hospital rather than a police station, causing chain-of-custody issues to bar admissibility. *State v. Busby*, 893 So.2d 161, 164-67 (La. App. 2005); *see also State v. Bedell*, 556 A.2d 101, 102-03 (Vt. 1989).

86. All states have passed “implied consent” statutes, which characterize a driver’s license as a privilege, rather than a right, which may be suspended if one does not submit to a test. *See, e.g., TEX. TRANSP. CODE* §§ 724.001, 724.031, 724.035 (Vernon 2010).

87. *Bransford*, 960 P.2d at 830; *People v. Mickle*, 187 Misc. 2d 718, 721-22 (N.Y. J. Ct. 2001).

88. *See Allman v. State*, 728 N.E.2d 230, 233-34 (Ind. Ct. App. 2000); *State v. Pendleton*, 849 P.2d 143, 148-49 (Kan. Ct. App. 1993) (explaining that the per se provision requires tests to be performed within two hours of the time the defendant was driving); *People v. Victory*, 166 Misc. 2d 549, 551-53 (N.Y. Crim. Ct. 1995) (discussing the history of the Two Hour Rule in New York). *But see People v. Wager*, 594 N.W.2d 487, 488-90 (Mich. 1999) (noting that administering a test within a reasonable time is not a requirement for purposes of the admissibility of the test results).

89. *Paty v. Dir. of Revenue*, 168 S.W.3d 625, 631-32 (Mo. Ct. App. 2005); *State v. Gregory*, No. 96-CO-89, 1999 WL 756440, at *3-4 (Ohio Ct. App. Sept. 23, 1999) (statute requires twenty minute observation period); *Korsakov*, 34 S.W.3d at 539-40; *State v. Grindstaff*, No. 03C01-9704-CR-00139, 1998 WL 126252, at *2-4 (Tenn. Crim. App. Mar. 23, 1998); *State v. Guidera*, 707 A.2d 704, 704-05 (Vt. 1998).

90. In *Daubert* terms, this may be thought of as a question of “fit.” Do the results of the study support the conclusion the state wishes to draw about a given individual? *See Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 591 (1993).

91. *See Connecticut v. Geisler*, 498 U.S. 1019 (1991); *State v. Robinett*, 106 P.3d 436, 439-40 (Idaho 2005); *Commonwealth v. Colturi*, 864 N.E.2d 498, 504-05 (Mass. 2007) (requiring extrapolation evidence if state prosecutes under the “per se” provision in the Massa-

tions are very difficult absent information about the individual's consumption behavior during the period immediately preceding the arrest. When the state does attempt to introduce retrograde extrapolation evidence, some courts have excluded the testimony as unreliable.⁹²

In sum, because various tests for blood alcohol concentration are formalized into a set of routine protocols, failure to follow the protocols casts the results into question and causes some courts to exclude testimony based on the results. Moreover, in areas such as retrograde extrapolation, because we know of the variables that influence the reliability of conclusions, the courts are more likely to exclude the evidence when experts cannot control for these variables.

Interestingly, a similar pattern has emerged with respect to field sobriety tests. These tests involve such steps as the "walk and turn" test, the "one-leg stand" test and a "finger count" test.⁹³ Courts frequently require that the officer administering the test receive special training.⁹⁴ Part of the training involves following routine procedures when administering the tests. As the field sobriety tests have become more routine, courts have become more willing to exclude expert testimony when the officer has deviated from test protocols.⁹⁵

One element of field sobriety tests deserves special mention. Nystagmus is an involuntary rapid movement of the eyeball, which may be horizontal, vertical, or rotary.⁹⁶ An inability of the eyes to maintain visual fixation as they are turned from side to side (in other words, jerking or bouncing) is known as horizontal gaze nystagmus (HGN).⁹⁷ Proponents of HGN tests believe that alcohol and drug use increases the frequency and amplitude of HGN and causes it to occur at a smaller angle of deviation.⁹⁸ Nystagmus

chusetts statute); *State v. Ladwig*, 434 N.W.2d 594, 595 (S.D. 1989); *State v. Dumont*, 499 A.2d 787, 788-89 (Vt. 1985).

92. See *Evans v. State*, 558 S.E.2d 51, 54-56 (Ga. App. 2001) (excluding defense expert testimony); *State v. Downey*, 195 P.3d 1244, 1251-53 (N.M. 2008); *Mata v. State*, 46 S.W.3d 902, 915-17 (Tex. Crim. App. 2001); *Groggins v. Commonwealth*, 537 S.E.2d 605, 607 (Va. 2000). *But see State v. Vliet*, 19 P.3d 42 (Haw. 2001); *Commonwealth v. Senior*, 744 N.E.2d 614, 619-20 (Mass. 2001) (admitting extrapolation evidence).

93. See *Volk v. United States*, 57 F. Supp. 2d 888, 891 (N.D. Cal. 1999).

94. See *United States v. Horn*, 185 F. Supp. 2d 530, 533 (D. Md. 2002); *Mullady v. State*, 606 S.E.2d 645, 646 (Ga. App. 2004).

95. *State v. Homan*, 732 N.E.2d 952, 956-57 (Ohio 2000); *State v. Hall*, No. E-98-088, 2000 WL 1061875, at *3-4 (Ohio Ct. App. Aug. 4, 2000).

96. A. Serra & R. Leigh, *Diagnostic Value of Nystagmus: Spontaneous and Induced Ocular Oscillations*, 73 J. NEUROLOGY, NEUROSURGERY & PSYCHIATRY 615 (2002).

97. See John P. Ludington, Annotation, *Horizontal Gaze Nystagmus Test: Use in Impaired Driving Prosecution*, 60 A.L.R.4TH, 1129 (1988).

98. See Joseph R. Meaney, *Horizontal Gaze Nystagmus: A Closer Look*, 36 JURIMETRICS J. 383, 384 (1996).

tests have become a routine part of field sobriety tests and are frequently introduced as part of the state’s case in drunk driving prosecutions. Nearly from their inception as part of the field sobriety test protocol, most courts viewed them as “scientific” and subjected them to an admissibility analysis.⁹⁹ Courts have excluded HGN results when conducted by an officer not trained in the technique¹⁰⁰ or when the test was not conducted in accordance with standard procedures.¹⁰¹

With respect to both chemical tests and field sobriety tests, other courts have applied more liberal admissibility standards, often concluding that problems with the evidence go to weight, not admissibility.¹⁰² My point is not that courts routinely bar all questionable chemical or field sobriety tests, only that they seem to take a much harder look at this type of expert evidence than is ordinary for forensic expert contexts.

Courts also seem to be sensitive to the error rates associated with various methods of detecting impairment. They distinguish the extraordinarily accurate gas chromatographic tests at one end of the accuracy spectrum to preliminary screening tests used in the field at the other end of the spectrum.¹⁰³ That is, courts appreciate the error rates of the various alcohol detection devices. As a consequence, they have restricted the conclusions that may be made based on certain types of evidence. This is the case with respect to preliminary alcohol screening tests. Courts hold that they may

99. State v. O’Key, 899 P.2d 663, 674 (Or. 1995); *Horn*, 185 F. Supp. 2d at 537-39. An appendix to the *Horn* opinion reviews the field sobriety test case law in all fifty states.

HGN tests are far from perfect. Among other things, the error rates associated with field administered HGN tests is not well understood. Nevertheless, existing evidence suggests that HGN tests are better predictors of impairment than other field sobriety tests. See Charles R. Honts & Susan L. Amato-Henderson, *Horizontal Gaze Nystagmus Test: The State of the Science in 1995*, 71 N.D. L. REV. 671, 688-89 (1995); Steven J. Rubenzer, *The Standardized Field Sobriety Tests: A Review of Scientific and Legal Issues*, 32 LAW & HUM. BEHAV. 293 (2008) (questioning the validity of SFSTs as indicators of impaired driving).

100. See State v. Crawford, 68 P.3d 848, 853-54 (Mont. 2003); State v. Parker, No. 51027-4-I, 2003 WL 21738778, at *1-2 (Wash. Ct. App. July 8, 2003).

101. See Coone v. Barnes, 266 B.R. 397, 406 (B.A.P. 8th Cir. 2001); *Homan*, 732 N.E.2d at 955-56.

102. See Miller v. State, 597 So.2d 767, 769-70 (Fla. 1991); State v. Vliet, 19 P.3d 42, 53-54 (Haw. 2001); State v. Tousley, 611 S.E.2d 139, 146-47 (Ga. App. 2005).

103. See, e.g., Mogg v. State, 918 N.E.2d 750, 759 (Ind. Ct. App. 2009) (refusing to recognize the results of an alcohol monitoring bracelet to indicate anything other than alcohol consumption).

be used only to show impairment, not a specific BAC level.¹⁰⁴ The same is true of field sobriety tests, including the HGN test.¹⁰⁵

From one perspective, this brief overview of admissibility issues in drunk driving cases is quite discouraging. One might conclude that courts are more solicitous of the rights of drunk drivers than those charged with serious felonies. However, from another perspective I find the courts' approach to scientific evidence in alcohol testing cases to be encouraging. It suggests that even if we accept that courts adopt a contextual approach to truth, under the right circumstances courts are prepared to exclude less probative evidence. In my estimation, the key difference between many areas of forensic evidence and alcohol testing is that in the latter case the courts are able to ascertain differences in the quality of evidence. Thus, any steps that assist courts to recognize differences in forensic evidence are steps in the right direction. This means that blind proficiency testing on samples of varying quality is an essential first step. The alcohol-testing arena suggests additional ways to achieve this objective.

The drunk driving example supports the idea that an important first step in improving forensic evidence is to routinize procedures and adopt specific protocols. This does not mean that courts should automatically exclude testimony that does not follow protocols. They may well conclude, as do many drunk-driving cases, that shortcomings go to weight. But the failure to follow protocol does alert courts to the need for special attention to the admissibility of expert testimony.

Similarly, the drunk driving example suggests that it is better to perform tests mechanistically. The advantage of such procedures is that they are more likely to make differences in diagnostic quality transparent. This transparency, in turn, allows courts to more easily distinguish between more probative and less probative evidence. This ability is so important that it might justify a small diminution in overall quality to create greater transparency. However, it is not clear that individual examiner judgments

104. See *State v. Bartlett*, 502 S.E.2d 53, 55 (N.C. Ct. App. 1998); *Verbois v. State*, 909 S.W.2d 140 (Tex. App. 1995). With continued improvements in the accuracy of alcohol screening test equipment, such challenges seem to be less frequent.

105. See *Horn*, 185 F. Supp. 2d at 556; *State v. Superior Court*, 718 P.2d 171, 181 (Ariz. 1986); *State v. Garrett*, 811 P.2d 488, 491 (Idaho 1991) ("HGN test results may not be used at trial to establish the defendant's blood alcohol level in the absence of the chemical analysis of the defendant's blood, breath, or urine."); *State v. Bresson*, 54 N.E.2d 1330, 1335-36 (Ohio 1990) (stating that an officer may not give an opinion as to the driver's actual BAC); *Burkett v. State*, 179 S.W.3d 18, 34 (Tex. App. 2005) (holding that HGN results cannot support officer testimony that the defendant is drunk beyond the legal BAC limit).

Limitations on the scope of testimony have, in fact, been used in some forensic areas such as handwriting and firearm examinations. See Paul C. Giannelli, *The NRC Report and Its Implications for Criminal Litigation*, 50 JURIMETRICS J. 53, 61-62 (2009).

2010]

“*UTTERLY INEFFECTIVE*”

569

are, on average, better than judgments based on a computer algorithm. Moreover, when we do rely on mechanistic assessment, the experience in alcohol testing and other areas such as DNA testing suggest that these techniques are likely to improve over time.

CONCLUSION

In this paper, I have offered an explanation as to why courts have been so lenient in admitting forensic evidence in criminal cases. The combination of a contextual approach to knowledge and the serious deficiency of systematic scientific evidence available to answer “the-cause-of-an-effect” questions push courts toward admissibility. However, I do not believe this means we must simply accept the status quo where the courts are “utterly ineffective” in policing forensic testimony. I indicate that this has not been the case in another criminal law area, convictions for drunk driving. I argue the key reason courts are more restrictive in drunk driving cases is that they have tools that allow them to relatively easily distinguish better from worse evidence and I suggest ways the courts might push forensic evidence in the same direction.

These suggestions are aspects of a larger idea. Whenever possible, we should move away from conclusions based on “clinical judgment” and toward conclusions based on more quantified and formal procedures. Other areas have recognized the value of this approach. Evidence-based medicine is premised on this idea.¹⁰⁶

Some in the forensics community are likely to resist these suggestions because they challenge clinical autonomy. Nevertheless, I believe these steps will encourage courts to play a more active role in achieving the twin goals of working toward the best possible forensic evidence and at the same time excluding the most marginal forensic testimony.

106. See David M. Eddy, *Evidence-Based Medicine: A Unified Approach*, 24 HEALTH AFFAIRS 9 (2005).