The Expert Witness: Lessons from the U.S. Experience*

Susan Haack†
shaack@law.miami.edu

ABSTRACT

The first section of this paper explains why assessing the worth of expert testimony poses special epistemological difficulties. The second traces the history of the various rules and procedures by means of which the U.S. legal system has tried to ensure, or at least control, the quality of the expert testimony on which it so often relies–from the *Frye* Rule, the Federal Rules of Evidence, and the *Daubert* trilogy to recent constitutional cases regarding the appearance of forensic witnesses in court and experiments with court-appointed experts and scientific education for judges. The third and final section suggests some lessons to be learned from the limited success of these efforts, and explores what might be better strategies going forward.

Keywords: epistemology; U.S. law on expert witnesses; *Frye* Rule; *Daubert* trilogy; court-appointed experts; *Melendez-Diaz* and its progeny; forensic science.

> In order that we may have the right to accept [another person's] testimony as ground for believing what he says, we must have reasonable grounds for trusting his *veracity*, that he is really trying to speak the truth as he knows it; his *knowledge*, that he has had opportunities of knowing the truth of this matter; and his *judgement*, that he has made proper use of those opportunities in coming to the conclusion which he affirms.
> W. K. Clifford (1877)[1]

> The whole object of the expert is to tell the jury, not facts ..., but general truths derived from his specialized experience. But how can the jury judge between two statements each founded upon an experience confessedly foreign in kind to their own? It is just because they are incompetent for such a task that the expert is necessary at all. [T]his is setting the jury to decide, where doctors disagree
> Judge Learned Hand (1901)[2]

† Distinguished Professor in the Humanities, Cooper Senior Scholar in Arts and Sciences, Professor of Philosophy, Professor of Law, University of Miami (Florida, U.S.A.).

[1] W. K. Clifford, "The Ethics of Belief" (1877), in *The Ethics of Belief and Other Essays*, eds. Leslie Stephen and Sir Frederick Pollock (London: Watts & Co., 1947), 70-96, p. 79.

[2] Learned Hand, "Historical and Practical Considerations Regarding Expert Testimony," *Harvard Law Review* 15 (1901): 40-58, p. 54. (In Hand's text, the sentence I have put last here occurs *before* the rest.)

If you look in the back of the *ABA Journal*, the official publication of the American Bar Association, you will find advertisements offering expert-witness services—in a recent issue: A & A Legal Nurse ("plaintiff or defense"); Independent Lab Testing; Pediatrics Experts; Domestic Violence Researcher; Surety Expert; Franchise Expert Witness; Attorney-Endorsed Medical Experts; Hospital Medical Director; Jail/Prison Medical Director; Emergency Medicine/Medical Toxicology; Emergency Medicine/Trauma; Neurologist ("on faculty of prestigious university"); Neurosurgeon; Accredited Psychiatry & Medicine ("Harvard alumni & faculty"); Nursing Home Medical Director; Medical Expert Available for Social Security Disability Claims; Forensic Accounting.[3] As this suggests, the expert-witness business is booming.

As it also suggests, medical experts seem to be particularly in demand; but there are many, many other kinds of expertise on which attorneys and, sometimes, judges, call. In fact, the sheer variety of experts who play a role in litigation of one kind and another is overwhelming: experts on asbestosis, accident reconstruction, automobile design, the authenticity of works of art; experts on blood spatter, bite-marks, bullets, behavioral analysis, Bendectin;[4] experts on construction techniques, cancers, causation evidence, criminology; experts on DNA, domestic violence, denture adhesive; experts on engineering, economic losses, epidemiology, evolution, eyewitnesses, the valuation of real estate; experts on fingerprints, footprints, forensic document examination, future dangerousness, Fosamax,[5] the design of folding lawn-chairs; ..., etc., etc.—in fact, experts on just about *everything*, all the way through the alphabet.

Moreover, experts appear in cases of almost every kind. In the criminal justice system we encounter not only DNA analysts, fingerprint examiners, specialists on handwriting and documents, tool-mark experts, etc. but also (among many, many others) psychiatrists testifying about Post-Traumatic Stress Syndrome, Battered Woman Syndrome, Rape Trauma Syndrome, Child Sexual Abuse Accommodation Syndrome, etc., and psychologists testifying about the weaknesses of eyewitness testimony and memory. DNA analysts also turn up in, for example, paternity and immigration cases, and handwriting and document specialists in cases of contested wills. Epidemiologists, toxicologists, experts on occupational safety, representatives of virtually every

[3] *ABA Journal*, November 2014, pp. 68-69.

[4] A drug for the treatment of morning-sickness in pregnancy, alleged to cause limb-reduction birth defects. See p. 50 and note 39 below.

[5] A drug for the treatment of osteoporosis, alleged to cause osteonecrosis of the jaw in some patients.

conceivable medical specialty, even experts on "weight of evidence methodology," show up in toxic-tort and medical-malpractice cases. Historians of art, computer specialists, forensic accountants, and, once again, those forensic document examiners, show up in fraud cases. In *Brown v. Board of Education*, the landmark civil-rights case in which the Supreme Court held that segregated "separate but [supposedly] equal" schools for minority children are unconstitutional, the appellants' brief included an appendix summarizing a fact-finding report at the White House Conference on Children and Youth "bringing together the available social science and psychological studies related to ... how racial and religious prejudices influence the development of a healthy personality."[6] And in other constitutional cases we encounter (among many, many others) professors of religion testifying as to whether being obliged to attend public high schools after the age of fourteen would cause psychological damage to Old Order Amish adolescents,[7] specialists in biology, paleontology, biochemistry, sociology, and philosophy testifying as to whether Intelligent Design Theory is a scientific theory or a theological one, even a professor of theology testifying as to whether it's a good theological theory or a bad one,[8] ..., and so on, again almost without limit.

Heavily as it has come to depend on them, however, the U.S. legal system has always found expert witnesses problematic; and from the beginning there have been complaints about how readily such witnesses conform their opinions to the interests of the party that hires them, and how often, rather than clarifying the factual matters at issue in a case, they confuse or obscure them.

My purpose here is, first—giving the epistemological backdrop—to explain why expert witnesses pose special difficulties (§1); second—sketching the very complex legal history—to describe the various rules and procedures by means of which the U.S. legal system has tried to ensure, or at least control, the quality of the expert testimony on which it so often relies (§2); and finally, to suggest some lessons to be learned from the limited success of these efforts, and to explore what might be better strategies going forward (§3).

[6] Appellants Brief, Brown v. Bd. of Educ. of Topeka, Kan., 347 U.S. 483 (1954), WL 47265 (1952), section II, *8-*13.

[7] Wisconsin v. Yoder, 406 U.S. 205, 210-13 (1972).

[8] Edward Humes, *Monkey Girl: Evolution, Education, Religion, and the Battle for America's Soul* (New York: HarperCollins, 2007), 271-72. Kitzmiller v. Dover Area Sch. Dist., 400 F. Supp. 2d. 707 (M.D. Pa. 2005).

1. The Epistemological Backdrop: Weathering A Perfect Storm

A core epistemological question—I might even say, *the* core epistemological question—is how to assess the worth of evidence.[9] This question encompasses a series of nested sub-questions: among them, how to determine in an effective, reasonable way whether (or to what degree)[10] what another person tells us is credible. This in turn encompasses further sub-questions: among them, how to determine in an effective, reasonable way whether what an expert (or purported expert) tells us on specialized matters of which we have none but the vaguest and most general knowledge ourselves is credible. And this in turn encompasses yet further sub-questions: among them, the question mot relevant here—*how to determine in an effective, reasonable way whether what we are told by an expert witness in a legal context is credible*.

This already suggests a preliminary explanation of why expert witnesses have proven so problematic: assessing the worth of testimonial evidence always involves a kind of indirection; assessing the worth of specialized and technical evidence requires substantive knowledge of relevant facts; assessing the worth of testimony presented in court involves not only taking account of how what's said may be skewed by legal constraints or by the interests of the parties, but also resisting the temptation to allow the horrific nature of a crime or the terrible injury suffered by a plaintiff to skew our appraisal of the evidence that the responsibility falls on this person or that company. And when a juror, or a judge, has to assess the credibility of an expert witness in a legal setting, he encounters *all* of these problems at once—facing, as the sub-title of this section signals, a perfect epistemological storm.

The remarkable paper of W. K. Clifford's from which I took my first opening quotation, though often read simply as an epistemological critique of religious belief,[11] is also an important, though seldom-acknowledged, contribution to what would nowadays be called "social epistemology." And the short passage I quoted identifies the core of what's involved in deciding whether what another person tells you can be trusted: making a sound

[9] See, e.g., Susan Haack, *Evidence and Inquiry* (1993: second, expanded ed., Amherst, NY: Prometheus Books, 2009); "Epistemologia: Chi Ne Ha Bisogno?" *Epistemologia*, XXXIV, 2011: 268-88.

[10] I won't keep repeating this qualification, but it should be understood as implicit in what follows.

[11] In part, no doubt, because this was the paper to which William James was responding in an even more famous paper, "The Will to Believe." William James, "The Will to Believe" (1897), in Frederick Burkhardt and Fredson Bowers, eds., *The Will to Believe and Other Essays in Popular Philosophy* (Cambridge, MA: Harvard University Press, 1979), 13-33.

assessment of (i) his truthfulness (his "veracity," as Clifford says), and (ii) his competence on the matter in question (his "knowledge" and his "judgement").[12] When we ask for directions, for example, we grow skeptical if we notice that our informant hesitates and changes his mind ("go right at the traffic lights—oh, no, sorry, go left—oops, no, silly me, right"), or has a lengthy disagreement with his friends in their own language before he finally tells us to take the no. 7 tram and get off at the fifth stop; and if we're wise, we bear in mind that in certain parts of the world *machismo* may demand that our informant give us directions whether or not he knows the way!

Deciding whether what an expert tells us on some specialized matter can be trusted is significantly *more* difficult. Why so? Our knowledge, and our experience, is limited. That's why we ask the advice of a doctor about the treatment of these symptoms, a plumber about the cause of this leak, a financial advisor about the risks involved in this investment—because we take ourselves to be in need of their expertise. And when we have virtually *no* relevant knowledge or experience of our own, we have to rely almost entirely on experts' knowledge and experience. Of course, if we're wise, we'll take steps to find out whether the plumber telling us we need a whole new drain field makes most of his money replacing such systems; we'll seek a second medical opinion; we'll check out several financial advisors; and so on. If we're lucky, these common-sense precautions will be enough; but, as we all know to our cost, they won't always be.

When an eye-witness describes an accident or a crime, a reasonable juror will try to assess his truthfulness by looking to his demeanor and thinking about his motivation: is he matter-of-fact, or evasive? Is he hesitant, or perhaps too emphatically confident? Does he have reason to lie? And he will try to assess the witness's competence: was he actually present at the scene, and in a position to see? Was the light was adequate? Is he short-sighted, and if so, he was wearing his glasses? Had he seen the defendant elsewhere, or previously identified another person as the perpetrator? But the legal context introduces further complications: e.g., was the witness induced to testify in return for some concession on his own case? Was the police line-up conducted or the photo-array presented in such a way as to encourage him to identify the suspect as the person he saw? And so on.

[12] As I have argued elsewhere, however, Clifford's account of when it's reasonable to believe what a scientist tells us is somewhat naïve. Susan Haack, "Credulity and Circumspection: Epistemological Character and the Ethics of Belief," forthcoming in *Proceedings of the American Catholic Philosophical Association* (2015).

No wonder, then, that deciding whether an expert witness can be trusted is more difficult yet. As Judge Learned Hand said in the celebrated paper from which the second of my opening quotations was taken, in the legal context we will likely be faced by the *opposite* opinions of *competing* expert witnesses—any one of whom is better-equipped and better-qualified to form an opinion on the contested matter than we are. Of course, an expert witness won't just offer a bare opinion, but will normally explain what the evidence is on the basis of which he arrived at that opinion. But that evidence will often be couched in a technical vocabulary that we can, at best, only partially understand; and will, moreover, often rely on background factual assumptions the truth of which we can't judge for ourselves. Even to know what evidence is relevant to a claim, after all, let alone to judge how strong or how weak that evidence is, *requires substantive knowledge of the subject-matter*.[13] Is the fact that a child has this mitochondrial disorder relevant to whether she is especially susceptible to a bad reaction to the MMR (mumps, measles, and rubella) vaccine?[14] Is the fact that the concrete used in building this parking structure has this composition rather than that relevant to why it collapsed? Is the fact that this DNA sample from the crime-scene matches that sample from the defendant at 13 loci sufficient to establish that the chance that the match is random is a one in a million? Etc., etc.

In short:

- Determining whether or to what degree it's reasonable to believe what another person tells us is always in some degree indirect, involving reliance (implicit or explicit) on surrogate indicators of his truthfulness and his competence.
- Sometimes even this poses considerable difficulties; but determining whether or to what degree it's reasonable to believe what an expert tells us tends to be *more* indirect, and so *more* difficult, because it's harder to identify such indicators.
- Determining whether or to what degree it's reasonable to believe what a witness says in court involves *further* complications.
- And determining whether or to what degree to believe an expert witness in a legal context is *even more* difficult, because:

[13] As I argue in, e.g., Susan Haack, "Legal Probabilism: An Epistemological Dissent" (first published, in Spanish, in 2013) in Susan Haack, *Evidence Matters: Science, Proof, and Truth in the Law (*New York: Cambridge University Press, 2014), 47-77, p.61.

[14] See. e.g., Poling *ex rel.* Poling v. Secretary of Health & Human Services (2008), No. 02-1466 V, *1.

(i) There is likely to be competing testimony from an expert witness, or witnesses, on the other side; and a lay juror (or judge) is likely to understand the contested factual issues only, at best, in part.
(ii) Expert witnesses are not likely to volunteer information that might be damaging to the party by which they were hired, even if this information would be helpful to the fact-finder.[15]
(iii) Moreover, legal cases often involve ugly crimes or grave injuries evoking powerful emotions, which can impede jurors' (or judges') assessment of evidence.

It's hardly surprising, then, that modern legal systems struggle to handle expert witnesses effectively; nor that in the recent history of the U.S. legal system there have been many efforts to ensure that, rather than being a hindrance or an impediment, expert testimony helps juries reach factually correct, substantially just verdicts

2. The U.S. Experience: Tracing a Tortuous Path

The conventional contrast between "adversarial" and "inquisitorial" systems is too crude to capture what is really a complex mesh of differences, and commonalities. Still, for readers unfamiliar with common-law procedures, I'll start with a brief description of some key characteristics of the U.S. system, beginning with the stress on precedents, i.e., on decisions in earlier cases, and the division of labor between the judge, charged with determining questions of law, and the "fact-finder," normally a jury, charged with determining questions of fact. Most to the present purpose: witnesses, including expert witnesses, are prepared and presented by the parties, and subject to cross-examination by the other side;[16] and a whole battery of rules makes certain kinds of relevant

[15] As this reveals, we depend (usually implicitly) on informants' good will, specifically, their intent to be informative, as well as on their truthfulness and their competence.

[16] By contrast, I understand, in Italy experts—who may be consulted either on the court's initiative or at the request of a party—must be selected from a register of technical consultants kept at each tribunal, and are considered, not witnesses, but auxiliaries of the court. Mauro Cappelletti and Joseph M. Perillo, *Civil Procedure in Italy* (The Hague, the Netherlands: Martinus Nijhoff, 1965), pp. 230 ff. A more recent source tells me that "the new Italian system .... has retained the system of official experts—whilst giving the parties the right to designate an expert of their own, who can check the work of the official expert, and be heard with him or against him at trial." J. R. Spencer, "Evidence," in Mireille Delmas-Marty and J. R. Spencer, eds., *European Criminal Procedures* (Cambridge: Cambridge University Press, 2002), 641-87, p. 634. The

evidence (for example, illegally-obtained evidence, and any further evidence obtained as a result of such evidence) inadmissible—meaning that the jury should not hear it or, if they hear it anyway, should be instructed by the judge to put it out of their minds. The admissibility or inadmissibility of evidence is a legal question, and hence the province of the judge; the weight of evidence, its sufficiency or insufficiency to meet the standard of proof, is a factual question, and hence the province of the finder of fact. Proffered evidence will be excluded if the opposing party challenges it under the evidentiary rules, and the judge upholds the exclusion.[17]

The world, and the U.S. legal system, have changed significantly since Hand wrote—1901, the year before fingerprint evidence was first used in a criminal case.[18] Now, as then, medical experts of one kind and another turn up all the time; but by now, as we saw, the legal system also calls on a vast range of other kinds of expertise. Moreover, by now relatively few cases go to a jury; indeed, relatively few cases ever go to trial—the great majority of criminal cases are plea-bargained, the great majority of civil cases settled.[19]

Most to the immediate purpose, at the time of Hand's article the distinctive characteristic of an expert witness was that he was exempt from the "opinion rule," under which ordinary, lay witnesses were required to confine themselves to reporting their perceptions, and not permitted to offer their opinions;[20] and all that was necessary for the testimony of an expert to be admissible was that the witness be suitably qualified, and his proffered testimony relevant to some

reference to the "new" Italian system seems to be to the 1988 Code of Criminal Procedure. See Antionette Perrodet, supervised by Mario Chiavario, revised by Elena Ricci, "The Italian System," in the same volume, 248-414, pp. 349 and 379. (There have been several amendments to the code since 1988 but, so far as I know, none is relevant to what is said in this note.)

[17] Moreover, except in cases of egregious legal error, evidentiary determinations cannot be appealed unless they were previously challenged at trial. The U.S. evidentiary regime is, in short, thoroughly adversarial.

[18] Michael Kurland, *Irrefutable Evidence: Adventures in the History of Forensic Science* (Chicago: Ivan Dee, 2009), p. 93.

[19] See, e.g., Marc Galanter, "The Vanishing Trial: An Examination of Trials and Related Matters in Federal and State Courts," *Journal of Empirical Legal Studies* 1, no.3 (November 2004): 459-70; Marc Galanter and Angela M. Frozena, "A Grin without a Cat: The Continuing Decline and Displacement of Trials in American Courts," *Daedalus* 140, no.3 (Summer 2014): 115-28.

[20] This is why—taking the crucial point to be that an expert witness, unlike a lay witness, isn't confined to testifying to his experience but may give his opinion—Hand tells us that the first case he can find of "real expert testimony," i.e., of the conclusions of skilled persons being submitted to the jury, was *Alsop v. Bowtrell* (1620); where physicians testified that a child born to a woman "forty weeks and nine days" after her husband died might well be his child. Hand, "Historical and Practical Considerations" (note 2 above), pp. 46-47.

fact at issue in the case. By now, the opinion rule has been relaxed somewhat;[21] and legally speaking, the distinctive characteristic of an expert witness is simply that he has specialized knowledge, skill, education, or training beyond that of the average juror.[22] Expert witnesses are still given much more latitude in expressing their opinions than lay witnesses are; but, as we'll soon see, the requirements for the admissibility of their testimony are now significantly more complex and demanding than they were a century ago.

(i) *Judicial Screening of Expert Testimony: The* Frye *Rule*

This part of the story begins with *Frye v. United States*,[23] a 1923 murder case in which, for the first time, a court placed restrictions not only on *the qualifications of a proffered expert witness*, but also on *the content of proffered expert-witness testimony*. James Alfonso Frye was accused of murdering a physician. He had confessed; but subsequently withdrew his confession. At trial, his attorney had proffered an expert who would testify that Frye had been subjected to a (then very new) blood-pressure deception test,[24] which allegedly showed that he was telling the truth when he claimed that his confession had been false– he *hadn't* killed Dr. Brown. This proffer was denied; and, in a very short, citation-free ruling, arguing that the technique underlying the proffered testimony was just too new to have established its evidentiary credentials, the D.C. Court of Appeals upheld the trial court's decision. The key passage reads:

> Just when a scientific principle or discovery crosses the line between the experimental and demonstrative stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the

[21] Federal Rule of Evidence 701 allows a lay witness to testify to opinions or inferences "rationally based on [his] perception," provided that these inferences are *not* based on scientific, specialized, or other technical knowledge.

[22] How to apply this very pragmatic conception of an expert has, however, sometimes been controversial. In *Downing*, for example, the trial court had denied the defendant's proffered (supposedly) expert testimony about the unreliability of eyewitnesses on the grounds that jurors would already know that eyewitnesses aren't very reliable: a claim that seems, at least in light of what we know now, dubious in the extreme. United States v. Downing, 735 F.2d 1224 (1985). Hal Arkowitz and Scott O. Lillefeld, "Why Science Tells Us Not to Rely on Eyewitnesses," *Scientific American*, 1.8.2009, available at http://www.scientificamerican.com/article/do-the-eyes-have-it/1 (last visited 9.1.15).

[23] Frye v. United States, 293 F. 1013 (D.C. 1923).

[24] The test–much simpler than a modern polygraph–measured changes in the subject's systolic blood-pressure under questioning. *Id.*, 1013.

> deduction is made must be *sufficiently established to have gained general acceptance in the field in which it belongs*.[25]

How is a judge to tell when a scientific principle or discovery has crossed the line between the "experimental" and the "demonstrative" stages? Judge Van Ordsel declines to say.[26] How is a judge to identify the "principle or discovery" on which proffered scientific testimony is based? Again, he doesn't say. And neither does he tell us how large a majority of those in a field must accept an idea if it is to qualify as "generally" accepted, or explain how fields are to be identified or individuated.

It's clear, however, that the "*Frye* Rule" (as it came to be known) is *conservative in intent*, i.e., meant to exclude as-yet untested, highly speculative scientific ideas; and that it does this by *deferring to the judgment of those in the field concerned*. It's also clear that it's very flexible—not to say easily manipulated. In particular, while *Frye* can be quite demanding if "the field to which [novel scientific testimony] belongs" is construed broadly, it can be very easily satisfied if the field is construed narrowly. This flexibility probably partly explains why, in due course, *Frye* proved so attractive.

"In due course" because, for many decades, *Frye* was rarely cited, and when it was, was usually construed as a precedent for excluding lie-detector evidence.[27] But over the years its influence grew. In fact, by 1975, when Congress ratified the Federal Rules of Evidence (FRE) to codify existing evidentiary practice, some version of the *Frye* Rule—usually in the abbreviated form of a requirement that novel scientific testimony be generally accepted it its field—was accepted in many jurisdictions, and construed, not just as excluding polygraph evidence, but as restricting the admissibility of novel scientific evidence of whatever kind.[28]

[25] *Frye* (note 23 above), 1014 (my italics).

[26] *Frye* is usually taken to apply only to "novel" scientific testimony. This can't mean simply that the idea or technique at issue is completely new *to the relevant scientific community*, since by definition such evidence couldn't have gained *any* acceptance in its field; nor can it mean simply that the idea or technique is new *to the legal system*, since this would mean that it could never get its foot in the legal door in the first place. It must mean, I assume, that scientific testimony should be excluded until it has gained acceptance in its field but, after that, be admitted as no longer "experimental" or novel.

[27] Kaminski v. State, 63 So. 2d339 (Fla.1953). On the history of the *Frye* Rule, see Paul C. Giannelli, "The Admissibility of Novel Scientific Evidence: *Frye v. United State*s, a Half-Century Later," *Columbia Law Review* 80 1980): 1197-1250.

[28] "While some courts have rejected the general acceptance standard, there remains considerable support for the *Frye* test." *Id.*, 1228.

(ii) *Judicial Screening of Expert Testimony: The Federal Rules of Evidence*

Section 7 of the FRE concerned opinion evidence generally. FRE 702 governed the admissibility of *expert* opinion testimony—whether scientific or not, whether novel or not. It made no mention of *Frye*, or of "general acceptance," but simply provided that an expert "qualified by specialized knowledge, education, skill, or training" might testify in the form of an opinion if his testimony would be helpful to the fact-finder, and was not otherwise legally excluded. Did this mean that FRE 702 had superseded *Frye*, at least in federal jurisdictions? Courts disagreed. Was FRE 702 best construed as requiring, simply, that such testimony be relevant to facts at issue in the case, or did being helpful to the fact-finder implicitly require something more, that expert testimony have some degree of reliability? Again, courts disagreed.

In *Barefoot* (1983),[29] for example, the Supreme Court brushed aside the suggestion that the defendant's constitutional rights were violated when psychiatric testimony that he would be dangerous in future was admitted—even though there was reason to believe that such predictions were wrong more often than they were right.[30] Both state and federal law, Justice White argued for the majority, anticipate that relevant evidence will be admitted, and its weight left to the jury to decide;[31] moreover, the defense had had the opportunity to challenge the reliability of the contested psychiatric testimony through cross-examination, and could have presented contrary witnesses.[32] But in *Downing* (1984) a federal court of appeals ruled that the lower court had erred in excluding psychological testimony about the factors that influence whether, and when, the testimony of an eyewitness is likely mistaken; and argued explicitly that judges should screen expert testimony not only for relevance, but also for reliability.[33]

By the early 1990s, the status of *Frye* under the new Federal Rules was still unresolved. Amid pressure for tort reform and complaints that bad science was flooding the courts, Peter Huber argued in *Galileo's Revenge*[34] that, while the

[29] Barefoot v. Estelle, 446 U.S. 880 (1983).
[30] *Id.*, 883, 898, 899.
[31] *Id,* 898. The testimony was undeniably relevant, since the Texas death-penalty statute required that, at the sentencing phase, jurors determine beyond a reasonable doubt that there is a probability that the defendant would be dangerous in future.
[32] *Id.*, 899.
[33] *Downing* (note 24 above).
[34] Peter Huber, *Galileo's Revenge: Junk Science in the Courtroom* (New York: Basic Books, 1993).

old *Frye* Rule had served to keep "junk science"[35] out, the more liberal Federal Rules of Evidence had opened the floodgates. Huber's legal history was all wrong; in point of fact the *Frye* Rule had virtually always been confined to criminal cases.[36] Nevertheless, his book struck a chord with those clamoring for reform of the tort system; and by 1991 then-Vice-President Dan Quayle was proposing, *inter alia*, adding a requirement of "widespread acceptance" in the field to FRE 702.[37]

In 1993, however, the Supreme Court stepped in, giving its first-ever ruling on the standard of admissibility of expert testimony, the status of the *Frye* Rule in federal jurisdictions, and the interpretation of FRE 702: *Daubert v. Merrell Dow Pharmaceuticals*.[38]

(iii) *Judicial Screening of Expert Testimony: The* Daubert *Trilogy*

Daubert was just one of many toxic-tort cases involving the morning-sickness drug Bendectin, which was alleged to cause limb-reduction birth defects in some of the babies born to women who took it.[39] Legally, however, *Daubert* was distinctive: it was a very rare instance of a civil case where the trial court had referred to "general acceptance," and the Court of Appeals had specifically cited *Frye*;[40] and hence provided the perfect opportunity to clarify the legal situation. FRE 702 *had* superseded *Frye,* Justice Blackmun wrote for a unanimous Supreme Court; but federal courts' responsibility to screen proffered exert testimony remained.[41]

This screening, Justice Blackmun continued—but now only on behalf of the majority—should ensure that expert testimony admitted is both relevant *and reliable*.[42] And how are courts to screen for reliability? The text of FRE 702 refers to "scientific, specialized, or other technical knowledge." But the

[35] By analogy, I assume, with "junk food," such as the burger that looks and smells like real food, but has no nutritional value. I don't know where the phrase originated, only that apparently it was Huber who made it famous.
[36] Kenneth J. Cheseboro, "Galileo's Retort: Peter Huber's Junk Scholarship," *American University Law Review* 42 (1993): 1637-1726.
[37] See Diana Culp Bork, "Reasonable Tort Reform," *The National Law Journal*, 14 (September 30, 1991): pp. 17, 18, 21.
[38] Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579 (1993) ("*Daubert* 1993").
[39] See e.g., Joseph Sanders, *Bendectin on Trial: A Study in Mass Tort Litigation* (Ann Arbor, MI: University of Michigan Press, 1996). Bendectin was withdrawn from the U.S. market in 1984; according to the manufacturers, this was because of the costs of litigation, not because the drug posed any real danger. It returned to the U.S. market (now under a new name, Diclegis, and made by a Canadian company) in 2013.
[40] *Daubert* 1993 (note 38 above), 583-84.
[41] *Id.*, 589.
[42] *Id.*, 591, citing *Downing* (note 22 above).

testimony at issue in *Daubert* was, specifically, scientific (epidemiological, toxicological, etc.). So, quietly dropping the phrase "technical or other specialized" from the text of the Rule, Justice Blackmun argued that such testimony should be genuine "scientific ... knowledge": i.e., really knowledge, not mere opinion,[43] and genuinely scientific. Being genuinely scientific, he continued, means arriving at your conclusions by the scientific method.[44] So courts should look, not to the conclusions an expert draws, but exclusively to the "methodology" he uses in arriving at those conclusions.[45]

In determining "evidentiary reliability," Justice Blackmun added, courts might consider these indicia of reliability (soon known as the "*Daubert* factors"):

- whether the theory or technique can be (and has been) tested;
- whether the theory or technique has been subjected to peer review and publication;
- the known or potential rate of error, and the existence and maintenance of standards controlling the operation of the technique in question;
- whether the theory or technique has gained widespread acceptance in its field[46]

The first of these—a result of Justice Blackmun's unfortunate confusion of "reliable" and "scientific"[47]—reflects a half-understood, quasi-Popperian misconception of the supposed "scientific method."[48] The second in part reflects the mistaken idea that peer-reviewed publication is a sign of the widespread acceptance mentioned in the fourth,[49] which is in turn a nod to the old *Frye* Rule. The third—though it looks potentially helpful with respect to, e.g., questionable forensic identification techniques, or those predictions of

[43] *Id.*, 589-90. Not surprisingly, potential scientific witnesses whose testimony is excluded by a judge who deems it not really scientific—the phrase is "Dauberted out"—are sometimes indignant at what they perceive as an insult to their professionalism.
[44] *Id.*, 590.
[45] *Id.*, 592-93.
[46] *Id.*, 593-95.
[47] See Susan Haack, "Trial and Error: Two Confusions in *Daubert*" (2005), in Haack, *Evidence Matters* (note 13 above), 104-21.
[48] See Susan Haack, "Federal Philosophy of Science: A Deconstruction—and a Reconstruction" (2010), in Haack, *Evidence Matters* (note 13 above), 122-55; "Just Say 'No' to Logical Negativism" (first published in Chinese in 2011), in Haack, *Putting Philosophy to Work: Inquiry and Its Place in Culture* (Amherst, NY: Prometheus Books, 2008, expanded second ed., 2013), 179-97 (text) and 298-305 (notes).
[49] See Susan Haack, "Peer Review and Publication: Lessons for Lawyers" (2007), in Haack, *Evidence Matters* (note 13 above), 156-79.

future dangerousness–is notably silent on the matter of *what* "known or potential" error-rate might disqualify expert scientific testimony as too unreliable to be admitted.[50]

Rather than make a final determination on the Dauberts' claim, after settling the question of the standard of admissibility the Supreme Court remanded the case to the ninth circuit. And in the course of this final ruling–once again deeming the Dauberts' expert testimony inadmissible, and once again granting summary judgment to the defendant company, Merrell Dow–Judge Kozinski added what is sometimes described as a fifth *Daubert* factor to the list:

- whether the work on which the testimony is based was litigation-driven, or undertaken independently of litigation[51]

–the idea being that scientific work undertaken for the purposes of litigation is inherently less likely to be reliable than work conducted in the normal course of scientific business. But in a footnote Judge Kozinski made an important exception: even though the testimony of forensic scientists is always litigation-driven, "the fact that [such an] expert has developed an expertise primarily for purposes of litigation will obviously not be a substantial consideration."[52] He doesn't say why not.

After *Daubert*, federal judges really did find themselves, as Judge Kozinski had observed, in a "Brave New World."[53] *Daubert* is much broader in scope than *Frye*; and it obliges judges, rather than deferring to the relevant scientific (or other expert) community, *to make determinations about the reliability of scientific expert testimony for themselves*–even though, as then Chief-Justice Rehnquist had pointed out in his partial dissent, they are untrained for such a task. The word "reliability" nowhere occurs in the text of FRE 702, Justice Rehnquist noted; Justice Blackmun's observations about falsifiability were baffling, sure to create confusion in the courts below; moreover, the stress on

[50] It is regrettable, in my opinion, that these dicta have now found their way into the entries under "scientific knowledge," "scientific method," and "falsifiability" in the most recent edition of *Black's Law Dictionary*. Bryan A. Garner, ed., *Black's Law Dictionary* (St. Paul, MN: Thomson Reuters, 10th ed., 2014).
[51] Daubert v. Merrell Dow Pharm., Inc., 43 F.3d 1311 (1995) ("*Daubert* 1995").
[52] *Id.*, 1317, n.5. See also Susan Haack, "What's Wrong with Litigation-Driven Science?" in Haack, *Evidence Matters* (note 13 above), 180-207.
[53] *Daubert* 1995 (note 51 above), 1315.

"scientific method" threatened trouble down the road about whether, and if so how, *Daubert* applied to *non-scientific* expert testimony.[54]

His dissent proved prophetic. Just a few years after constructing the new *Daubert* regime, the Supreme Court began quietly *de*constructing it. The first thing to go, in the second of the Supreme Court's trilogy of cases on expert testimony, *General Electric v. Joiner* (1997), was the distinction between methodology and conclusions that had played a starring role in 1993. The core legal issue in this case was the standard of appellate review of decisions excluding such testimony; which, the Court ruled, remained the same—abuse of discretion—even when, as happened in *Joiner*, excluding one side's expert testimony determined the outcome of the case.[55] Moreover, the ruling continued, the lower court had *not* abused its discretion in excluding the experts Joiner had proffered to show that his occupational exposure to PCBs[56] had promoted his lung cancer.[57]

But Joiner's attorneys had argued that their experts used precisely the *same* methodology that G.E.'s experts used—"weight of evidence" methodology—so that the lower court must have looked beyond their experts' methodology to their conclusions; which, under *Daubert*, was legally an error. Sidestepping this argument, Justice Rehnquist wrote for the majority that a court may legitimately conclude that "there is simply too great an analytical gap between the data and the opinions offered." He doesn't tell us what an "analytical gap" is, nor how courts are to judge when an analytical gap is "too great." But he does add, firmly: "[C]onclusions and methodology are not entirely distinct from one another."[58] As Justice Stevens observed in his partial dissent, this was already a significant shift away from *Daubert*.[59]

And, precisely as Justice Rehnquist had predicted, Justice Blackmun's dicta about "scientific ... knowledge" left federal courts puzzling over whether *Daubert* applied to non-scientific exert testimony and, if it did, whether those *Daubert* factors also applied. As one judge put it, federal courts were "balkanized":[60] some held that both *Daubert* and the *Daubert* factors applied

[54] *Daubert* 1993 (note 38 above), 598-601 (Justice Rehnquist, dissenting in part).
[55] Gen. Elec. Co. v. Joiner, 522 U.S. 136 (1997).
[56] Polychlorinated biphenyls (a class of man-made organic compounds). The production and sale of PCBs has been banned in the U.S. since 1977, after they were found to be seriously carcinogenic.
[57] *Joiner* (note 55 above), 143.
[58] *Id.*, 146.
[59] *Id.*, 151 (Justice Stevens, dissenting in part).
[60] Moore v. Ashland Chemical Inc., 152 F.3d 269, 280 (5th Cir. 1998) (Judge Dennis, dissenting).

to all expert testimony, scientific or otherwise;[61] some that *Daubert* and *a fortiori*, the *Daubert* factors, applied only to scientific testimony.[62] One court conducted a *Daubert* hearing as a result of which it determined that, since forensic document examination testimony wasn't science, *Daubert* didn't apply;[63] another seemed to fall into hopeless confusion: "[Mr. Bihlmeyer's testimony] will assist the jurors to understand whether or not there is a design or manufacturing defect involved in this case. So to the extent the *Daubert* case is applicable, it's applicable."[64]

In 1999, stepping in to settle whether, and if so how, *Daubert* applied to non-scientific experts, the Supreme Court continued the deconstruction of *Daubert* begun two years earlier in *Joiner*. The specialized evidence at issue in *Kumho Tire Co. v. Carmichael*[65]—the testimony of an expert on motor tires that the blowout that caused Mr. Carmichael's accident was the result of faulty design, not of the tire's having been abused—was by no stretch of the imagination scientific. The lower courts had been divided, with the trial court holding that this evidence was inadmissible because it flunked all the *Daubert* factors, and the appeals court reversing on the grounds that *Daubert* only applied to scientific testimony.[66] The Supreme Court split the difference: *Daubert* applied here, as to all expert testimony; those *Daubert* factors, however, might or might not be relevant, depending on the nature of the expert testimony in question.

Justice Breyer wrote for a (nearly)[67] unanimous Court:

> [W]e can neither rule out nor rule in, for all cases and all time, the applicability of the factors mentioned in Daubert, nor can we now do so for subsets of cases categorized by category of expert or kind of evidence. [...] Indeed, those factors do not necessarily apply even in every instance in which the reliability of scientific testimony is challenged. It might not be surprising in a particular case, for example, that a claim made by a scientific witness has never been the subject of peer review, for the particular application of it at

[61] See. e.g., Berry v. City of Detroit, 25 F.3d 1342 (6th Cir. 1994).
[62] See, e.g., Carmichael v. Samyang Tires, Inc., 131 F.3d 1433 (11th Cir. 1997).
[63] United States v. Starzecpyzel, 880 F. Supp. 1027 (S.D.N.Y. 1995).
[64] Compton v. Subaru of American Inc., 83 F.3d 1515, 1517 (10th Cir. 1996) (citing the District Court's ruling).
[65] Kumho Tire Co., Ltd. v. Carmichael, 525 U.S. 137 (1999).
[66] Carmichael v. Samyang Tires, Inc. 923 F. Supp. 1514 (S.D. Ala. 1996). *Samynang Tires, Inc., v. Carmichael* (note 62 above).
[67] There are two brief partial dissents, neither of which is relevant to present concerns. *Kumho Tire* (note 65 above), 158-59.

> issue may never previously have interested any scientist. Nor, on the other hand, does the presence of Daubert's general acceptance factor help show that an expert's testimony is reliable when the discipline itself lacks reliability... .[68]

He says nothing, however, about how a judge is to determine whether a discipline has, or lacks, reliability. And what, in the end, does this key passage tell federal judges about how to assess the reliability of proffered expert testimony?—That they should use any, all, or none of the *Daubert* factors, and/or such other factors as they deem appropriate; i.e., that they should use their judgment, and do the right thing.

It's true, and important, that what matters isn't (as Justice Blackmun had suggested) whether expert testimony is scientific, but whether it's really knowledge.[69] It's true, and important, that the sheer variety of kinds of expertise with which courts may have to deal makes it impossible to give indicia of reliability that will work for any and every kind of expert testimony. And it's true, and important, that widespread acceptance of an idea, theory, or technique in some field is no indication of its reliability unless the field itself is legitimate. All that said, however, the fact remains that, since this third ruling in the *Daubert* trilogy, federal courts are left with large responsibility and broad discretion in screening expert testimony in all its nearly limitless variety, but little substantive guidance about how to do this.

In 2000, FRE 702 was revised so as to say explicitly what, according to the Supreme Court, it had said implicitly all along. It was "restyled" in 2011,[70] but its content remains the same:

> 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 the product of reliable principles and methods; and
> (d) the expert has reliably applied the principles and methods to the facts of the case.

[68] *Id.*, 150.
[69] *Id.*, 138.
[70] That is, re-written for style but not changed in content.

Arguably, by suggesting that courts need to look not only at the pedigree of the testimony, but also at its application to the case at hand, this makes a small epistemological step forward; but that ritual incantation of "reliable," "reliably," "sufficient," does nothing to guide judges as to how, in the specific, to screen those proffered expert witnesses.

(iv) *Judicial Education*

As Justice Rehnquist had observed, their training hardly prepared federal judges for the new responsibilities imposed on them by *Daubert*. The year after *Daubert*, however, the National Institute of Justice issued the first edition of its *Reference Manual on Scientific Evidence*.[71] A second edition appeared in 2000,[72] and a third in 2011.[73] By now a very substantial volume, the *Reference Manual* brings together chapters by different specialists, both scientific and legal, on various scientific topics (the nature of science, DNA analysis, epidemiology, probability theory, etc., etc.), likely to be encountered at trial. Naturally, the chapters are variable in quality and accessibility to a lay reader; naturally, they occasionally focus more on judicial rulings than on strictly scientific matters. Still, the manual is a useful tool for judges needing to mug up on some scientific topic pertinent to a case—though sometimes, probably, they use it, instead, as a source of authoritative-sounding quotations to bolster the conclusion they would have reached anyway. But of course, even this big brick of a book can't cover *every* scientific topic that might be legally relevant; and it doesn't touch non-scientific kinds of expertise at all.

There have also been various programs offering judicial education on various scientific topics. Some are now defunct: e.g., the "Science for Judges Program," held for several years at Brooklyn College of Law, ended in 2007; the Federal Judicial Center's educational programs, publications, and videos are no longer current on its website; the long-standing program Advanced Science and Technology: Adjudication Resource (ASTAR) lost its funding in 2013. Some are still active: e.g., in April 2015 the Law and Economics Center at George Mason Law School held a conference on Forensic Statistics. Others are in prospect: e.g., at the National Commission on Forensic Science—a joint effort by the Department of Justice and the National Institute of Standards and Technology (part of the Department of Commerce) created in 2013—one

[71] *Reference Manual on Scientific Evidence* (Washington, D.C.: Federal Judicial Center, 1994).
[72] *Reference Manual on Scientific Evidence* (Washington, D.C.: Federal Judicial Center, 2000).
[73] *Reference Manual on Scientific Evidence* (Washington, D.C.: Federal Judicial Center, 2011).

committee focuses on "Training on Science and Law," and will "explore mechanisms for judges, lawyers, and forensic scientists to engage in collaborative training ... "; the fifth planning meeting [!] took place in January 2015. And the newly-created National Courts and Science Institute (NCSI) includes a Center for Basic and Continuing Judicial Science Education that will oversee annual conferences and a judicial certification in "scientific method, tools and measures."[74] But, while there are, and have been, many admirable efforts to improve judges' education in the sciences, these are at best a drop in the bucket—and scientific testimony, remember, is by no means the only kind of expert testimony a judge may need to screen for admissibility.

(v) *Court-Appointed Experts*

Aware (like Justice Rehnquist) that, since *Daubert*, federal judges' new responsibilities for screening expert testimony posed formidable difficulties, Justice Breyer had urged in his concurrence in *Joiner* that they use their powers under FRE 706 to appoint expert witnesses of their own choosing.[75] Some did. But this process proved less straightforward than, perhaps, Justice Breyer anticipated—and *much* less straightforward than Judge Hand, who apparently thought it would handle all the main problems with expert witnesses,[76] imagined it would be.

By now, there have been many experiments with court-appointed experts, the best-known and most ambitious of which was Judge Samuel Pointer's National Science Panel of four scientists charged with sifting through the medical evidence in the thousands of silicone breast-implant cases consolidated to his court. In 1998, the panel reported that that there was no evidence that, as the plaintiffs claimed, these implants caused connective-tissue diseases. But Judge Pointer's experience revealed all too clearly just how hard it is to identify competent experts who have no connection to one or other of the parties. Indeed, despite all his efforts to ensure neutrality, *all* the panel

[74] See generally, John Cher, "Judges and Lawyers Work to Understand Courtroom Science," *New York Law Journal* 252, no.3 (July 2014), p.1, column 3. On the program at George Mason, see George Mason University Law and Economics Center, "The Value of Judicial Training in Quantitative and Scientific Methods," Draft White Paper, April 8, 2013, available at www.masonlec.org; on the NCSI, see http://ncsi.institute; on the Federal Judicial Center's programs, see www.fjc.gov. On the National Commission on Forensic Science, see p. 61 and note 101 below.

[75] *Joiner* (note 55 above), 149 (Justice Breyer, concurring). FRE 706 reads in part: "The court may appoint any expert that the parties agree on and may appoint expert witnesses of its own selection."

[76] Hand, "Historical and Practical Considerations" (note 2 above), pp. 56-57.

members turned out to have some such indirect connections; and one, a Canadian rheumatologist—who had been chosen in part because the U.S. Association of Rheumatologists had already taken a position on the safety of these implants—was found, while serving on the panel, to have solicited, and received, funds from one of the defendant companies to support a professional conference.[77]

The same year, with much less fanfare and at much more modest cost, Judge Robert E. Jones had appointed expert advisors to help him assess the expert evidence in the much smaller number of silicone breast-implant cases Judge Pointer had returned to his district for trial; after which he excluded all testimony to the effect that the implants cause connective-tissue disorders.[78] But this wasn't entirely reassuring, either, given that—unlike Judge Pointer, who had kept the selection process at arms' length by having a team of advisors choose his panel members—Judge Jones had asked his medical-scientist cousin to help him choose suitable experts to advise him; moreover, all but one of his experts came from the same university.[79] On top of which, in the end, Judge Jones simply disregarded the opinion of one of the experts he had himself appointed.[80]

We have also learned—as should have been obvious from the beginning—that, since the science involved in litigation is often controversial, there is absolutely no guarantee that court-appointed experts won't disagree among themselves.[81] None of this is to deny that, in some circumstances, the use of

[77] The conference was on an unrelated topic. However, the plaintiffs moved that Dr. Tugwell's appointment to the panel be vacated, and the panel's report be withdrawn; but Judge Pointer denied this motion. See: "Breast implant plaintiffs say scientific panel was tainted," CNN Interactive (April 13, 1999), available at http://www.cnn.com/HEALTH/9904l3/breast.implant.panel (last visited February 22, 2015); In re Silicone Gel Breast Implants Products Liab. Litig., MDL 926, Case No. CV 92-P-10000-S, Order 311 (denying plaintiff's "Motion for Relief from Prejudicial Bias"); and, more generally, Laural L. Hooper, Joe S. Cecil, and Thomas E. Willging, "Assessing Causation in Breast Implant Litigation: The Role of Science Panels," *Law and Contemporary Problems* 64, no.4 (Autumn 2001): 139-89, 170-71.

[78] Hall v. Baxter Healthcare Corp., 947 F. Supp. 1387 (D. Ore. 1996). Judge Jones used his authority under FRE 104 (a), appointing these scientists, not as witnesses, but as advisors to the court. See generally Hooper, Cecil, and Willging, "Assessing Causation in Breast Implant Litigation" (note 77 above). Joseph Sanders and D. H. Kaye, "Expert Advice on Silicone Implants: *Hall v. Baxter Healthcare Corp.*," *Jurimetrics* 37 (Winter 1997): 113-28.

[79] Hooper, Cecil, and Willging, "Assessing Causation in Breast Implant Litigation" (note 77 above), p.146-47 and n. 33.

[80] Judge Jones set aside Dr, Greenlick's critique of the idea that a showing of more than doubled risk is necessary for proof of specific causation. See Susan Haack, "Risky Business: Statistical Proof of Specific Causation," in Haack, *Evidence Matters* (note 13 above), 264-93, p. 284.

[81] See, e.g., Soldo v. Sandoz Pharm. Corp., 244 F. Supp. 2d434 (W. D. Pa. 2003).

court-appointed experts may be the best option; but it has not turned out to be the panacea Judge Hand hoped.

(vi) *Confrontation and Cross-Examination*

Justice Blackmun's rhetoric had suggested that *Daubert* was intended to liberalize the standard of admissibility of expert testimony. *Frye* was "an austere standard," he wrote, at odds with the "liberal thrust" of the FEE. [82] But in practice the effect of *Daubert*, *Joiner*, and *Kumho Tire* seems to have been to make the standard of admissibility not less, but more restrictive—at least, in civil cases.

The effect of *Daubert* on criminal cases, however, has been much less:[83] as witnessed, for example, by the numerous failed *Daubert* challenges to fingerprint-identification testimony.[84] And yet there's reason to suspect that some of the forensic sciences (such as hair analysis or bite-mark identification), and some of the psychiatric specialties (such as supposedly "recovered" memories, or predictions of future dangerousness) that have played a significant role in the criminal justice system are distinctly *un*reliable; and even DNA identification evidence, the "gold standard" of forensic science, is susceptible to the same kinds of human error—sloppiness, mislabeling or contamination of samples, confirmation bias, etc.—as other forensic sciences.

But when issues about forensic evidence in criminal cases came to the fore in *Melendez-Diaz* (2009), the Supreme Court's attention was focused, as in *Barefoot*, not on the minutiae of judicial screening for admissibility of expert testimony, but on constitutional matters, specifically on the implications for forensic evidence of the Confrontation Clause of the Sixth Amendment to the U.S. Constitution, guaranteeing defendants the right to confront witnesses against them.

Luis Melendez-Diaz had been convicted of drug-trafficking. At trial, the prosecution had provided three sworn certificates of analysis affirming that the substance the police had seized from him was cocaine. The defendant's objection that the analysts concerned should testify in person was overruled; the appeals court upheld the decision; and the Supreme Judicial Court of

[82] *Daubert* 1993 (note 38 above), 588 ("liberal thrust") and 598 ("austere standard").

[83] As, perhaps, Judge Kozinski's footnote on forensic science in *Daubert* 1995 prefigured. See, e.g., Peter J. Neufeld, "The (Near) Irrelevance of *Daubert* to Criminal Justice and Some Suggestions for Reform," *American Journal of Public Health* 95 (2005): S107-13.

[84] See, e.g., U.S. v Havvard, 117 F.Supp. 2d 848 (S.D. Ind. 2000); U.S. v. Mitchell, 365 F.3d 215 (3rd Cir. 2004).

Massachusetts denied review. The U.S. Supreme Court, however, reversed and remanded, ruling 5-4 that Melendez-Diaz had the right under the Confrontation Clause to have these analysts appear in court to testify and be cross-examined.[85] But four Justices, led by Justice Kennedy, worried that this ruling might have the consequence of requiring *all* the several technicians who may be involved in conducting a single forensic test to appear in court. Since "the defendant does not even dispute the accuracy of the analysts' work, confrontation adds nothing," Justice Kennedy notes; and yet, by obliging forensic scientists to set aside their real work in the laboratory to go testify in court, "for the sake of ... negligible benefits, the Court threatens to disrupt forensic investigations across the country ... ."[86]

Since then, the Supreme Court has been struggling to articulate exactly what makes a forensic report "testimonial" in the legally-relevant sense (and so, subject to the Confrontation Clause), and exactly *which* forensic analysts should be required to testify, and why. A couple of years after *Melendez-Diaz*, in *Bullcoming v. New Mexico* (2011), the judgment of the Court was that the Confrontation Clause required that the technician who actually completed the form reporting the results of the defendant's blood-alcohol test should testify; the testimony of his laboratory supervisor was insufficient.[87] The technician in question, Curtis Caylor, was on unpaid leave; his supervisor didn't know why; and Bullcoming's counsel had had no opportunity to ask questions that might have revealed whether he was removed from his work station for incompetence or dishonesty.[88] Once again, Justice Kennedy dissented. This decision, he argued, went well beyond *Melendez-Diaz*: in this instance an employee of the testing laboratory *did* appear in court to authenticate the findings and be cross-examined.[89] What would the presence of the technician who actually signed the form have added?—after all, the test in question was run on the gas chromatograph overnight, after everyone had already gone home.[90]

But the following year, in the plurality ruling in *Williams v. Illinois*,[91] Justice Alito—who had been with the minority in *Melendez-Diaz* and

[85] Melendez-Diaz v. Massachusetts. 557 U.S 305 (2009). U.S. Constitution, Amendment VI (Confrontation Clause).
[86] *Id.*, 339 ff. (Justice Kennedy, dissenting).
[87] Bullcoming v. New Mexico, 131 S.Ct. 2705 (2011).
[88] *Id.*, 2715.
[89] *Id.*, 2723.
[90] *Id.*, 2724.
[91] Williams v. Illinois, 132 S.Ct. 2221 (2012).

Bullcoming–wrote the judgment of the Court. At Williams' trial for rape, an expert witness had referred to the DNA profile submitted by Cellmark[92] as having been produced from semen found in the victim's vaginal swabs. Justice Alito echoes the language of the rule excluding hearsay evidence,[93] designed to combat the same evil as the Confrontation Clause–reliance on the word of an out-of-court declarant who can't tested under cross-examination: this witness was not testifying as to the *truth* of the claim that the DNA profile was produced from the victim's swabs, he argues; and so did not trigger the requirement that the technician(s) concerned appear in court to testify.[94] And this time it was Justice Kagan–who had argued in support of the decisions in *Melendez-Diaz* and *Bullcoming*–who wrote an impassioned dissent. Focusing on the power of cross-examination to reveal mistakes, incompetence, sloppiness, and dishonesty, she opens with an extraordinary excerpt from the transcript of a trial where a forensic witness realized only after cross-examination that, oh my God, she had the names on the DNA samples mixed up: "I'm a little hysterical right now, but I think .... the two names should be switched."[95]

The story is certainly disturbing. But, recalling that most criminal cases (up to 95% in some jurisdictions)[96] are resolved by plea-bargain and never go to trial, one wonders whether there aren't better ways to avoid this kind of forensic fiasco. An ounce of prevention, as the saying goes, is worth a pound of cure. Shouldn't the priority be to do what we can to ensure that such mistakes don't happen in the first place?

(vi) *Forensic Science Laboratories*

A thought much like this, apparently, motivated another important development the same year as *Melendez-Diaz* when, at the instigation of Congress, the National Research Council (NRC) of the National Academies of Science (NAS) produced a substantial volume entitled *Strengthening Forensic*

[92] Cellmark is a company (founded in 1987) based in the U.K. which also provides forensic DNA testing in the U.S. See http://www.cellmark.foresnics.us/welome-cellmark-forensics (last visited 2.9.15).

[93] FRE 801 (c), defining "hearsay," tells us that a hearsay statement is one that "a party offers in evidence to prove the truth of the matter asserted... ."

[94] *Williams* (note 91 above), 2240.

[95] *Id.*, 2264 (Justice Kagan, dissenting), referring to the rape trial of John Kozack, Tr. in no. SCD 110465 (Super. Ct. San Diego Cty., Cal., Nov. 17 1995).

[96] Murphy, Erin, 2014. "The Mismatch between Twenty-First Century Forensic Evidence and our Antiquated Criminal Justice System," *Southern California Law Review* 87: 633-72, p. 661.

Science in the United States,[97] suggesting ways in which the quality of forensic work might be improved *at the source*. The NRC notes, as I did earlier,[98] that the *Daubert* trilogy has done disappointingly little to improve the quality of forensic testimony:

> *Daubert* and its progeny have engendered confusion and controversy. [...] Federal appellate courts have not with any consistency or clarity imposed standards ensuring the application of scientifically valid reasoning and reliable methodology in criminal cases involving *Daubert* questions.[99]

Even a summary of the NRC's recommendations–which contain much good sense[100]–would require a paper of its own; the point I want to stress here is that this report looks for ways to strengthen the practice of forensic science before it ever gets to court, not to control the admissibility of forensic testimony or flush out mistakes and dishonesty in such testimony at trial. The report has prompted the establishment of the National Commission on Forensic Science referred to earlier (2013);[101] a report from the Subcommittee on Forensic Science of the National Science and Technology Council Committee on Science on what could be done to implement the NRC recommendations, including estimates of the cost of, e.g., implementing a proficiency-testing program (2014);[102] and, the same year, policy recommendations from the National Institute of Standards and Technology on the presentation of expert testimony.[103] But, as far as I can determine, the substantive changes that the NRC proposed remain largely prospective.

[97] National Research Council, *Strengthening Forensic Science in the United States: A Path Forward* (Washington, D.C.: National Academies Press, 2009).

[98] I mean, both earlier in this paper (p. 59 above) and several years before the NRC report. See Susan Haack, "Trial and Error" (2005), in Haack, *Evidence Matters* (note 13 above), 104-21, pp. 116, 120.

[99] *Id.*, p. 11 (citing Neufeld, "The (Near) Irrelevance of *Daubert* to Criminal Justice" (note 83 above). The NRC's somewhat naïve reference to "scientifically valid reasoning" and "reliable methodology" is worthy of note.

[100] As well as the predictable calls to set up and fund a new federal body, a National Institute of Forensic Science, to establish and enforce better practices at forensic laboratories, and for funds for more research.

[101] Department of Justice, *National Commission on Forensic Science* (2013).

[102] National Science and Technology Council Committee on Science Subcommittee on Forensic Science, *Strengthening the Forensic Sciences* (Washington, DC: Office of the President, 2014).

[103] National Institute of Standards and Technology, Presentation of Expert Testimony: Policy Recommendations (October 29, 2014). Some of the recommendations seem very sensible; others, e.g., that "[e]xperts should remain neutral, and attorneys should respect this neutrality," sound to me like whistling in the dark.

3. Lessons to be Learned: Shifting, and Broadening, our Focus

All these efforts to control the quality of expert testimony—though none could be described as an unqualified success—have brought to light what a tangled mesh of problems, theoretical and practical, courts' handling of such testimony involves, among them:

- the extraordinary variety of fields of expertise—some stronger, some weaker, and some so feeble as scarcely to constitute real fields of expertise at all;
- the existence of more and less competent practitioners in every field, even the strongest;
- the guild mentality that affects some areas, perhaps especially the weaker forensic sciences;
- the conceptual difficulties of distinguishing scientific from other kinds of expertise (or for that matter "hard" from "soft" science);
- and the squishiness of the idea of "methodology";
- the elusiveness of the contrast between the neutral expert and the biased, and the potential for tension between an expert's being "neutral" and his being competent to the task;
- etc., etc.

They should also have taught us a number of other important lessons.

Our experience with the Frye *Rule*[104] *should have taught us* that the very reasonable thought that consensus among scientists in a field is the best indication a lay judge can have that this theory or that technique can be trusted isn't quite as helpful as it initially seems. It should have been obvious long before Justice Breyer's ruling in *Kumho Tire* that the fact that, e.g., a psychiatric theory or a forensic technique is "generally accepted" in its field is little or no assurance of its reliability when the field itself is weak, small, cliquish, and/or self-serving.

Some purported fields of "specialized knowledge, experience, skill, or training"—mind-reading, say, astrological prediction, or phrenological

[104] Frye is still the law in a number of states. According to Demosthenes Lorandos and Terence Campbell, "Mental Health Experts: Science and the Law," *Cross Examining Experts in the Behavioral Sciences* (St. Paul, MN: Thomson Reuters, 2014), §1:16.1, notes 4 and 5, 36 states are now "*Daubert* or *Daubert*-leaning," and 12 continue to use *Frye* (in the text, however, the authors seem to have miscounted!).

diagnoses of criminal personality—are really no such thing.[105] But mostly it's a matter of degree, of more reliable forms of expertise, and less. DNA identifications, properly conducted, are much more reliable than hair analysis or bite-mark identifications; physical analysis of the canvas, paint, varnish, etc., of a painting probably more reliable than historians' intuitive judgment of its likely date and provenance; a chemical analysis of a drug likely more reliable than a sociological analysis of the roots of crime. Consensus in a "field" that is misconceived or fraudulent is *no indication* of reliability; more generally, consensus in a field is a *less robust indication* of reliability, the weaker the field in question.

And unfortunately—as we in Florida know from the long-running saga of Joseph Ramirez, convicted three times of a stabbing murder on the basis of a knife-mark examiner's testimony that he could identify this specific knife, to the exclusion of all other knives in the world, as the one that made the half-inch wound in the victim's neck[106]—it's all too easy for a tightly-knit guild of specialists in a relatively weak field, because *they* all agree it works, to convey a quite unjustified sense that their technique is sound enough for a jury to hear.

Our experience with the Daubert *trilogy should have taught us* that the preoccupation with the demarcation of science and the question of "methodology" was, at best, a distraction; that to ask judges to assess the reliability of any and every kind of specialized knowledge, technique, or skill imposes a burden they are ill-equipped to carry; and that, because of the huge range of types of kinds of expertise, guidelines for determining whether proffered expert testimony is reliable inevitably to end up leaving judges with plausible-sounding verbal formulae the effective application of which requires them—well, as I said earlier, to "use their judgment, and do the right thing."

[105] To give a more realistic example: in the 1980s testimony about supposedly "recovered" memories played a significant role in numerous cases of alleged sexual abuse of small children. Some events are so traumatic, the theory was, that all memory of them they will be blocked from consciousness; but the memories are still there, unsuspected, and may return spontaneously years later, or be brought to consciousness in therapy or under hypnosis. Whether or not this theory is true, it's not clear that, absent independent evidence of abuse, genuine instances of recovered memories can be reliably distinguished from false "memories" planted, consciously or otherwise, by therapists, or simply cooked up in the imaginations of the suggestible or psychologically disturbed. See generally Daniel Brown, Alan W. Scheflin, and D. Corydon Hammond, *Memory, Trauma Treatment, and the Law* (New York: W. W. Norton, 1998); and, on the handling of recovered memory testimony since *Daubert*, Robert Timothy Reagan, "Scientific Consensus on Memory Repression and Recovery," *Rutgers Law Review* 51, no.2 (Winter 1999): 275-321.

[106] The story is summarized in Ramirez v. State, 810 So. 2d 836 (Fla. 2001).

Justice Blackmun's confusion of "reliable" with "scientific" diverted courts' attention from what should have been an obvious fact: that *not all* scientific experts are reliable—some are honestly mistaken, some incompetent, some self-deceived, and probably a few outright dishonest; and *not only* scientific experts are reliable, either—I'm sure there are reliable experts in forensic accounting, the valuation of real estate, computer hacking, "lifecare" costs,[107] etc., etc., too. Moreover, by equating "reliable" and "scientific," *Daubert* focused courts' attention on whether proffered expert testimony is, or isn't, science. Some cases are clear: the expert testimony of an epidemiologist or a toxicologist is scientific evidence; the expert testimony of an art historian or a theologian is not. But where does a physician's differential diagnosis fall,[108] or a psychiatrist's testimony as to a defendant's fitness to stand trial, or a construction engineer's testimony that the building collapsed because the joists used weren't strong enough, or ...., etc.? Are social-scientific experts to be held to the same standards as natural-scientific experts, or should the bar be set lower for the "soft" sciences than the hard?[109] But their preoccupation with such questions didn't do much to help courts ensure the quality of expert testimony.

And by suggesting that what makes expert testimony scientific is that it uses the "scientific method" to arrive at its conclusions,[110] *Daubert* generated a fruitless, and sometimes laughable, preoccupation with "methodology." Looking at all the evidence and using your judgment about the degree to which it warrants a causal conclusion was elevated to the status of "Weight of Evidence Methodology";[111] fingerprint examiners' procedure of analyzing and comparing prints and then asking a second examiner whether he agrees there's a match became the "ACE-V [analysis-comparison-evaluation-verification]

[107] Experts who calculate the medical and related costs that a personal-injury victim will incur over his or her lifetime.

[108] See, e.g., *Moore v. Ashland Chemical* (note 60 above), 280 (Justice Dennis, dissenting).

[109] See, e.g., Nenno v. State, 970 S.W.2d 482 (Tex. Crim. App.1998).

[110] I have argued elsewhere that there *is no* "scientific method," at least if what that means is a method used by all scientists and only scientists, and responsible for the success of scientific inquiry. See Susan Haack, *Defending Science—Within Reason: Between Scientism and Cynicism* (Amherst, NY: Prometheus Books, 2003), especially chapter 4; "Six Signs of Scientism" (first published, in Chinese and Spanish, in 2010) in Haack, *Putting Philosophy to Work: Inquiry and Its Place in Culture* (note 48 above), 105-120 (text) and 278-83 (notes).

[111] See Susan Haack, "Proving Causation: The Weight of Combined Evidence" (2008), in Haack, *Evidence Matters* (note 13 above), 208-38.

methodology";[112] even Dennis Carlson, the tire-design expert in *Kumho Tire*, claimed to have a methodology: "visual inspection methodology"[113]—meaning, apparently, that he looked at the blown-out tire!

And now, since the Supreme Court abandoned the distinction between methodology and conclusions in *Joiner* and acknowledged, in *Kumho Tire*, that what's important isn't, after all, whether expert testimony is scientific, but whether it's reliable, the impossibility of identifying operationally effective indicia of reliability applicable to any and all of the host of potentially legally-relevant fields of expertise is unmistakable.

Granted, if judges were better-educated scientifically, they would be better able to use their judgment and do the right thing. But *our experience with scientific education for judges should have taught us* that—while it's certainly desirable that they have some understanding of, e.g., the ways in which epidemiological studies may be well, or poorly, designed and conducted, the basics of probability theory, how the scientific peer-review system really works, why a DNA identification is almost certainly more reliable than a fingerprint match to a latent print that amounts to 20% of one finger, ..., etc.—there's simply no way to bring (let alone keep) judges up to speed on every kind of expert testimony with which they may be faced. There are just too many potentially legally-relevant fields of expertise. It's impossible to solve the problem Judge Hand drew to our attention more than at century ago, that we set laymen to decide where experts disagree, by making judges experts on *everything*.

Our experience with court-appointed experts should have taught us that, while such experts are by definition "non-partisan" in the superficial sense that they weren't hired by one or other of the parties to a case, it's extraordinarily difficult to ensure even that these experts have no conflict of interest, let alone that they are neutral in an epistemologically robust sense. In fact, it's simply not realistic to expect to find someone competent to the task with *no* opinion, *no* tendency to conclude one way or the other, at the outset. In short, while it's common to hear "biased" experts contrasted with "neutral" ones, and sometimes assumed that this is co-extensive with the contrast between experts hired by a party and experts appointed by a judge, we should know by now that this is a big muddle.

[112] *Mitchell* (note 84 above), 221. It's also worth noting that the "verification" stage of this procedure is sometimes elevated to the status of "peer-review." *Id.*, 238.
[113] *Kumho Tire* (note 65 above), 146.

When the parties to a case choose expert witnesses, naturally they seek out specialists in the field who will testify in a way that favors their side; and naturally they prepare their expert witnesses to offer the strongest testimony they can. Moreover, expert witnesses often seem to become increasingly dogmatic as they testify over and over. But the fact that an expert is chosen by a judge, rather than by one of the parties to a case, doesn't guarantee "neutrality" even in the sociological sense of "having no professional contact, direct or indirect, with either party," let alone in the epistemological sense of "having no preconceived opinion."

In fact, it's not clear that this would be desirable even if it were feasible. In specialized medical-scientific fields such as those at issue in the silicone breast-implant cases and the like, anyone competent to offer an opinion will almost certainly have *some* professional interaction, direct or indirect, with others who have some professional interaction, direct or indirect, with one of the parties. Moreover, as I said, anyone competent to offer an opinion on some specialized matter will surely have *some* ideas on contested issues ahead of time. A supposed "expert" unaware that a drug with an atomic weight of less than 1,000 taken by a pregnant woman can cross the placental barrier, and so *might* harm the fetus,[114] for example, would surely not be competent to opine on whether a morning-sickness drug is teratogenic.[115] And when, as is almost always the case in litigation of this kind, the science at issue is thus far unsettled, even the most competent, honest, and conscientious experts may, quite reasonably, disagree.

Our experience with the ongoing saga of Melendez-Diaz *and its aftermath should have taught us* that, if several forensic technicians had to testify in court in every criminal case that goes to trial,[116] our already-overburdened forensic services might soon be even less adequate to their task than they are now; but also have prompted the thought that, while cross-examination *may* reveal crucial weaknesses in forensic testimony, there's not only absolutely no guarantee that it will, but also, when a case is resolved by plea-bargain, a guarantee that it won't.

* * *

[114] Rock Brynner and Trent Stephens, *Dark Remedy: The Impact of Thalidomide and Its Revival as a Vital Medicine* (Cambridge, MA: Perseus Publishing, 2001), p. 12.

[115] That is, causes birth defects.

[116] An appendix to Justice Breyer's concurrence in *Williams* indicates that as many as 13 different analysts may be involved in producing a single DNA profile. *Williams* (note 91 above), 2253-55.

Am I saying, then, that we should just throw up our hands and admit that the problems with expert witnesses are insoluble? No. I *am* saying, though, that improving the quality of exert testimony will require a recognition that we face, not *one* problem, but a whole tangle of interrelated problems. It was never realistic to hope that any legal form of words, however carefully crafted, could by itself enable judges or jurors to discriminate reliable expert testimony from unreliable; nor that court-appointed experts would prove a simple solution to all the problems; nor that cross-examination would always flush out the weak or dishonest expert; nor, ..., etc.

But, once we acknowledge the tangled complexity of the issues, we might see how to make some headway going forward. Rather than focusing on how to tweak the rules of evidence or to boost the role of cross-examination, I suggest, we should think about what could be done (i) to reduce the incidence of bad stuff reaching the courts in the first place, and (ii) to increase the likelihood that it will get exposed quickly if it does. This would mean (as the sub-title of this section says) *shifting* our focus to earlier in the process, and *broadening* it to include more just legal rules and procedures.

Just as I reached this point of the paper, an article in the *Wall Street Journal* described the fallout from the discovery that one technician in the Boston drug-testing lab, the now nationally-notorious Annie Dookhan, had been faking her results: more than 40,000 convictions tainted, and now a big legal brouhaha, with the American Council for Civil Liberties (ACLU) asking that all the potentially tainted cases be reopened and the District Attorney arguing that each such case should be dealt with individually, plus a dispute over what to do about defendants who accepted plea-bargains.[117] Ms. Dookhan's malfeasance was discovered in 2011, when her supervisor caught her taking ninety samples from the evidence vault without signing them out; but it had begun even before she joined the lab nine years earlier: she had lied about her qualifications on her job application.[118] How, I wonder, might we have ensured that such gross dishonesty would be discovered sooner?

[117] Jennifer Levitz, "Lab Flaws Cast Doubt on Drug Convictions," *Wall Street Journal*, January 8, 2015, A3.

[118] Sean K. Driscoll, "'I Messed Up Bad': Lessons on the Confrontation Clause from the Annie Dookhan Scandal," *Arizona Law Review* 56, no. 3 (2014): 717-40. Brian Ballou and Andrea Estes, "Chemist Admitted Wrongdoing in Lab Scandal," *Boston Globe* online (September 26, 2012), available at http://www.bostonglobe.com/2012/09/26/chemist-annie-dookhan-lab-scandal-investigators-messed-bad/uORxid5JamieMK1wumq2uO/story.html.

You might think that Ms. Dookhan's misconduct would surely have been exposed if she had been cross-examined. But no: in the three years before her arrest she *was* cross-examined—in around 150 trials; and no defense attorney ever uncovered even one of her faked and skimped tests.[119] Moreover, the ongoing legal dispute over the disposition of all those plea-bargained cases in Boston,[120] and the similar recent scandals at forensics labs in New, York, Delaware, and Colorado,[121] all bring home the point that, even if cross-examination were better at uncovering forensic malfeasance than, apparently, it is, it would have been far better had the management of such labs been more attentive, their culture healthier, and their hiring practices more vigilant, etc., in the first place.

Well, yes, you may say; but surely this isn't a lesson that can be extrapolated beyond the forensic sciences. Not in any simple way, I agree; but the underlying thought—that it's better, so far as it's possible, to prevent a problem than to fix things later—applies here, too. Think of the storm of litigation over those silicone breast-implants, which had been "grandfathered in" when the FDA's (Food and Drug Administration's) remit was extended to medical devices, but then were banned when the manufacturers failed to submit the evidence of safety that they had been asked to provide by the date the FDA had specified.[122] There was no evidence that the implants were unsafe; but the announcement of the ban seems to have been handled, and reported, in such a way that many of the many women who had such implants panicked, and began to attribute every twinge to them.

You have to wonder: if the ban had been handled differently, mightn't the panic, and the legal fiasco, have been, at least, mitigated? For that matter, wouldn't it have been better if no medical journal had been willing to publish the *only* study that ever found even a small increased risk in women with the implants, given that it relied on the women's own reports?[123] Again: the year after Judge Pointer's and Judge Jones's panels reported that there was no

[119] Levitz, "Lab Flaws Cast Doubt" (note 117 above).

[120] These are many, because many of the tainted cases concern low-level drug offenses, which rarely go to trial. *Id.*

[121] *Id.*

[122] For the background story, see Marcia Angell, *Science on Trial: The Clash of Medical Evidence and the Law in the Breast Implant Case* (New York: W. W. Norton, 1996), chapters 1 and 3; for more details of the panels, see Hooper, Cecil, and Willging, "Assessing Causation in Breast Implant Litigation" (note 77 above).

[123] Charles H. Henneckens, "Self-Reported Breast Implants and Connective-Tissue Diseases in Female Health Professionals: A Retroactive Cohort Study," *Journal of the American Medical Association* 275, no.8 (1996): 616-21. Sanders and Kaye, "Expert Evidence on Silicone Implants" (note 78 above), p 127, n.87.

evidence that the implants caused the connective-tissue diseases they were feared to, a significantly larger panel set up by the Institute of Medicine (IOM) at the National Institutes of Health (NIH) reached the same conclusion. You have to wonder: wouldn't it have been better if the IOM had stepped in sooner?

My theme throughout has been that ensuring the quality of expert testimony involves a whole tangled knot of tricky problems; so it should come as no surprise that, rather than offering a panacea, I conclude by urging that we think harder and more imaginatively not only about legal rules and procedures but also about the many *other* ways in which we might, bit by bit, make expert testimony more often genuinely helpful to fact-finders, less often confusing or misleading.

ACKNOWLEDGEMENTS

My thanks to Mark Migotti for his helpful comments on draft versions, and to Pamela Lucken and Bianca Anderson in the University of Miami Law Library for skilled help in finding relevant materials.