



2023 FORUM FOR STATE APPELLATE COURT JUDGES
EXPERT TESTIMONY: JUDGES, SCIENCE, AND TRIAL BY JURY

EXPERT EVIDENCE: EVOLUTION OF RULES AND PRACTICES

By Michael J. Saks

I. Introduction: Expert Evidence is Different

The most fundamental rule of evidence, of course, is that relevant evidence is admissible while evidence that is not relevant is not admissible. Ostensibly relevant expert evidence, however, is required to clear a higher hurdle in order to be admitted. Why?

The standard answer is that, while ordinary witnesses normally may testify only to facts they have observed, experts are permitted to testify to opinions—that is, inferences drawn from the observations they (or others) have made.

Those experts, it has long been feared, are too easily selected and influenced to present the proffering party’s view of the matter. While ordinary fact witnesses are limited in number (to those having personal knowledge of the transaction at issue), expert witnesses can be plentiful. If one expert does not see things as counsel wishes them to be seen, another expert might be found who does. After selection, counsel has ample opportunity to subtly influence the expert’s views of the case. As far back as 1843, a court observed of experts: “They come with such a bias on their minds to support the cause in which they are embarked, that hardly any weight should be given to their evidence.”¹

¹ Lord Chancellor Campbell, in the Tracy Peerage, 10 Clark & F. 154 (1839, 1843). See, other illustrations, old and recent:

“[W]e think [expert testimony] should not be much encouraged and should be received only in case of necessity. [Their views] cannot fail to be warped by a desire to promote the cause in which they are enlisted.” Ferguson v. Hubbel, 97 N.Y. 507 (1884).

“Many judges have expressed their thorough dissatisfaction with the prevalent method . . . of making use of the services of experts in the conduct of judicial inquiries. . . . [N]o judge has, in recent times, said aught in praise of the system.... Law writers are equally condemnatory of the system and severe in their reflections upon the product

Sometimes experts go so far as to offer testimony that points in a direction that they know perfectly well is opposite to where the principles of their expertise actually point.² Still other times, expert witnesses offer sincere nonsense. In this circumstance, the entire field of experts believes things that upon later testing turn out to be incorrect. “In the Witches’ case, in 1665, Dr. Brown, of Norwich, was desired to state his opinion of the accused persons, and he was clearly of opinion that they were witches, and he elaborated his opinion by a scientific explanation of the fits to which they were subject.” We need not, however, look back so far, to find examples of sincere nonsense. Among modern examples is arson investigation, which for decades relied on nearly two dozen “arson indicators” until they were belatedly determined by empirical testing to be incapable of distinguishing set fires from accidental ones.³

On the receiving end, lay factfinders, it has long been feared, are prone to over-weighting the testimony of experts.⁴

Before being allowed to testify, therefore, proffered experts must be screened to ensure that they possess sufficient “knowledge, skill, experience, training, or education” (“qualifications”) and their proffered testimony must satisfy Rule 702. In earlier times, somewhat different criteria were employed.

of that system—the expert. Experts themselves do not like it.” C. Herschel, *Services of Experts in the Conduct of Judicial Inquiries*, XXI AM. L. REV. 571 (1887), at 571.

“[T]here is a constant complaining and mistrust on the part of judges, juries and lawyers of the expert witness.” L.M. Friedman, *Expert Testimony, Its Abuse and Reformation*, 19 YALE L. J. 247 (1910), at 247.

An experiment in which 108 experts were paid to review the same case files, having been randomly assigned to work for one side of a case or the other, found the experts’ views automatically, without prompting, tilted toward the interests of the hiring party. Daniel Murrie et al., *Are Forensic Experts Biased by the Side That Retained Them?*, 24 PSYCH. SCI. 1889 (2013).

A study by the FBI found greater than 98% of reports and testimony by the government’s microscopic hair comparison experts exaggerated their conclusions beyond what the discipline was capable of finding. See, ABS GROUP, *ROOT AND CULTURAL CAUSE ANALYSIS OF REPORT AND TESTIMONY ERRORS* (August 2018), at <https://vault.fbi.gov/root-cause-analysis-of-microscopic-hair-comparison-analysis/root-cause-analysis-of-microscopic-hair-comparison-analysis-part-01-of-01/view>.

² E.g., *Gregory v. City of Louisville et al.*, 444 F.3d 725 (2006). False and misleading testimony by forensic scientists has, overall, been one of the major causes of wrongful conviction. See, National Registry of Exonerations, at <https://www.law.umich.edu/special/exoneration/Pages/about.aspx>.

³ DAVID L. FAIGMAN ET AL., *MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY* (West, updated Annually), Fire and Arson chapter.

⁴ Some commentators question the assumptions on which differential treatment of expert evidence is based, and offer data to support their doubts. Frederick Schauer and Barbara A. Spellman, *Is Expert Evidence Really Different?*, 89 NOTRE DAME L. REV. 1 (2013).

The need for such screening (gatekeeping) presumably increases in proportion to the extent of the public's mindless faith in the asserted expertise, the difficulty of evaluating the expert's claims, and the risk of exaggeration or fraud. Where a type of asserted expertise is systematically more available to one side than the other, the situation is worse while appearing to be better: the absence of disagreement makes it appear there is nothing to disagree about.

All of this is, of course, a dilemma. Judges have been aware of both the benefits and the risks of expert witnesses for as long as there have been adversarial legal proceedings. The worries outlined above are offset by the fact that expert witnesses have the potential to "help the trier of fact to understand the evidence or to determine a fact in issue."⁵ "No one will deny that the law should in some way effectively use expert knowledge wherever it will aid in settling disputes."⁶

A central problem is that the specialized knowledge base—which is exactly what the court wishes the factfinder to access—is unknown to laypersons. How, then, can the jury evaluate how much of what they are hearing is accurate? How can they determine which of two conflicting experts is providing the more complete and candid account? And how can the judge, whose job is to help the jury by separating the sound from the unsound, know which is which? Falling back on "credibility" is more of an illusion than a solution. This is the "expert dilemma," most famously articulated by Learned Hand.⁷

The law has a number of procedural tools to help reduce the risk of misleading expert testimony. These include the oath, cross-examination, the possibility of counter-experts proffered by opposing counsel. Further, Rule 706 empowers judges to appoint the court's own experts whenever they choose to do so. The drafters of Rule 706 shared the widespread skepticism concerning whether experts who were chosen by the parties, briefed by them, prepared by them, and paid by them, would provide the education that the judge and jury need. So they enabled judges to appoint their own expert witnesses. But the effectiveness of this appointment power was thought to reside not so much in its use but in its mere existence. "[T]he assumption may be made that the availability of the procedure in itself decreases the need for resorting to it. The ever-present possibility that the judge may appoint an expert in a given case must inevitably exert a sobering effect on the expert witness of a party and upon the person utilizing his services."⁸

⁵ Federal Rules of Evidence, Rule 702.

⁶ Learned Hand, *Historical and Practical Considerations Regarding Expert Testimony*, 15 HARV. L. REV. 40 (1901).

⁷ *Id.*, at 48.

[H]ow 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.

Moreover, there can be no competent tribunal, except one composed of those who have possessed themselves of the specialized experience and the trained powers of observation necessary to bring to a valid test the truth of the various propositions offered.

⁸ Rule 706, Advisory Committee Comments.

II. How Did Courts Screen Expert Evidence in the Centuries Before 1923?

Long before *Frye v. United States* (1923) was decided, physicians were invited to courts and asked about wounds and pregnancies and the cause of death; merchants were asked about the condition of food and drink; engineers about wells and bridges, and so on. Experts were called as early as 1345. In that instance, a court summoned surgeons to aid in determining whether a wound was fresh, though this was for the purpose of deciding a motion.⁹ Perhaps the earliest true expert witness—called to present testimony to a jury—was in *Alsop v. Bowtrell* (1620).¹⁰

How did courts of old screen those proffered experts? One test found in 19th Century America, was an implicit one. That is, courts did not announce a formal rule, but their approach can be discerned from discussion in their opinions. Unsurprisingly, courts assessed whether the proffered expert was prepared to offer information that was beyond the ken of the court and the jury, and whether the proposed witness was “qualified.” Although courts spoke of the expert’s “greater study respecting certain subjects,” they seem also to have been attentive to whether the witness had achieved some degree of success in the practice of the occupation or profession claiming that knowledge. The implicit reasoning seems to have been: if a person could prosper selling the knowledge or skill at issue, then the person was expert enough to testify. In effect, the marketplace determined whether valid knowledge existed by endowing it with commercial value. This has been termed the “marketplace test” of admissibility.¹¹

The marketplace test did not directly assess the validity of the claimed knowledge. It allowed consumers (of the services of doctors, food merchants, well-diggers, etc.) to perform the critical evaluation. This was a clever shortcut for judges. Except that consumers have been known to be fooled by sellers claiming to have expertise they do not have or asserted skills and products that were of no actual value (e.g., snake oil and other quack cures, dowsing, astrology, and their modern descendants). Also, the marketplace test is unable to evaluate areas of expertise for which there is no consumer market, or which are too new to have gained a clientele.

III. *Frye v. United States* (1923): The General Acceptance Test

Frye v. United States (1923)¹² can be better understood against the background of the marketplace test that preceded it. In *Frye*, the defense proffered an expert using an early form of lie detector. At that time, no market existed for the services of polygraph examiners. At trial, the proffered expert testimony was rejected. On appeal, Judge Van Orsdel solved the problem that no marketplace yet existed to which courts could turn for an assessment of the claimed expertise

⁹ Anonymous, Lib. Ass. 28, pi. 5 (28 Ed. III).

¹⁰ Cro. Jac. 541.

¹¹ David L. Faigman et al., *Check Your Crystal Ball at the Courthouse Door, Please: Exploring the Past, Understanding the Present, and Worrying about the Future of Scientific Evidence*, 15 Cardozo L. Rev. 1799 (1994).

¹² 293 F.1013 (D.C. Cir 1923).

by substituting a scientific or intellectual market for the commercial one, thereby creating the *Frye* test of general acceptance—and affirmed the district court’s exclusion. A novel scientific principle “must be sufficiently established to have gained general acceptance in the particular field in which it belongs.” In the particular instance, “the systolic blood pressure deception test has not yet gained such standing and scientific recognition among physiological and psychological authorities as would justify the courts in admitting expert testimony deduced from the discovery....”

A. Shortcomings of the *Frye* Test

Significantly, however, the *Frye* test replaced consumers with producers as the crucial assessors of validity. Control over the assessment of validity was transferred to the people who produced the asserted knowledge and offered it (and themselves) to the courts. Consumers might not be perfect in their assessments, but at least their interests aligned better with those of the courts.

Another problem with *Frye* is the malleability of “the particular field in which [proffered expert evidence] belongs.” The “particular field” has sometimes been interpreted to mean the narrow circle of purported experts who conduct the technique at issue. Other times it has been understood more broadly to include relevant fields that have meaningful knowledge about the concepts or techniques. Let’s call those the narrow versus the broad versions of *Frye*.

For example, when voice spectrography was proffered, purporting to identify the individual whose voice was recorded (e.g., making a threatening phone call), some judges limited the experts on the expertise (who would say whether it was generally accepted or not) to practicing voiceprint examiners.¹³ Other judges were interested in hearing from additional knowledgeable experts—acoustical engineers, linguists, statisticians. In every instance when the *narrow* version of *Frye* was employed, the court found voiceprints to be admissible. In every instance when the *broad* version of the *Frye* test was employed, the court excluded the voiceprint evidence.¹⁴ Thus, if a judge applying *Frye* wanted to maximize the probability that the evidentiary hearing would support admission, the judge could choose the narrow form of the test. To increase the probability that flaws in the proffered technique’s theory or data or practice would emerge, the broad version could be employed. By choosing one flavor of *Frye* over the other, the judge has gone a long way towards determining the outcome.

For a long time, *Frye* was taken to apply to criminal cases but not civil. And, because *Frye* referred to “novel scientific evidence,” many later courts thought the test applicable only when “novel”

¹³ Suppose astrological predictions were relevant to a case. Under *Frye*, astrologers would have to be relied on to tell us whether astrology or one of its techniques is generally accepted by astrologers.

¹⁴ MODERN SCIENTIFIC EVIDENCE, supra note 3, Chapter 1.

science was being proffered.¹⁵ Numerous other difficulties with *Frye's* logic and its application have been discussed in the legal literature.¹⁶

B. *Frye* Not So Dominant as Commonly Believed

The U.S. Supreme Court offered a simple history of expert evidence admissibility rules. First there was *Frye*; then came *Daubert*. “In the 70 years since its formulation in the *Frye* case, the ‘general acceptance’ test has been the dominant standard for determining the admissibility of novel scientific evidence at trial,” said *Daubert*.¹⁷

Historically, however, *Frye* was not so top-of-mind as that picture suggests. To begin, the *Frye* variant went unnoticed for quite some time. It was not cited by any other court, federal or state, for a decade. Judge Van Orsdel himself did not so much as mention *Frye* or any requirement of general acceptance when he reviewed the novel technique of firearms identification—also untested, “experimental,” not yet generally accepted, and which had recently been found inadmissible by a prominent state supreme court¹⁸—and held the technique to be admissible.¹⁹ During the first quarter century of its existence, *Frye* was cited in only eight federal cases and five state cases. That amounts to one case every other year in the entire country. During its second quarter century, citations increased to 54 times in federal cases and 29 times in state cases.

The *Frye* test was not truly “discovered” until the drafting of the Federal Rules of Evidence was under way and those discussions awakened interest in a rule for judicial gatekeeping of expert evidence. Then, ironically, *after* the adoption of the Federal Rules of Evidence (which we now know, and should always have known, did not incorporate *Frye*), *Frye* started to be employed as much each year as it had been in all of its first fifty years added together. *Frye's* sun rose just as it had been ordered to set.

IV. Other Tests and No Test

When they did not employ the general acceptance test, what were courts using to guide their evaluation of expert evidence between *Frye* (1923) and *Daubert* (1993)? Some used a “relevancy” test, best articulated by Professor McCormick, which treats the validity of the underlying principle and the validity of the technique as matters of relevancy. In addition, of course, the information must be testified to by an expert qualified to present that subject matter, and exclusion could be based on other considerations:

¹⁵ This implies a belief that once a scientific discipline comes to believe in the validity of something, that something becomes true for all time—as if science never corrects its errors.

¹⁶ See., e.g., GIANNELLI. MODERN SCIENTIFIC EVIDENCE, *supra* note 3; Paul C. Giannelli, *The Admissibility of Novel Scientific Evidence: Frye v. United States, A Half-Century Later*, 80 COLUM. L. REV. 1197 (1980).

¹⁷ *Daubert v. Merrell Dow Pharmaceuticals*, 509 U.S. 579 (1993).

¹⁸ *People v. Berkman*, 307 Ill. 492, 139 N.E. 91 (1923).

¹⁹ *Laney v. United States*, 294 F. 412 (App. D.C. 1923). Signed the same day as *Frye*.

“General scientific acceptance” is a proper condition upon the court’s taking judicial notice of scientific facts, but not a criterion for the admissibility of scientific evidence. Any relevant conclusions which are supported by a qualified expert witness should be received unless there are other reasons for exclusion. Particularly, its probative value may be overborne by the familiar dangers of prejudicing or misleading the jury, unfair surprise and undue consumption of time.²⁰

Empirical demonstration was another way to try to establish the reliability of a new technique. On this basis, a number of forensic sciences were admitted, but the demonstrations were often surprisingly flimsy.²¹ Questions that were not squarely faced were how sound the methodological quality of the empirical demonstration (or testing) has to be and how much testing is required to establish validity.²²

Some courts employed a balancing test, weighing the probative value of the proffered scientific evidence against other factors, among them the significance of the issue to which the evidence is directed, the availability of other evidence, and the utility of limiting instructions.²³

Florida, Iowa, New York, and Utah required only that the expert vouch for the theory and technique. In Ohio, New Mexico and at least one federal jurisdiction, courts held that the defendant had a constitutional right to present critical scientific evidence regardless of whether it satisfied *Frye* or not.

Some courts, claiming to apply *Frye*’s general acceptance test, were plainly doing something else. See, for example, *United States v. Stifel* (1970) which used the McCormick “relevancy” test while purporting to be relying on the *Frye* test.²⁴

No Test at All

Another popular option was no test at all. Numerous kinds of purportedly scientific evidence were admitted without anything that looked like formal scrutiny.

As mentioned earlier, the author of *Frye* did not employ it himself in another case decided the same day as *Frye*. Instead, his analysis came only to this: “the testimony given by the expert witnesses, tending to establish that the bullet, extracted from the head of the deceased, was shot

²⁰ McCormick on Evidence, §§ 202, 203 (2d ed. 1972).

²¹ See, Michael J. Saks, *Merlin and Solomon: Lessons from the Law’s Formative Encounters with Forensic Identification Science*, 49 HASTINGS L.J. 1069 (1998).

²² The parallels to *Daubert* should be obvious.

²³ WEINSTEIN & M. BERGER, WEINSTEIN’S EVIDENCE 1702[03] (1980).

²⁴ 433 F.2d 431 (6th Cir. 1970), cert. denied, 401 U.S. 994 (1971). Just as we shall see examples of courts claiming to be applying *Daubert* but plainly doing something else.

from the pistol found in the defendant's possession, was competent, and the examination in this particular was conducted without prejudicial error...."²⁵

In another example—actually, five examples—in 1999 the Kentucky Supreme Court, claiming to be employing Kentucky's adoption of *Daubert*, upheld the admission of microscopic hair comparison evidence on the basis of the technique's having been held to be "generally accepted" in five prior Kentucky cases.²⁶ Yes, five prior decisions did hold microscopic hair comparison expert testimony to be admissible. But, the court noted, there was an "absence in our previous opinions of any in-depth analysis under the 'general acceptance' test." That's too generous. Not one of the cited cases engaged in any analysis of admissibility of any kind. However, said the court, "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 trouble with assuming its prior decisions applied *Frye* though they said not a word about it, overlooked the fact that Kentucky was one of those states that hadn't discovered *Frye* until far into the 20th Century. The first citation to *Frye*'s general acceptance test by a Kentucky court occurred in 1983. The cases that supposedly relied *Frye sub silentio* were decided in the 1950s and 1970s.

The problem is not only that a court in 1999 had to conjure holdings out of the silence of its prior opinions, or that it mistakenly assumed general acceptance was the law in its state when it had not been. The larger problem is that, in case after case, an unvalidated form of expert testimony was admitted on the basis of nothing.

Other forms of expert evidence were grandfathered in. They had been coming to courts before 1923, and *Frye* was taken to apply only to "novel" science. So there was no occasion to reconsider their admissibility under the new standard.

V. *Daubert v. Merrell Dow Pharmaceuticals* (1993): The Scientific Validity Test

In 1993, *Daubert* held that the Federal Rules of Evidence, adopted in 1975, did not incorporate the *Frye* test or any requirement of "general acceptance."²⁷ Given that neither the Rules, the commentary, drafts of rules and commentary, nor debates leading to the Rules make any mention of that case or its central concept, the Supreme Court's unanimous holding ought to have been unsurprising.²⁸

²⁵ Laney v. United States, supra note 19.

²⁶ Johnson v. Commonwealth, 12 S.W.3d 258 (Ky. 1999).

²⁷ *Daubert*, which concerned whether the drug Bendectin had teratogenic effects.

²⁸ The Second Circuit and the Maine Supreme Judicial Court promptly construed the new Federal Rules of Evidence to impliedly overrule *Frye*.

Instead, the Court explained, the “overarching subject” of “[t]he inquiry envisioned by Rule 702” “is the scientific validity and thus the evidentiary relevance and reliability—of the principles that underlie a proposed submission.”²⁹ Validity is the central command of *Daubert*.

A. Comparing *Frye* and *Daubert*

Initially, many judges and lawyers believed that *Daubert* was a less demanding filter. The opinion itself spoke of “the Rules’ liberal thrust and their general approach of relaxing the traditional barriers to ‘opinion’ testimony.” But sometimes what an opinion does is more significant than what it says.

The essential distinction between *Frye* and *Daubert*, and therefore the impact of *Daubert* on gatekeeping, is summarized in Figure 1. Any given theory or principle or technique can be based on a strong scientific foundation or a weak one (the *Daubert* inquiry). Or, independently of that, it can enjoy high or low “general acceptance” in the particular field or fields in which it belongs (the *Frye* inquiry). As Figure 1 depicts, proffered knowledge that is based on valid science *and* enjoys general acceptance should be admitted by either test. Conversely, proffered knowledge that is not based on valid science *and* does not enjoy general acceptance should be excluded by either test.

But there are two circumstances in which the two attributes can be discordant for any given type of asserted scientific evidence. Where proffered knowledge is based on a sound scientific foundation but has not (or not yet) gained general acceptance, *Frye* would exclude while *Daubert* would admit. This is the situation that is usually envisioned when the two tests are discussed, leading *Daubert* to be seen as the more liberal test. But where proffered knowledge has only a weak scientific foundation and yet enjoys general acceptance in its field, *Frye* would admit but *Daubert* would exclude. In this situation, the *Frye* test is the more porous filter.

Figure 1. Comparison of *Frye* and *Daubert*

	Daubert: Valid Foundation	
Frye: General Acceptance	Strong	Weak
High	Both admit	Frye admits Daubert excludes
Low	Frye excludes Daubert admits	Both exclude

That fourth situation occurs more frequently than one might expect. Indeed, it is the quadrant into which many forensic sciences fell. Judges were surprised to find that techniques and

²⁹ *Daubert*. See n. 9 (“In a case involving scientific evidence, *evidentiary reliability* will be based upon *scientific validity*” (italics in original).) The “relevance” in the words quoted in the text reflects the Court’s argument that evidence which is not valid cannot be relevant. See also Rule 104(a-c).

knowledge claims that had passed muster under *Frye* had a much harder time getting through the *Daubert* filter. For example, the first time fingerprint identification was challenged under *Daubert*, the government (the proponent of the evidence) was astonished to learn that its experts had *no studies* (no data, no evidence) with which to support their century of claims about uniqueness and the flawless accuracy of fingerprint identification.³⁰ Fingerprint identification was generally accepted among fingerprint examiners, but lacked a foundation of empirical testing—the *sine qua non* of science and of *Daubert*. This is but one example of expertise that passed the *Frye* test easily but found *Daubert* more formidable.

B. The “*Daubert* Factors”

The *Daubert* opinion offered some “general observations” about how judges might evaluate scientific evidence in carrying out their gatekeeping duties under Rule 702. These included looking for empirical testability (and presumably actual testing), peer review (especially to facilitate evaluation of research design and procedures so that “substantive flaws in methodology will be detected”), error rates (which presumably includes, more broadly, the findings of empirical studies relevant to the validity of the proffered expert evidence), “the existence and maintenance of standards controlling the technique’s operation,” and, still, a dash of “general acceptance,” which “can yet have a bearing on the inquiry.”

Though the opinion took care to explain that these are not “a definitive checklist,” most lawyers and judges have treated them as if they were exactly that.

C. Other Aspects and Elements

Besides making validity the touchstone of admissibility, in dicta *Daubert* clarified a number of other things. The distinction between validity-of-the-asserted-expertise and qualifications-of-the-expert was sharpened. Thus, if a technique or body of asserted knowledge is not valid, even the most highly qualified expert on that subject should be barred from giving testimony.³¹ The test of admissibility applies to non-novel as well as to novel proffers. And the proponent of the expert evidence must establish validity by a preponderance of the evidence.

D. *Daubert* on Remand

When *Daubert* was remanded, the circuit panel was clear-eyed enough to see that the *Daubert* factors were not a definitive checklist, that “[t]he inquiry envisioned by Rule 702 is... a flexible one,” and that the key requirement was to effectively assure the validity of admitted expert evidence. So the opinion devised an approach to evaluating admissibility that was thought to accomplish that goal, mixing additional criteria with *Daubert*’s suggestions.³²

³⁰ United States v. Llera-Plaza, 179 F. Supp. 2d 492, 188 F. Supp. 2d 549 (E.D. Pa. 2002). See extensive discussion in MODERN SCIENTIFIC EVIDENCE, supra note 3, Fingerprint Identification chapter.

³¹ Think: astrologers.

³² *Daubert v. Merrell Dow Pharmaceuticals*, 43 F.3d 1311 (9th Cir. 1995)

Notably, the remand opinion required that expert witnesses must come to their knowledge through pre-existing research not conducted in contemplation of litigation. This reflects familiar concerns about the impact of lawyers and the litigation process on the development of experts' opinions. Also, the experts must publish their work, so that their conclusions are available for their field to evaluate. This criterion makes explicit what in *Daubert* was only implicit: that the reactions of peers that arise after publication are more illuminating than the editorial peer review that takes place before publication.

Recognizing that those criteria, used in this case to evaluate the testimony of experts on toxicology and epidemiology, would apply equally to forensic science, many of which would fare poorly, the author of the opinion added a footnote exempting crime laboratory forensic sciences from the test just constructed.³³ Twenty-one years later, after serving with scientists on a presidential committee and learning about the weaknesses of forensic science, that same judge did a striking about-face:

The new study from the President's Council of Advisors on Science and Technology (PCAST) examines the scientific validity of forensic-evidence techniques—DNA, fingerprint, bitemark, firearm, footwear and hair analysis. It concludes that virtually all of these methods are flawed, some irredeemably so.

Americans have long had an abiding faith in science, including forensic science. Popular TV shows like "CSI" and "Forensic Files" stoke this confidence. Yet the PCAST report will likely upend many people's beliefs, as it should. Why trust a justice system that imprisons and even executes people based on junk science?³⁴

VI. Other members of the *Daubert* Quartet

A. *General Electric v. Joiner* (1997)

In *General Electric v. Joiner*³⁵ the central issue was the standard of review on appeal of a district court's decision to admit or exclude expert testimony. A unanimous (8-0) Court held that abuse of discretion is the proper standard. The reasoning was simply, "We have held [in past cases] that abuse of discretion is the proper standard of review of a district court's evidentiary rulings."³⁶

³³ Id., at note 5 ("There are, of course, exceptions. Fingerprint analysis, voice recognition, DNA fingerprinting and a variety of other scientific endeavors closely tied to law enforcement may indeed have the courtroom as a principal theatre of operations. As to such disciplines, the fact that the expert has developed an expertise principally for purposes of litigation will obviously not be a substantial consideration."). Whispers leaking from chambers say that Judge Kozinski's panel colleagues debated with him the wisdom of the footnote.

³⁴ Alex Kozinski, "Rejecting Voodoo Science in the Courtroom," WALL ST. J. (Sept. 19, 2016), at <https://www.wsj.com/articles/rejecting-voodoo-science-in-the-courtroom-1474328199>.

³⁵ *General Electric Company v. Joiner*, 522 U.S. 136 (1997) (another toxic tort case, this involving whether PCBs can promote the kind of cancer suffered by the plaintiff).

³⁶ Id. at 137.

Though that seems straightforward enough, it is worth considering the havoc deferential review can play when the facts at issue are scientific facts. The typical evidentiary ruling is something like whether or not a hearsay exception applies given the particulars of the case at bar. If the trial judge makes a mistake, the error affects only the parties to the case; the ruling has no trans-case implications.

Rulings about science are different. They inherently implicate other cases in which the same scientific question arises. It is one thing to rule that the requirements of the excited utterance exception have not been met in a particular case. If they are met in the circumstances of another case, the two rulings are not inconsistent. But one cannot plausibly find that chemical X is capable of causing cancer in one case but in another case that it is not capable of causing cancer. Or find that the comparison of trace elements of bullets is an unsound (and therefore inadmissible) identification technique in one case but that it is sound science in another case.

Moreover, appellate affirmances made under clear error have no precedential effect.³⁷ Parties are therefore entitled to raise the same issue again and again in different courts. (Voice spectrography might be rejected as invalid in the southern district, but could be found valid in the northern district.) Parties' briefs often cite appellate decisions they like, but competent judges will recognize that they are not bound by rulings made under deferential review.

Imagine that two district courts within the same circuit reach opposite conclusions with respect to the admissibility under *Daubert* of testimony based on some scientific matter. Both cases go up on appeal together. What is the court of appeals to do? Under *Joiner*, unless one of those trial judges was far off base, both would have to be affirmed.³⁸

The chaos described above could be managed most effectively, efficiently (in the long run), and rationally through de novo review. De novo review is the rule for questions of law because rule-making has trans-case impact. Consistency (correctness across cases) in the law is an aspect of the rule of law. In that same sense—because they have trans-case implications—scientific questions are law-like at the same time that they are factual and need to be treated similarly.

³⁷ On deferential review, an appellate court can believe a trial court made the wrong decision, but had not abused its discretion such that the court above is entitled to reverse it. When a court of appeals overrules, we have a much clearer situation.

³⁸ A decade earlier, confronting the same problem in a different context, Justice Rehnquist, the author of *Joiner*, saw the problem more clearly. See *Lockhart v. McCree*, 476 U.S. 162 (1986):

McCree argues that the “factual” findings of the District Court and the Eighth Circuit . . . may be reviewed by this Court only under the “clearly erroneous” standard of Federal Rule of Civil Procedure 52(a) We are far from persuaded, however, that the “clearly erroneous” standard of Rule 52(a) applies to the kind of “legislative” facts at issue here. See generally *Dunagin v. City of Oxford, Mississippi* (C.A.5 1983). The difficulty with applying such a standard to “legislative” facts is evidenced here by the fact that at least one other Court of Appeals, reviewing the same social science studies as introduced by McCree, has reached a conclusion contrary to that of the Eighth Circuit. (citations omitted)

Another issue was addressed in *Joiner*. Justice Blackmun’s majority opinion in *Daubert* had stated that in evaluating studies, the “focus must be solely on principles and methodology, not on the conclusions that they generate.”³⁹ *Joiner* argued that in light of the language quoted just above, it was error for the District Court to reject *Joiner*’s expert’s conclusions based on a number of studies, and the Court of Appeals therefore properly reversed. Chief Justice Rehnquist responded to that argument:

But conclusions and methodology are not entirely distinct from one another. Trained experts commonly extrapolate from existing data. But nothing in either *Daubert* or the Federal Rules of Evidence requires a district court to admit opinion evidence that is connected to existing data only by the ipse dixit of the expert. A court may conclude that there is simply too great an analytical gap between the data and the opinion proffered. That is what the District Court did here, and we hold that it did not abuse its discretion in so doing.

Commentators have discussed and debated which view is correct, Blackmun’s or Rehnquist’s. But the two concepts are so different that there really is no conflict. Counsel confused the issue. The court could easily have swept it aside, by saying: that is a different issue; *Daubert* was talking about an apple and here we have an orange.

Blackmun’s point was that the strengths and weaknesses of a study are determined by the methodological qualities of the study (its design, sampling, measures, etc.), independent of whether a judge (or anyone else) likes the study’s results or not.

Rehnquist’s point was that where a large gap exists between an expert’s opinion and the empirical results on which it purports to stand, the expert’s opinion may be excluded.⁴⁰ (That sensible point would have been even clearer had it not been introduced by needless entanglement: “conclusions and methodology are not entirely distinct from one another.”)

B. *Kumho Tire v. Carmichael* (1999)

*Kumho Tire v. Carmichael*⁴¹ addressed several noteworthy issues. First and foremost was the question of “how *Daubert* applies to the testimony of engineers and other experts who are not scientists.” A unanimous Court held that a federal trial judge’s “gatekeeping” obligation applies not only to “scientific” testimony, but to all expert testimony. In his opinion for the Court, Justice Breyer wrote, “We conclude that *Daubert*’s general holding—setting forth the trial judge’s general ‘gatekeeping’ obligation—applies not only to testimony based on ‘scientific’ knowledge, but also to testimony based on ‘technical’ and ‘other specialized’ knowledge.”

³⁹ *Daubert* at 580.

⁴⁰ If a set of (well-designed, well-conducted, well-analyzed) studies suggests with sufficient clarity that X does not cause Y, the expert may not testify that the studies support an opinion that X causes Y.

⁴¹ *Kumho Tire v. Carmichael*, 526 U.S. 137 (1999).

Any and all proposed expert testimony must be found to rest on valid foundations as a condition of admission. “The objective of [the gatekeeping] requirement is to ensure the reliability and relevancy of expert testimony.” Whatever the proffered expertise is, the trial court has to fashion an appropriate test for determining whether or not it is valid. The particular factors mentioned in *Daubert* might or might not be useful in evaluating proffers of non-science expert testimony. “The trial court must have the same kind of latitude in deciding how to test an expert’s reliability, and to decide whether or when special briefing or other proceedings are needed to investigate reliability, as it enjoys when it decides whether that expert’s relevant testimony is reliable.”⁴²

Kumho Tire involved a blowout leading to a serious crash, and the expert testimony proffered by plaintiffs was that of a tire failure analyst. Applying the *Daubert* factors as best it could under the circumstances, the district court rejected the testimony. The Eleventh Circuit reversed, holding that *Daubert* applied only to scientific expert evidence, which this was not. And the Supreme Court reinstated the district court’s exclusion.

Lurking in the background were other cases in which *Daubert* was held inapplicable to non-science expert testimony. Proponents of such testimony were beginning to discover that if they framed their experts as non-scientists, some courts would let them in the door without having to survive the rigors of *Daubert*.

In one case, following a *Daubert* hearing the district judge ruled: “[T]he testimony at the *Daubert* hearing firmly established that forensic document examination, despite the existence of a certification program, professional journals and other trappings of science, cannot, after *Daubert*, be regarded as ‘scientific ... knowledge.’” Because the witness was not offering anything scientific, however, *Daubert* was inapplicable; because *Daubert* was inapplicable, the proposed testimony could be evaluated against a looser standard; and therefore the testimony was admissible.⁴³

Another case in which the problem emerged involved two proffered expert witnesses on the question of fire causation.⁴⁴ One expert was a fire scientist whom the district court felt did not make it over the *Daubert* hurdle and was therefore excluded. But the other proffered expert was a fireman who had acquired experience examining fire scenes and reaching opinions on whether the fire had been started accidentally or intentionally. This was clearly not a scientist, thought the court; therefore *Daubert* did not apply; therefore the witness’s conclusions were admissible.

Kumho Tire impliedly reversed those kinds of cases and prohibited such reasoning in the future.

⁴² *Kumho Tire*, at 152.

⁴³ *United States v. Starzecpyzel*, 880 F. Supp. 1027 (S.D.N.Y. 1995). No doubt the testimony would have been admitted if the proponent could have demonstrated the requisite validity. Thus, the testimony would be admissible if it satisfied *Daubert* and also admissible if it failed *Daubert*. The point was not lost on handwriting examiners and other forensic scientists.

⁴⁴ *Michigan Millers Mut. Ins. Corp. v. Benfield*, 140 F.3d 915 (11th Cir. 1998).

Another important issue was the nature of the continuing role of “general acceptance.” If the key to *Daubert* is proof of validity of whatever the proponent is proffering, and if the “*Daubert* factors” were guides to establishing validity, could one of those factors alone supply a sufficient basis for a finding of validity, and might that factor be “general acceptance”? The Supreme Court realized that to equate general acceptance with validity would permit *Frye* to swallow *Daubert*. This was not allowed to happen: “*Daubert*’s general acceptance factor [does not] help show that an expert’s testimony is reliable where the discipline itself lacks reliability...” —giving as examples “any so-called generally accepted principles of astrology or necromancy.”⁴⁵

Finally, *Kumho Tire* reiterated the notion that knowledge which is determined to be valid must also be “relevant to the task at hand.”⁴⁶ The point here is that demonstrating that a body of knowledge or a field of expertise is reliable and relevant with respect to some tasks does not establish that it is valid with respect to the “task at hand” in the case at bench.

C. *Weisgram v. Marley* (2000)

The fourth case in the *Daubert* line is *Weisgram v. Marley*.⁴⁷ Damages were sought from the manufacturer of a heater alleged to have been defective, starting a fire, and causing the death of Bonnie Weisgram. Plaintiff offered three expert witnesses on the question of defect and its causal relationship to the fire. Over defendant Marley’s objection, the trial judge admitted the expert testimony. On appeal by defendant, the Eighth Circuit reversed the admission, holding that the experts’ testimony was not sufficiently sound and therefore should not have been admitted. The appeals court then considered the remaining evidence in the light most favorable to Weisgram, found it insufficient to support a jury verdict for the plaintiff, and directed judgment as a matter of law for Marley.

The question before the Supreme Court was whether the Court of Appeals, upon deciding that the plaintiff’s expert evidence was inadmissible, was required to remand for a new trial rather than direct judgment on its own.

The Supreme Court, in an opinion by Justice Ginsburg, unanimously rejected Weisgram’s argument that allowing courts of appeals to direct the entry of judgment for verdict-losing defendants will “punish plaintiffs who could have shored up their cases by other means had they known their expert testimony would be found inadmissible.”

Since *Daubert* ... parties relying on expert evidence have had notice of the exacting standards of reliability such evidence must meet. It is implausible to suggest, post-

⁴⁵ *Kumho Tire*.

⁴⁶ Id. See, also, D. Michael Risinger, *Defining the Task at Hand: Non-Science Forensic Science after Kumho Tire Co. v. Carmichael*, 57 WASH. & LEE L. REV. 767 (2000).

⁴⁷ *Weisgram v. Marley Co.*, 528 U.S. 440 (2000).

Daubert, that parties will initially present less than their best expert evidence in the expectation of a second chance should their first try fail.⁴⁸

Thus, in the space of seven years, *Daubert* had morphed from standing for “the Rules’ liberal thrust and their general approach of relaxing the traditional barriers to ‘opinion’ testimony” to a doctrine that imposed “exacting standards of reliability” on expert evidence.

VII. Amendments to Rule 702

Rule 702 has been amended “in response to *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993), and to the many cases applying *Daubert*, including *Kumho Tire Co. v. Carmichael* (1999).”⁴⁹ This might seem ironic, since *Daubert* was interpreting Rule 702 as it stood since being promulgated in 1975. Nevertheless, there has been something of a dialogue between judicial interpretations of the rule and the language of the rule itself. Bringing the two into closer alignment added clarity and made some of the Supreme Court’s requirements harder to miss.

An Appendix to this paper provides the language of Rule 702 from its origin in 1975 to the most recent amendments, scheduled to take effect on December 1, 2023.

A. 2000 Amendments

The 2000 Amendments served two purposes. One was to have the rule better reflect the Supreme Court’s views expressed in *Daubert*. The amended rule begins by preserving the original standard for expert witness qualification and for insuring that the expert’s testimony “will assist” the jury. Next, three numbered statements aim to reflect *Daubert*’s validity requirement: that the testimony is “based upon sufficient facts or data,” “is the product of reliable principles and methods,” and that “the witness has applied the principles and methods reliably to the facts of the case.” The Rule thus reflected the need for helpfulness, adequate foundation, relevance, and fit.

The second purpose was to provide an extensive Committee Note to assist judges in performing their gatekeeping responsibilities. The Advisory Committee⁵⁰ and the Court recognized that *Daubert*’s interpretation of Rule 702 created a different test for determining whether expert testimony is admissible, thereby imposing very different demands on trial judges than what some or many had experienced previously.

⁴⁸ *Weisgram*, at 455.

⁴⁹ Advisory Committee Comments to 2000 Amendments (citations omitted).

⁵⁰ Details about the work of the advisory committees on rules governing the federal judiciary can be found at: <https://www.uscourts.gov/rules-policies/about-rulemaking-process>. There, information can be found on the Rules Enabling Act (1934) governing the work of the rules committees, the process of rulemaking, how to propose changes to federal court rules, how to offer input on pending proposals, the membership of the various committees, and other information.

B. 2011 Restyling

In 2011, Rule 702 was amended as part of the restyling project that had been undertaken for the entirety of the Federal Rules of Evidence in order to make them “more easily understood and to make style and terminology consistent throughout the rules.” In keeping with the guiding principle of that project, restyled Rule 702 contains no changes in content or meaning. Its major elements became subparts. Subpart (c), “the testimony is the product of reliable principles and methods,” contains the heart of *Daubert’s* teaching.

C. 2023 Amendments

Rule 702 has been amended again (effective December 1, 2023), to address two concerns.

First, *Daubert* held that a trial court’s expert evidence admissibility decisions are governed by Rule 104(a). Under that rule it is the court’s duty to make the admissibility decision and not punt it to the jury. Moreover, citing *Bourjaily v. United States* (1987), the Supreme Court in *Daubert* held that Rule 104(a) requires the proponent to demonstrate that the requirements of admissibility are met by a preponderance of the evidence. In addition, the Advisory Committee’s note to the 2000 amendment to Rule 702 reiterated the applicability of that standard of proof. Nevertheless, many federal district court opinions employed an incorrect standard, and handed to the jury some of the issues that were the judge’s duty to resolve—incorrectly treating the sufficiency of an expert’s basis and the reliability the expert’s application as questions of weight for the jury. The 2023 amendment seeks to prevent future such errors by stating the applicable standard of proof in the rule itself.

A second purpose of the amendment is to address a common application problem, namely, experts’ overstating the conclusions that can be drawn from application of their discipline’s concepts and methods to the data of the case at bar. For example, an FBI study found that microscopic hair comparison experts exaggerated their reports and testimony beyond what their field was capable of—not occasionally, but more than 90% of the time.⁵¹

It is not enough for a gatekeeping judge to find that the expert relied on “sufficient facts or data” and that the expert applied “reliable principles and methods.” It is also necessary that expert opinions “stay within the bounds of what can be concluded from a reliable application of the expert’s basis and methodology.” Accordingly, subpart (d) now states that “the expert’s opinion [must reflect] a reliable application of the principles and methods to the facts of the case.” The Advisory Committee note to the amendment explains that the problem of experts exaggerating the conclusions that can be drawn from their techniques is “especially pertinent to the testimony of forensic experts in both criminal and civil cases.”

⁵¹ *Supra* note 1. This is the sort of problem that led to amending Rule 702 to require policing of expert opinions to prevent exaggeration. Rule 702(d) (requiring that each expert opinion must “stay within the bounds of what can be concluded from a reliable application of the expert’s basis and methodology.”).

VIII. Federal Expert Evidence Admissibility Rules in a Nutshell

We can summarize existing federal expert evidence admissibility doctrine quite briefly.

- A distinction is made between validity-of-the-asserted-expertise and qualifications-of-the-expert. [Rule 702]
- The trial court hears evidence and arguments in a Rule 104(a) hearing, also referred to as a *Daubert* hearing. [*Daubert*]
- At that hearing, the proponent must establish validity by a preponderance. [*Daubert*, *Bourjaily*, Rule 104(a), Rule 702]
- The court shall then admit, exclude, or order limitations on the expert’s testimony. [Rule 104(a)]
- The requirement of proving validity applies to non-novel as well as to novel proffers [*Daubert*]
- The requirement of proving validity applies to all expert evidence, not just “science” [*Kumho Tire*]
- General acceptance alone is insufficient to prove validity. [*Kumho Tire*]
- If the proponent’s expert evidence is admitted, the opponent may attack the testimony again at trial—but only with respect to weight and credibility. [Rule 104(e)]
- The standard of review on appeal is abuse of discretion. [*Joiner*]
- The court of appeals may itself enter judgment against proponents of expert evidence if their evidence had been admitted at trial, admission is found on appeal to have been an abuse of discretion, and exclusion leaves proponent without sufficient evidence to prevail at a retrial.⁵² [*Weisgram v. Marley*].

IX. Difficulties Applying Rule 702 as Aligned with the Supreme Court’s Cases

About any of the tests of admissibility, one could ask several questions. How good is the test theoretically? If the test is conceptually sound, are judges able to apply it effectively? Even if

⁵² Reflecting, in the space of seven years, a shift in the Court’s view of Rule 702 from reflecting the “liberal thrust of the Federal Rules” in *Daubert* to a test of “exacting standards” in *Weisgram*.

judges are able to perform the gatekeeping function, will they do so faithfully and honor its results even when they do not like where it leads? At the end of the day, does the legal test for admissibility of expert evidence, as applied, contribute to the effective sorting of valid from invalid proffers?

A. Are Judges Unable or Unwilling?

The question of whether *Daubert's* validity requirement places unattainable demands on judges was raised by Chief Justice Rehnquist in his concurrence in *Daubert*. While agreeing “that the *Frye* rule did not survive the enactment of the Federal Rules of Evidence,” he voiced doubts about what judges can be expected to do when evaluating expert testimony:

I defer to no one in my confidence in federal judges; but I am at a loss to know what is meant when it is said that the scientific status of a theory depends on its “falsifiability,” and I suspect some of them will be, too.

I do not doubt that Rule 702 confides to the judge some gatekeeping responsibility in deciding questions of the admissibility of proffered expert testimony. But I do not think it imposes on them either the obligation or the authority to become amateur scientists in order to perform that role.⁵³

Rehnquist’s concern echoes the “expert dilemma” discussed earlier⁵⁴ and repeated by others.⁵⁵ If experts are needed precisely because laypersons do not understand what the experts understand, how can the non-expert judge evaluate the validity of assertedly expert evidence for the benefit of the non-expert jury?

The Ninth Circuit panel evaluating the plaintiff’s expert proffers on remand of *Daubert* saw the paradox clearly, even poignantly:

As we read the Supreme Court’s teaching in *Daubert* ... though we are largely untrained in science and certainly no match for any of the witnesses whose testimony we are reviewing, it is our responsibility to determine whether those experts’ proposed testimony amounts to “scientific knowledge,” constitutes “good science,” and was “derived by the scientific method.”

The task before us is more daunting still when the dispute concerns matters at the very cutting edge of scientific research, where fact meets theory and certainty dissolves into probability. As the record in this case illustrates, scientists often

⁵³ *Daubert*, Rehnquist concurrence.

⁵⁴ *Supra* note 7 and accompanying text.

⁵⁵ See, e.g., Samuel R. Gross, *Expert Evidence*, 1991 WIS. L. REV. 1113 (1991) (“We call expert witnesses to testify about matters that are beyond the ordinary understanding of lay people (that is both the major practical justification and a formal legal requirement for expert testimony), and then we ask lay judges and jurors to judge their testimony.”).

have vigorous and sincere disagreements as to what research methodology is proper, what should be accepted as sufficient proof for the existence of a “fact,” and whether information derived by a particular method can tell us anything useful about the subject under study.

Our responsibility, then, unless we badly misread the Supreme Court’s opinion, is to resolve disputes among respected, well-credentialed scientists about matters squarely within their expertise, in areas where there is no scientific consensus as to what is and what is not “good science,” and occasionally to reject such expert testimony because it was not “derived by the scientific method.” Mindful of our position in the hierarchy of the federal judiciary, we take a deep breath and proceed with this heady task.⁵⁶

Even if judges are capable of carrying out *Daubert*’s teachings with aplomb, they might sometimes choose not to. Apparently fearing that some district judges would try to evade their gatekeeping obligations by taking the “flexibility” granted them as license to dilute their testing of proffered expert testimony, Justice Scalia warned that deficient scrutiny justifies reversal:

Trial-court discretion in choosing the manner of testing expert reliability is not discretion to abandon the gatekeeping function it is not discretion to perform the function inadequately. Rather, it is discretion to choose among reasonable means of excluding expertise that is *fausse* and science that is junky. The failure to apply one or another of [the *Daubert* factors] may be unreasonable, and hence an abuse of discretion.⁵⁷

B. Illustrations

Perhaps the clearest illustrations of judicial inability or unwillingness are found in their treatment of numerous forensic sciences. A review by a large interdisciplinary panel of the National Academy of Sciences (NAS) scrutinized those fields and their techniques, and concluded: “Much forensic evidence ... is introduced in criminal trials without any meaningful scientific validation, determination of error rates, or reliability testing”⁵⁸

That same panel—co-chaired by a prominent federal appellate judge—found the judges to be as bad at science as forensic scientists have been: “[F]orensic science professionals have yet to establish either the validity of their approach or the accuracy of their conclusions, and the courts have been utterly ineffective in addressing this problem.”⁵⁹

⁵⁶ *Daubert v. Merrell Dow Pharmaceuticals*, 43 F.3d 1311 (9th Cir. 1995) (citations omitted).

⁵⁷ *Kumho Tire*, Scalia concurrence, (joined by O’Connor and Stevens).

⁵⁸ NATIONAL RESEARCH COUNCIL, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES, A PATH FORWARD (Washington D.C.: National Academies of Science (2009)). See, also, PRESIDENT’S COUNCIL OF ADVISORS ON SCI. & TECH., EXEC. OFFICE OF THE PRESIDENT, FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY IF FEATURE-COMPARISON METHODS (2016).

⁵⁹ NRC Report, *supra* note 58.

Whether gatekeeping judges lacked the necessary understanding of how to evaluate scientific (empirical) claims or refused to apply their acuity to forensic “expertise that is *fausse* and science that is junky,” evidence of deficient screening is not hard to find.

A number of forensic sciences that had routinely been admitted into evidence began to disappear from the courts. But their departures came about not because courts recognized their junkiness, but because the fields themselves, or the wider scientific community, discovered that their techniques lacked validation if not validity. These fields include voiceprint identification,⁶⁰ comparative bullet lead analysis,⁶¹ and numerous arson indicators.⁶² Bitemark identification seems poised to join those others in the cemetery of unsound forensic sciences.⁶³

On occasion, judges recognized the weaknesses of some forms of expertise and imposed limitations on what experts could assert in their testimony.⁶⁴ More often, however, courts remained oblivious to fatal weaknesses.

Some forensic sciences have achieved such a degree of unthinking cultural acceptance that few judges have been able to rise above their own immersion in that culture.⁶⁵ When fingerprint identification was challenged for the first time under *Daubert*, prosecutors learned from their FBI examiners that no research existed testing the validity of their claims or supporting the breathtaking accuracy (zero error, 100% certainty) commonly asserted for it.⁶⁶

Apparently realizing that the proponent lacked the scientific evidence needed to survive a proper *Daubert* analysis, many judges found ways to avoid the analysis required by *Daubert* and *Kumho Tire*. As the treatise *Modern Scientific Evidence*, explains:

With few exceptions... judicial opinions reacting to challenges to asserted fingerprint identification expertise are united by their failure—typically, their refusal—to conduct any thoughtful analysis under *Daubert* and *Kumho Tire*. Some of the opinions contain virtually no *Daubert* analysis at all; others an inquiry that

⁶⁰ NATIONAL RESEARCH COUNCIL, ON THE THEORY AND PRACTICE OF VOICE IDENTIFICATION (1979).

⁶¹ NATIONAL RESEARCH COUNCIL, FORENSIC ANALYSIS: WEIGHING BULLET LEAD EVIDENCE (2004).

⁶² MODERN SCIENTIFIC EVIDENCE, supra note 3, Fire and Arson chapter.

⁶³ BITEMARK ANALYSIS: A NIST SCIENTIFIC FOUNDATION REVIEW (NIST Interagency Report 8352) (March, 2023). Also, see <https://www.txcourts.gov/media/1454500/finalbitemarkreport.pdf> (Texas Forensic Science Commission recommends moratorium on forensic bitemark identification in Texas courts).

⁶⁴ United States v. Glynn, 578 F. Supp. 2d 567 (S.D.N.Y. 2008) (firearms identification).

⁶⁵ One opinion revealed its author’s cultural immersion: “The court’s decision may strike some as comparable to a breathless announcement that the sky is blue and the sun rose in the east yesterday.” United States v. Havvard, 117 F.Supp.2d 848 (S.D.Ind. 2000). Compare with another judge’s awareness of the problem: “[F]ingerprint evidence has been afforded a near magical quality in our culture. In essence, we have adopted a cultural assumption that a government representative’s assertion that a defendant’s fingerprint was found at a crime scene is an infallible fact, and not merely the examiner’s opinion.” State v. Quintana, 103 P.3d 168 (Utah Ct. App. 2004).

⁶⁶ Llera-Plaza, supra note 30.

is no more than a parody of *Daubert* analysis; and those empty opinions then become something later cases can cite as justification for ducking their own gatekeeping responsibilities. Judges, like virtually everyone else in our culture, have grown up believing, without evidence or critical thought, that fingerprints are unique and that examiners are extremely accurate in all that they do. ***

Judges who have had to resolve such challenges appear to have been unable to adopt the necessary posture of skepticism long enough to see whether or not the proponents of the expert testimony can lift their claims over the *Daubert* hurdle. With few exceptions, the opinions all resort to one or another sort of evasion so that they can arrive at what they already “know” to be the “correct” conclusion, namely, that asserted fingerprint identification expertise satisfies the law’s admissibility requirements. What they actually do is to refrain from subjecting the proponents’ claims to the rigors of *Daubert*.⁶⁷

The treatise then reviews the various stratagems found in the cases. Here is that list:

Refusal to conduct a *Daubert* hearing

Reversal of the burden of persuasion

Ignoring *Kumho Tire*’s task-at-hand requirement

Avoidance of actual *Daubert* analysis⁶⁸

Turning *Kumho Tire* on its head⁶⁹

Reliance on admission by other courts

⁶⁷ MODERN SCIENTIFIC EVIDENCE, supra note 13, Fingerprint Identification chapter.

⁶⁸ The *Havvard* judge proceeded seriatim through the *Daubert* “factors,” substituting for each one an easier test. For example, scientific testing was replaced with courtroom testing which involved “the highest possible stakes—liberty and sometimes life.” Peer review of scientific studies was replaced with review of the examiner’s work by a peer. *Havvard* rejected the expert’s claim of zero error (based on no data) but substituted the court’s own belief that the error rate is “vanishingly small” (still based on no data).

⁶⁹ The court reviewing *Havvard* noted that, “The standards of *Daubert* . . . are not limited in application to ‘scientific’ testimony alone.” . . . Therefore, the idea that fingerprint comparison is not sufficiently ‘scientific’ cannot be the basis for exclusion under *Daubert*.” Therefore, admissibility was affirmed. *United States v. Havvard*, 260 F.3d 597 (7th Cir. 2001). As MODERN SCIENTIFIC EVIDENCE comments: “The Court of Appeals sought to remove fingerprinting from the realm of the empirical in order, apparently, to move it out of *Daubert*’s reach. Bizarrely, it relied on *Kumho Tire* to accomplish that, even though *Kumho Tire* stands for exactly the opposite proposition, namely, that there is to be no escape from appropriate scrutiny.” *Kumho Tire* had not reduced the obligation to ensure validity, it reinforced that duty. See, also, Scalia’s concurrence in *Kumho Tire*.

Reliance on general acceptance⁷⁰

Emphasis on flexibility of criteria⁷¹

Bringing the standards down to meet the expertise⁷²

Relegate to weight, not admissibility⁷³

Ironically, such efforts to find ways around the *Daubert/Kumho Tire* gauntlet do not so much shelter the favored expertise from serious scrutiny as they underscore that the proffered evidence failed to meet the law's requirements.

Also ironically, exempting weak expert evidence from the legal test helps insure it will remain weak. Shortly after the publication of the earthshaking NAS Report on forensic science, Barry Fisher, then director of the Los Angeles County Crime Laboratory and president of the American Academy of Forensic Sciences addressed judges at a conference: "You have given us a free ride. Unless there's some push to deal with these things [validation studies], there's not going to be any push on the part of any parent agency that runs laboratories to do these kinds of things. We are reactive. And if it turns out that an appellate court or a supreme court decision comes down and says 'you can't say this kind of stuff, you need to do the needed validation,' that will send a message"

Thus, what the courts do affects what the experts do. This impact is inevitably greatest for fields that have no life outside of courts.⁷⁴ Scientific knowledge developed within conventional industry, healthcare, universities, or even the consumer marketplace faces cultural and economic pressures for continued testing and improvement. But a field of asserted expertise that has only one ultimate audience (the courts) is in a far different situation. If its audience insists upon high

⁷⁰ See, *Kumho Tire* (rejecting general acceptance as a sufficient basis by itself for approving admission). See, *supra* note 41. See, also, Scalia's concurrence in *Kumho Tire*, *supra* note 57.

⁷¹ See, *id.*

⁷² On realizing that asserted fingerprint identification expertise was not scientific, had not been well tested, lacked "traditional" scientific testing, and lacked a body of peer reviewed research literature, a court concluded that those shortcomings meant that alternative, less rigorous, criteria needed to be employed (in order to facilitate admission).

U.S. v. Cline, 188 F. Supp. 2d 1287 (D. Kan. 2002). For a similar example involving handwriting identification, see *U.S. v. Starzeczpyzel*, *supra* note 43. Again, also, Scalia's concurrence in *Kumho Tire*, *supra* note 57.

⁷³ See Commentary to 2023 Amendments to Rule 702 ("[M]any courts have held that the critical questions of the sufficiency of an expert's basis, and the application of the expert's methodology, are questions of weight and not admissibility. These rulings are an incorrect application of Rules 702 and 104(a)."

⁷⁴ Fields that have a life outside the courtroom—serving industry, patients, consumer markets, an academic science community—have other pressures driving their continual improvement.

quality, sound science, the expert field would almost certainly rise to the challenge. Because the courts do not insist, the forensic disciplines do not rise.⁷⁵

X. Evolution of Admissibility Rules, Past and Future

So long as the law believes that expert evidence must cross a higher threshold to gain admission, and so long as existing rules and practices fail to achieve the desired filtering, the rules governing expert evidence admissibility will continue to evolve. As we have seen, expert evidence admissibility rules have been advanced, tweaked, revised and tweaked again in a centuries-long search for the right recipe. Rules of procedure have also been part of that mix.⁷⁶

In any given era, there have been optimists, urging that with another tweak, or the right kind of help, judges will do better what the rules call upon them to do.⁷⁷

Others believe that fundamental flaws in the process of presenting expert evidence to factfinders, with or without judicial gatekeeping, can be mended only through a legal paradigm shift. Learned Hand offered one of the earliest of these proposals.⁷⁸ Most recently, a leading evidence scholar, Edward Cheng, has argued that the whole *Daubert* project has proved unworkable. Cheng begins by observing:

Founded on good intentions but unrealistic expectations, the dominant *Daubert* framework for handling expert and scientific evidence should be scrapped. *Daubert* asks judges and jurors to make substantively expert determinations, a task they are epistemically incompetent to perform as laypersons.⁷⁹

Having placed its lay decisionmakers in impossible positions, the *Daubert* regime dooms itself to suboptimal decisions. And while critics are quick to blame the decisionmakers, the fault lies not with them, but with the underlying structure.⁸⁰

⁷⁵ We can see the irony in Judge Kozinski's 1993 footnote seeking to protect forensic sciences from *Daubert* scrutiny when he realized they would flunk, and his later castigation of them in the *Wall Street Journal*. He scolded them for being what he helped make them.

⁷⁶ See, most notably, rules governing pre-trial discovery, permitting advisory juries, permitting judicial advisors, and court-appointed experts (Rule 706—which seems as much a rule of procedure as of evidence). See, Daniel L. Rubinfeld and Joe S. Cecil, *Scientists as Experts Serving the Court*, 147 DAEDALUS 152 (Fall, 2018).

⁷⁷ As an example of another amendment, see, David E. Bernstein and Eric G. Lasker, *Defending Daubert: It's Time to Amend Federal Rule of Evidence 702*, 57 WM. & MARY L. REV. 1 (2015) (proposing small changes in the language of Rule 702 they believe will produce large improvements in gatekeeping). As an example of educational assistance, the Federal Judicial Center puts on judicial education programs and publishes its SCIENTIFIC EVIDENCE REFERENCE MANUAL. Academics try to assist by producing other treatises, e.g., MODERN SCIENTIFIC EVIDENCE, supra note 3; and PAUL C. GIANNELLI ET AL., SCIENTIFIC EVIDENCE (2012).

⁷⁸ Hand, supra note 6.

⁷⁹ Edward K. Cheng, *The Consensus Rule: A New Approach To Scientific Evidence*, 75 VAND. L. REV. 407, 407 (2022).

⁸⁰ Id., at 419-420.

Discussion of the substance of these proposals, and others,⁸¹ is a topic for another day. But they are out there. When dissatisfaction with the current rules and procedures grows sufficiently large or loud,⁸² we shouldn't be surprised to see more evolution in the future—perhaps tweaks, perhaps changes more fundamental.

⁸¹ See, e.g., Samuel Gross, *supra* note 55; Arthur Kantrowitz, *Proposal for an Institution for Scientific Judgment*, 156 SCIENCE 763(1967); Andrew W. Jurs, *Science Court: Past Proposals, Current Considerations, and a Suggested Structure*, 15 VA. J. L. & TECH. 1 (2010).

⁸² The legal literature contains many critiques. For a few good examples, see, Bernstein & Lasker, *supra* note 77; *Modern Scientific Evidence*, *supra* note 3; NRC REPORT, *supra* note 58. Another example might be the Commentary to the 2023 Amendments to Rule 702.

**Appendix: Changes in Rule 702, Testimony by Experts,
from the Original (1975) Rule through Current Amendments (2023)**

Original 1975

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

Amended 2000

If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise, if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of reliable principles and methods, and (3) the witness has applied the principles and methods reliably to the facts of the case.

Restyled 2011

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.

Amended 2023 (Effective December 1, 2023)

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 the proponent demonstrates to the court that it is more likely than not that:

- (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's opinion reflects a reliable application of the principles and methods to the facts of the case.