

Judging Admissibility

Michael J. Saks*

I. INTRODUCTION	135
II. ADMISSIBILITY LAW	136
<i>A. The Centuries Before Frye</i>	136
<i>B. Frye's Innovation</i>	137
<i>C. The Decades After Frye</i>	139
<i>D. Daubert and the Federal Rules</i>	139
<i>E. The Daubert Quartet</i>	142
III. APPLYING <i>DAUBERT</i> : ONE RULE (MORE OR LESS) IN MANY CONTEXTS.....	144
<i>A. Civil Versus Criminal Cases</i>	144
<i>B. Forensic Science and the Courts Generally</i>	146
<i>C. Voiceprint Identification</i>	147
<i>D. Fire and Arson Investigation</i>	149
<i>E. Fingerprint Identification</i>	150
<i>F. Microscopic Hair Comparison</i>	150
<i>G. Toxicology and Epidemiology</i>	151
<i>H. Economics Expert Witnesses</i>	154
IV. CONCLUSION	155

I. INTRODUCTION

The aim of this Article is to provide those thinking about the admissibility of economics expert testimony on damages in mass securities litigation under *Daubert* with a broader perspective on the law of expert evidence admissibility. This will be accomplished by looking at what courts have done about admissibility more generally, with other kinds of expertise, under varying tests, over a longer span of history.

* Regents Professor, Sandra Day O'Connor College of Law and Department of Psychology, Arizona State University. Ph.D., Ohio State University; M.S.L., Yale Law School.

II. ADMISSIBILITY LAW

A. *The Centuries Before Frye*

Expert testimony was proffered to courts long before 1923, when the *Frye* test was announced.¹ The first reported decision affirming the propriety of the use of what were then referred to as “skilled witnesses” was in 1782, in the case of *Folkes v. Chadd* (engineers).² One of the earliest trials using skilled witnesses of which we have a record occurred in 1699 in the murder trial of Spencer Cowper (physicians),³ though the judge in that trial plainly seemed to be accustomed to the use of expert witnesses. So it seems that at least as early as the seventeenth century, expert witnesses had become familiar in common law trials.

What did courts do in the centuries before *Frye* to screen expert testimony, if they did filter expert testimony? The practice varied. Many courts appeared not to be using any special test to assess the admissibility of expert testimony. (As today, in English courts the judges applied the basic test of relevancy plus whatever discretionary judgment they chose to apply.⁴) Others took another step, assessing expert witnesses’ qualifications (meaning their credentials to be regarded as possessing the expertise of those in their field).⁵ Exclusive use of that test presupposed that an expertise existed, and the admissibility question was whether the proffered witness was a person who had command of that expertise.

Cases going back at least to the American Civil War show judges using what could be termed the “marketplace test.”⁶ In many cases, judges showed an unmistakable interest in the commercial success (outside of litigation) of the proffered witness in his or her field.⁷ The implicit measure of expertise seems to have been how the expert witness fared in the commercial market for the witness’s learning.⁸ If a person could make a living selling the knowledge at issue, then expertise presumably existed.⁹ “Although courts sometimes spoke of an expert’s ‘greater study respecting certain subjects’ or having ‘made the subject upon which he gives his opinion a matter of particular study,’” it is clear that some degree of prosperity in the practice of the occupation or profession

1. *Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923).

2. *Folkes v. Chadd*, (1782) 99 Eng. Rep. 589, 590 (K.B.).

3. *The Trial of Spencer Cowper*, 13 Howell’s State Trials 1105 (1699). The factual issue was whether the decedent had first been killed and then thrown into a river, or whether she drowned in the river. *Id.* at 1107. Some of the physician expert witnesses based their opinions of the autopsy observations on experiments they conducted in which they compared postmortem findings of dogs they killed and then submerged in water to those of dogs they drowned. *Id.* at 1159.

4. British courts traditionally, and as of this writing, have no rules attempting to regulate the reliability of expert testimony. See MIKE REDMAYNE, *EXPERT EVIDENCE AND CRIMINAL JUSTICE* (2001) (especially ch. 5).

5. DAVID L. FAIGMAN ET AL., *MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY* § 1:2 (2008–2009).

6. David L. Faigman, Elise Porter & Michael J. Saks, *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, 1804–05 (1994).

7. *Id.* at 1804.

8. *Id.*

9. *Id.* at 1804–05.

claiming that knowledge almost always accompanied the expertise.¹⁰ “In effect, the marketplace determined whether valid knowledge existed by endowing it with commercial value.”¹¹

The marketplace test can be seen to have some serious problems.¹² One problem is that the market can tell us only what people select; it cannot tell us whether what they select is any good. Thus, for example, the marketplace test is incapable of distinguishing astronomy from astrology. The market values both. Commercial value is not a measure of scientific or any other kind of validity.¹³ Another problem is that some fields have little or no life in any commercial marketplace.¹⁴ That is true of cutting-edge knowledge which has yet to develop a market for itself, and of fields that have little or no function outside of their possible courtroom utility (sometimes signaled by the adjective “forensic,” as in “forensic science”). “The courtroom is their marketplace.”¹⁵

B. Frye’s Innovation

In light of the tests of admissibility that existed when *Frye v. United States* arose, we can both understand the necessity for inventing the *Frye* test and appreciate the actual importance of what could be viewed as a minor corollary to the marketplace test. In proffering an early form of polygraph testing, the defendant presented the court with an unfamiliar problem.¹⁶ How was the validity of such asserted expertise to be evaluated and a judgment made as to its admissibility? The technique was new.¹⁷ There was no professional field of polygraph examination yet, and no developed market for their services.¹⁸ Perhaps the technique was valid, perhaps not. Judge Van Orsdel might have realized that no solution could be found in the marketplace and that a modification of the older test was needed to accommodate the unusual circumstances.

Judge Van Orsdel’s entire opinion took up only two pages of the Federal Reporter. The critical language is:

Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.¹⁹

10. *Id.* at 1804.

11. Faigman, Porter & Saks, *supra* note 6, at 1804.

12. *See* FAIGMAN ET AL., *supra* note 5, § 1:2.

13. *Id.*

14. *Id.*

15. *Id.* (emphasis omitted).

16. *Frye v. United States*, 293 F. 1013, 1013–14 (D.C. Cir. 1923).

17. *Id.*

18. *See id.* (referring to polygraph examination as a theory only and as a “systolic blood pressure deception test,” thus alluding that there was neither a professional field of polygraph testing nor a developed market).

19. *Id.* at 1014.

At bottom, this is nothing more than the old marketplace test deployed in a different market. Where there is no commercial market, the intellectual or professional marketplace could be substituted. The *Frye* test is still a marketplace test, the real evaluation is still conducted outside of the court and outside of the law, and it is still incapable of distinguishing astrology from astronomy. But the *Frye* corollary did several new and important things. First, of course, the alternative marketplace allows knowledge to be evaluated even if the knowledge is too new to be marketed commercially or even if there is no hope of ever marketing it commercially.

Second, the *Frye* test separated the expertise from the expert. This created explicit legal recognition of the notion that a body of asserted knowledge has an existence apart from any individual, and that it is that body of asserted knowledge that must first be evaluated, and then, only if it is judged sound, does it make sense to assess the qualifications of the particular individuals who seek to bring that knowledge to court.

Third, *Frye* replaced buyers with sellers as the real judges of the validity of the offerings. While the commercial marketplace test posed the implicit question, “Do consumers of an asserted expertise believe those asserted experts are able to do what they claim to be able to do?,” the intellectual marketplace test (*Frye*) posed the question, “Do producers of an asserted expertise believe themselves able to do what they claim to be able to do?” This is a serious problem. The commercial marketplace test, despite its serious weaknesses, had the virtue of allowing buyers to assess the value of purported expertise and whether the expertise and the expert were “therefore” valid. Under the *Frye* variant, that control was transferred to the people who produced the knowledge and who offered it (and themselves) to the courts.

Fourth, and somewhat paradoxically, while the *Frye* test seems to place all of the real control in the hands of the producers of asserted knowledge to say whether that knowledge is sound (and therefore admissible) or not, it developed that the judges controlled a lever which surreptitiously put much or most of the control in their hands. The question framed above—“Do producers of an asserted expertise believe themselves able to do what they claim to be able to do?”—is the narrower of two versions of the same question. The broader version asks: “Do producers of an asserted expertise *and others with relevant knowledge* believe the asserted experts are able to do what they claim to be able to do?” If not made clear in law (and it was rarely if ever made clear in law),²⁰ a judge could ask one of those questions of one proffered expertise and the other of those questions of a different proffered expertise. And if the narrow question were to be posed (e.g., Do astrologers agree that astrology is valid?) the answer would almost certainly be yes, while if the broader question were posed (Do astrologers and other knowledgeable fields (astronomy, psychology, statistics, etc.) agree that astrology is valid?), the answer would very likely be no. One of the clearest examples of this is the body of cases evaluating voiceprint identification under *Frye*. Every court that posed the narrow question concluded that voiceprints were generally accepted and therefore admissible. Every court that posed the broad question concluded that voiceprints were not generally accepted and therefore were not admissible.²¹

20. See FAIGMAN ET AL., *supra* note 5, §§ 1:5, 1:6, 37:1.

21. FAIGMAN ET AL., *supra* note 5, § 37:1, at 2.

C. The Decades After Frye

If the first great myth of the *Frye* test is that it was a revolutionary judicial invention, the second great myth is that it thereafter dominated the courts' scientific admissibility decisions until the Supreme Court's decision in *Daubert v. Merrell Dow Pharmaceuticals*.²² In actuality, the *Frye* corollary went unnoticed for decades. *Frye* was not cited by a single other court, federal or state, for a decade. Even its inventor, Judge Van Orsdel, ignored it in another landmark scientific evidence case he handed down on the very same day he issued the *Frye* opinion.²³ During the first quarter-century after its publication, *Frye* was cited in only eight federal cases and five state cases. During its second quarter-century, it was cited 54 times in federal cases and 29 times in state cases.²⁴ The *Frye* test was not really "discovered" until around the time codification of the Federal Rules of Evidence was underway and, ironically, even more so after the Federal Rules were adopted. Consequently, in the 1980s *Frye* was being cited as much each year as it had been in its first 50 years added together.

D. Daubert and the Federal Rules

The Federal Rules of Evidence, adopted in 1975, rejected the *Frye* corollary in favor of a test that focused on the demonstrable validity of the proffered expert testimony. Of course, this was not appreciated about the Federal Rules until the Supreme Court's unanimous decision in *Daubert* in 1993. Until that time, most of the federal circuits and half the states incorporated *Frye* into Rule 702—despite the absence of any language in the Rules or the Advisory Committee Comments invoking *Frye* or the concept of general acceptance. *Daubert* concluded that *Frye*'s "general acceptance"²⁵ test "was superseded by the adoption of the Federal Rules of Evidence" and that "the Rules occupy the field."²⁶

The test embodied in Rule 702, according to *Daubert*, requires the trial court to conduct a preliminary hearing in which it decides whether "the testimony's underlying reasoning or methodology is scientifically valid and properly can be applied to the facts at issue."²⁷ The application of that test requires a court to assess the scientific validity of the proffered testimony with the help of several non-exclusive guidelines. The first three of these *Daubert* "factors," as they often are termed, bear a striking resemblance to the major sections of a conventional scientific journal article.

First is testability or falsifiability: "whether [the subject matter] can be (and has

22. *Daubert v. Merrell Dow Pharms.*, 509 U.S. 579 (1993).

23. *See generally* *Laney v. United States*, 294 F. 412 (D.C. Cir. 1923) (ruling on the admissibility of firearms identification). *Laney* not only made no use of the *Frye* test, it made no mention of it, and did not explain why it was not applied or applicable to the novel question of firearms identification. *Id.* Moreover, a leading state supreme court had that same year denied admissibility to firearms identification. *People v. Berkman*, 139 N.E. 91 (Ill. 1923). Perhaps Judge Van Orsdel foreshadowed later judges by using the *Frye* test merely as a legal tool to be used or not used depending on the outcome desired.

24. Faigman, Porter & Saks, *supra* note 6, at 1808 n.25.

25. *Daubert*, 509 U.S. at 588.

26. *Id.* at 587.

27. *Id.* at 592.

been) tested.”²⁸ This resembles the introduction section to a scientific journal article, in which past tests are described and the currently contemplated test of the subject is explained.

Second is assessment of the quality of the methodology: “Peer review and publication . . . [reflecting] submission to the scrutiny of the scientific community . . . increases the likelihood that substantive flaws in methodology will be detected.”²⁹ This parallels the second section of a scientific journal article, the delineation and justification of the research methods to be used. This factor is widely misconceived as mere “peer review and publication” because those are the first few words of the paragraph in the Court’s opinion explaining it, and perhaps because that superficial test is easier than the real thinking called for. But a careful reading of the full paragraph leads one to a deeper realization: the district court must assess the methodology of the proffered research (offered to support the asserted expertise which is the subject of a *Daubert* hearing) and satisfy itself that the results that emerge from such studies are therefore sound. Judges would need help in doing that, and the peer review that precedes publication and (probably more so) the peer review that follows publication (as other researchers comment on a study’s strengths and weaknesses) will provide “the scrutiny of the scientific community” that will help a court “detect” “substantive flaws in methodology.”³⁰ Sound research design, leading to sound results, is the real point—not the superficial facts of peer review and publication.

The third factor parallels the results section of a scientific journal article: “known or potential rate of error.”³¹ What are the findings of the testing that was conducted using sound research methods?

Concluding the marketplace analogy, by moving the central axis of decision-making from marketplace consumers and producers to courtroom judges, *Daubert*, at least formally, directs judges to behave like modern and informed participants in the transaction, which befits the ultimate consumers of expert knowledge that they are: no more shortcuts.

Many courts and commentators have puzzled over whether *Daubert* lowered the threshold for admission of expert testimony or erected a higher barrier. An answer can be found by looking at the varied terrain of cases that have tried to apply *Daubert*, and from looking more at what *Daubert* does than what it says.

Frye and *Daubert* focus on two different attributes of asserted expert knowledge. Whether a proffered scientific theory or technique passes muster should depend upon how it fares on the chief attribute of the particular test.

28. *Id.* at 593.

29. *Id.*

30. *Id.*

31. *Daubert*, 509 U.S. at 594.

Figure 1. A Simplified Comparison of Admission Decision-Making Under *Frye* and *Daubert*

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

As Figure 1 illustrates, any given theory or technique can be based on a strong scientific foundation or a weak one and, independently, it can enjoy “general acceptance in the particular field in which it belongs”³² or not enjoy general acceptance. A theory or technique that is based upon valid science and enjoys general acceptance would be admitted by either test (upper left cell). A theory or technique that is not based upon valid science and does not enjoy general acceptance would be excluded by either test (lower right). The interesting situations are those in which the two attributes are discordant for any given type of scientific evidence. Where proffered knowledge is based on a solid scientific foundation but has not yet gained general acceptance within its field, *Frye* would exclude the testimony but *Daubert* would admit it (lower left). This is the situation that is usually envisioned when the two tests are compared, leading some to see *Frye* as the more conservative test. But where proffered knowledge has only a weak scientific foundation and yet enjoys general acceptance within its field, the *Frye* test will admit the testimony, but the *Daubert* test will exclude it (upper right). In this situation, the *Frye* test is not conservative at all. Among those fields where this particular discordance exists, *Daubert* has from time to time provoked deep concern. The fields that occupy this cell—and have historically received the most generous welcome for the weakest science—have been the forensic sciences.³³ *Daubert* subjects those fields to more rigorous scrutiny than *Frye* had, and the field’s practitioners realized that before the courts did.³⁴

Despite language in *Daubert* itself suggesting that the Supreme Court regarded *Daubert* as a more liberal test (referring as it does to the “liberal thrust of the Federal Rules”³⁵), in only seven years the Court re-perceived *Daubert* as a much more austere

32. It should be obvious to the reader that this diagram and accompanying discussion are a considerable oversimplification, but helpful in making a subtle point. At a minimum, one ought to realize that these two tests are not two hard categories, or the only categories, but that the scientific soundness and general acceptance of any proposition exist in degrees and fall along a continuum. Second, as discussed earlier, the *Frye* question itself comes in at least two flavors (broad and narrow), and probably also in degrees. A host of other complications emerge from topic to topic and case to case.

33. See FAIGMAN ET AL., *supra* note 5, chs. 32–34 (discussing traditional forensic sciences, namely handwriting, fingerprints, and toolmarks); NAT’L RESEARCH COUNCIL, COMM. ON IDENTIFYING THE NEEDS OF THE FORENSIC SCI. CMTY., STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD (2009) [hereinafter STRENGTHENING FORENSIC SCIENCE].

34. Brief for Americans for Effective Law Enforcement, et al. as Amici Curiae Supporting Respondents, *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1998).

35. *Daubert*, 509 U.S. at 588.

standard. In *Weisgram v. Marley Co.*³⁶ the Court held that *Daubert* had put parties on notice that they faced a high hurdle, and if they failed to surmount it in a district court, the court of appeals could both affirm the trial court's evidentiary ruling and dispose of the case entirely, without remand.³⁷

The final *Daubert* "factor" was a tip of the hat to *Frye*: "Finally, 'general acceptance' can yet have a bearing on the inquiry."³⁸ The Justices must have suspected that district judges would look for easy solutions and try to slip evidence around the barriers the Supreme Court had erected, and that general acceptance might be the vehicle for doing so. In *Kumho Tire Co. v. Carmichael*,³⁹ the Supreme Court explained that if a field can find support in no factor other than "general acceptance," then it has inadequate support to be admitted. In *Kumho Tire*, Justice Scalia concurred, saying:

[D]iscretion in choosing the manner of testing expert reliability . . . is not discretion to abandon the gatekeeping function. [I]t 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. Though, as the Court makes clear today, the *Daubert* factors are not holy writ, in a particular case the failure to apply one or another of them may be unreasonable, and hence an abuse of discretion.⁴⁰

E. The Daubert Quartet

Daubert is accompanied by three other Supreme Court decisions which expanded and refined its meaning. *General Electric v. Joiner*⁴¹ held that appellate review of trial court decisions to admit or exclude is to be deferential, reversing only for abuse of discretion.⁴² Moreover, the Supreme Court held that the logic by which the expert traveled from principles and evidence to a conclusion also is subject to appraisal by the court.⁴³ The notion that old decisions to admit or exclude can be rechallenged again and again is reinforced by the standard of review adopted in *Joiner*, which was abuse of discretion.⁴⁴ This deferential standard of review assumes that each submission will vary from case to case, and needs to be dealt with in an individualized way by the trial judge; that, by definition, appellate judges are not to review the evidence on the evidence de novo; that affirmances generally do not establish any legal precedent; and therefore affirmances by reviewing courts have no necessary application to subsequent cases. There are good arguments that this is an unwise and inappropriate standard of review in the context of many, if not most, questions of admissibility of expert testimony.⁴⁵ But,

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

37. *Id.* at 455–57.

38. *Daubert*, 509 U.S. at 594.

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

40. *Id.* at 158–59.

41. *Gen. Elec. v. Joiner*, 522 U.S. 136 (1997).

42. *Id.* at 139.

43. *Id.*

44. *Id.* at 146.

45. See generally Michael J. Saks, *The Aftermath of Daubert: An Evolving Jurisprudence of Expert Evidence*, 40 JURIMETRICS J. 229 (2000).

for now, that is the law.

*Kumho Tire Co. v. Carmichael*⁴⁶ held that *Daubert*'s essential evidentiary reliability requirement applies to all fields of expert evidence, not only to science.⁴⁷ For non-science expertise (that is, expertise on questions that are seldom the topic of systematic empirical investigations), therefore, courts might have to develop new criteria for evaluating the soundness of proffered expert evidence. *Kumho Tire* underscores *Daubert*'s direction to judges to focus narrowly on the task at hand, that is, not to evaluate expertise in global generalities, but instead to focus on specific asserted knowledge proffered to answer specific factual issues in the case before them. All of this is a daunting assignment. It may be more apparent now than it was in earlier centuries why judges sought ways to avoid such responsibility (and why, notwithstanding the commands of *Daubert*, many of them still do). Because *Daubert* seemed to apply only to "scientific" expert evidence, some fields had taken refuge in that distinction, arguing that they were not sciences and therefore need not be evaluated under *Daubert* (even though in decades past those same fields had slipped through the gates of admissibility by claiming to be sciences).⁴⁸ The Supreme Court put an end to that gambit by its holding in *Kumho Tire*. Evaluation criteria would have to change, depending on the nature of the claimed expertise. The ways to evaluate science are not the ways to evaluate art history. Courts would have to find suitable criteria and apply them in order to satisfy the gatekeeping requirement that only sound, valid expert evidence is admissible. And, as noted earlier, under *Kumho Tire*, general acceptance alone is not enough.

Finally, *Weisgram v. Marley* held that if a party's expert evidence was excluded at trial, the exclusion was upheld on appeal, and the expert evidence was necessary for the party's prima facie case, then the case was over and that party lost.⁴⁹ Thus, in the eyes of the Supreme Court, in a mere seven years *Daubert* went from being a test that represented "the liberal thrust of the Federal Rules"⁵⁰ to being an austere and rigorous criterion that counsel must approach with great seriousness: "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-*Daubert*, that parties will initially present less than their best expert evidence in the expectation of a second chance should their first try fail."⁵¹

Some other details are worth noting. Whereas *Frye* assumed that any "novel" evidence would be challenged at once, scrutinized, and admitted or excluded, apparently as a matter of law, *Daubert* rejected the notion that only "novel" evidence was subject to challenge and judicial evaluation. Wrote the Court: "[a]lthough the *Frye* decision itself focused exclusively on 'novel' scientific techniques, we do not read the requirements of Rule 702 to apply specially or exclusively to unconventional evidence. Of course, well-established propositions are less likely to be challenged than those that are novel, and

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

47. *Id.* at 147.

48. See, e.g., FAIGMAN ET AL., *supra* note 5, at chs. 33, 38 (explaining the admissibility of questionable sciences such as handwriting analysis and fire investigation).

49. *Weisgram v. Marley*, 528 U.S. 440, 441 (2000).

50. *Daubert v. Merrell Dow Pharms.*, 509 U.S. 579, 588 (1993).

51. *Weisgram*, 528 U.S. at 455–57.

they are more handily defended.”⁵²

Finally, although *Daubert* was written as though it was doing no more than explicating the Federal Rules of Evidence, the Rules themselves were amended in the wake of *Daubert*, in 2000, in order to come more fully in line with *Daubert*.⁵³ In all, the Supreme Court’s four major *Daubert* cases reflect an increasingly strong insistence on a demonstration of validity as a condition of admission.

III. APPLYING *DAUBERT*: ONE RULE (MORE OR LESS) IN MANY CONTEXTS

Recall Justice Holmes’s axiom that “[t]he prophecies of what the courts will do in fact, and nothing more pretentious, are what I mean by the law.”⁵⁴ Knowing the announced admissibility rules, as reviewed in the preceding section, gets one only so far. Any statement of the black letter law should be accompanied by knowledge of what judges actually do in categories of actual cases when faced with actual proffers of expertise and actual challenges by parties.

A. Civil Versus Criminal Cases

In civil cases and especially tort cases, judges can be seen to enforce *Daubert* aggressively and often insightfully, showing considerable acumen about research methodology.⁵⁵ In other categories of cases, judges appear to be either incapable of applying *Daubert* to the expertise before them, or unwilling to do so, and find ways to evade the burden or to hedge the result that would have emerged if they had conscientiously undertaken the burden *Daubert* imposes on judges.⁵⁶ These latter categories certainly include criminal cases, especially where the government proffers crime laboratory experts whose expertise purports to link evidence from the crime scene

52. *Daubert*, 509 U.S. at 592 n.11.

53. Rule 701, which governs lay witness testimony, was amended to prevent proffered experts who failed *Daubert* scrutiny to offer their opinions as lay witnesses. FED. R. EVID. 701. Rule 702 essentially codified *Daubert* by adding three numbered clauses:

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.

FED. R. EVID. 702. Rule 703 was amended to limit otherwise inadmissible hearsay statements from coming in as evidence. The Rule now provides, in relevant part: “[f]acts or data that are otherwise inadmissible shall not be disclosed to the jury by the proponent of the opinion or inference unless the court determines that their probative value in assisting the jury to evaluate the expert’s opinion substantially outweighs their prejudicial effect.” FED. R. EVID. 703.

54. Oliver Wendell Holmes, *The Path of Law*, 10 HARV. L. REV. 457, 460–61 (1897).

55. See D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?*, 64 ALB. L. REV. 99 (2000); Peter J. Neufeld, *The (Near) Irrelevance of Daubert to Criminal Justice and Some Suggestions for Reform*, 95 AM. J. PUB. HEALTH S107 (2005).

56. See Risinger, *supra* note 55. See also *infra* Parts III.C–E (discussing various cases involving issues of admissibility of forensic science expert evidence).

to the defendant. In these categories of cases, the principles of *Daubert* seem to vanish.⁵⁷

Since the explanation is not to be found in the quality of the science or the witnesses,⁵⁸ one must look elsewhere. Perhaps lawyers in criminal cases are not as vigilant or their challenges as cogent as their counterparts on the civil side. Perhaps the courts assume that criminal defendants are so likely to be guilty that anything offered against them contributes to a correct verdict, and if admitted in error the error will almost inevitably be deemed harmless.⁵⁹ Perhaps the problem is the failure of the adversary system to present issues the judges could rule upon, and at the root of that is the asymmetry of resources that has allowed the government, for decade upon decade, to present expert testimony while the defense had nothing with which to challenge or counter.⁶⁰ When only one side has experts, and there is no fight, perhaps courts get the mistaken impression that all is well with the expertise. Ironically, exactly the opposite will be true.⁶¹ When experts come from only one side, and they go unchallenged for an extended period of time, they are free to do whatever they want: present non-science, stray from real science, even fabricate findings, and do it all with impunity.⁶² As one crime laboratory director said to judges and lawyers at a recent conference on the NRC Report, “You have given us a free ride.”⁶³ Whatever the explanation, judges do not appear to be as vigilant in criminal cases as they are in civil cases.

We turn next to some specific areas of expert testimony, beginning with the forensic sciences, seeking to draw some general lessons about what courts do when evaluating the admissibility of expert evidence.

57. After spending more than two years reviewing forensic science at the direction of Congress, a committee of the National Research Council concluded, “The bottom line is simple: In a number of forensic science disciplines, forensic science professionals have yet to establish either the validity of their approach or the accuracy of their conclusions, and the courts have been utterly ineffective in addressing this problem.” STRENGTHENING FORENSIC SCIENCE, *supra* note 33, at 53. The point was made even more bluntly elsewhere: “There is almost no expert testimony so threadbare that it will not be admitted if it comes to a criminal proceeding under the banner of forensic science.” Jane Campbell Moriarty & Michael J. Saks, *Forensic Science: Grand Goals, Tragic Flaws, and Judicial Gatekeeping*, 44 JUDGES J. 16, 28 (2005).

58. The expert evidence offered (usually by the prosecution) in criminal cases embodies science that is no more sound, nor the experts better educated or more accomplished than that which is seen in civil cases. As to the substance of the science on the criminal side: “there is a notable dearth of peer-reviewed, published studies establishing the scientific bases and validity of many forensic methods.” STRENGTHENING FORENSIC SCIENCE, *supra* note 57, at 8. As to the personnel, the great majority of crime laboratory experts have no more than bachelor’s degrees, and sometime less than that. K. G. Furton, Y. L. Hsu & M. D. Cole, *What Educational Background Do Crime Laboratory Directors Require from Applicants?*, 44 J. FORENSIC SCI. 128, 128–29 (1999).

59. Which is literally to pre-judge the case.

60. The government has crime labs dedicated to serving police and prosecution needs. The defense has no institutional resources and typically no resources at all with which to hire ad hoc experts to scrutinize, re-analyze, or help think about the government’s expert’s report and testimony.

61. As anyone who believes in the adversary system should expect.

62. See Brandon L. Garrett & Peter J. Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions*, 95 VA. L. REV. 1 (2009) (finding misstatements of the science, exaggeration, and fabrication to be surprisingly frequent, defense attorney challenges rare, and judicially granted relief rarer).

63. Barry A. J. Fisher, L.A. County Sheriff’s Dep’t, comment made during colloquy following his presentation, *Monopoly is a Bad Institutional Structure*, at the Sandra Day O’Connor College of Law Center for the Study of Law, Science, and Technology Conference: Forensic Science for the 21st Century (Apr. 3, 2009).

B. Forensic Science and the Courts Generally

The recent report by a committee of the National Research Council, *Strengthening Forensic Science in the United States: A Path Forward*,⁶⁴ made clear that much of forensic science is weak, pre-science, or non-science, and that the courts have been unable or unwilling to exercise the vigilance that the law requires of them. The report states:

The simple reality is that the interpretation of forensic evidence [in the identification fields, such as handwriting, bitemarks, fingerprints, etc.] is not always based on scientific studies to determine its validity. This is a serious problem. Although research has been done in some disciplines, there is a notable dearth of peer-reviewed, published studies establishing the scientific bases and validity of many forensic methods.⁶⁵

[E]xaggerated expert testimony has sometimes contributed to the admission of erroneous or misleading evidence [which] may have contributed to wrongful convictions of innocent people.⁶⁶

[T]he courts have been utterly ineffective in addressing this problem.⁶⁷

Even when the most vulnerable forensic sciences—hair microscopy, bite marks, and handwriting—are attacked, the courts routinely affirm admissibility citing earlier decisions rather than facts established at a hearing.⁶⁸

Most courts long remained unaware of, or intentionally disregarded what they learned about, the state of the forensic sciences.⁶⁹ These courts, of course, enthusiastically threw open the gates to the proffered expertise.⁷⁰ Awareness of serious weaknesses of some forensic sciences when evaluating them through *Daubert*'s lens does not seem to affect admission or limitation decisions as one might have expected. One federal court evaluating proffered fingerprint identification expertise noted: "I conclude that the one *Daubert* factor which is both pertinent and unsatisfied is the first factor—'testing.'"⁷¹ Still, the testimony was fully admitted.⁷² Another court, evaluating proffered forensic handwriting expertise, concluded: "Were the Court to apply *Daubert* to the proffered FDE [forensic document examiner] testimony, it would have to be excluded.

64. *Supra* note 57 and accompanying text.

65. STRENGTHENING FORENSIC SCIENCE, *supra* note 33, at 8.

66. *Id.* at 4.

67. *Id.* at 53.

68. *Id.* at 107 (quoting Neufeld, *supra* note 55, at S110).

69. *Id.* (stating that "the courts have been utterly ineffective in addressing" the problem of unvalidated and sometimes error-prone forensic science).

70. *See generally* STRENGTHENING FORENSIC SCIENCE, *supra* note 33; FAIGMAN ET AL., *supra* note 5 (discussing historical and contemporary cases in the forensic science chapters).

71. *United States v. Llera-Plaza*, 188 F. Supp. 2d 549, 571 (E.D. Pa. 2002). The court understates the problem. If there has been no testing, that also means there are no studies the quality of which to assess (therefore no "peer review and publication"), and no findings (no "error rates") with which to evaluate the proffered testimony.

72. *Id.*

This conclusion derives from a straightforward analysis of the suggested *Daubert* factors,⁷³ and that the court “might well have concluded that forensic document examination constitutes precisely the sort of junk science that *Daubert* addressed.”⁷⁴ Nevertheless, the court admitted the expert testimony.⁷⁵ The court’s reasoning was that *Daubert* applied only to science; since the hearing disclosed that asserted handwriting identification expertise had no scientific basis, *Daubert*, therefore, did not apply and the evidence was admissible.⁷⁶ Another court, evaluating proffered firearms identification expertise, observed that the examiner

conceded, over and over again, that he relied mainly on his subjective judgment. There were no reference materials of any specificity, no national or even local database on which he relied. And although he relied on his past experience with these weapons, he had no notes or pictures memorializing his past observations.⁷⁷

The court nevertheless (though reluctantly) admitted the testimony.⁷⁸

Whether all of this will change in the wake of the NRC Report remains to be seen. But, undoubtedly, there is much inertia to be overcome. The following sections take a closer look at some interactions of the courts and several forensic sciences.

C. Voiceprint Identification

In the United States, the proffering of “voiceprint” experts has been in a slow decline since a National Academy of Sciences (NAS) scientific review in 1979 indicated great skepticism that the technique could do what it purported to do: dependably link a recorded voice to the person who uttered the words.⁷⁹ Upon the report’s publication, the FBI ceased to offer such expert testimony as part of a prosecution’s case in chief. Local police agencies continued to offer such testimony. When challenged, it was no less likely to be admitted after 1979 than before.⁸⁰ By the end of the century, the proffering of asserted voiceprint experts had nearly ceased, but not because of any recognition by courts that they were not able to do what they asserted.⁸¹

Thirty years after Massachusetts provided the very first appellate affirmance in the United States of the admission of voice spectrographic expert testimony, a Massachusetts Superior Court revisited the same question in the same case, on a motion for post-conviction relief.⁸² To the court in 2005, the 1979 NAS Report was offered as new

73. *United States v. Starzeczyzel*, 880 F. Supp. 1027, 1036 (S.D.N.Y. 1995).

74. *Id.* at 1028.

75. *Id.*

76. *Id.*

77. *United States v. Green*, 405 F. Supp. 2d 104, 107 (D. Mass. 2005).

78. *Id.* at 108.

79. NAT’L RESEARCH COUNCIL, COMM. ON THE EVALUATION OF SOUND SPECTROGRAMS, ON THE THEORY AND PRACTICE OF VOICE IDENTIFICATION (1979).

80. *See generally* 5 FAIGMAN ET AL., *supra* note 5, § 37:1 (discussing patterns of admit/exclude decisions in talker identification cases).

81. *Id.* §§ 37:2–37:3.

82. *Commonwealth v. Lykus*, No. 43558, 2005 WL 3804726 (Mass. Super. Ct. 2005), *rev’d*, 885 N.E. 2d 769 (Mass. 2008). By then, a considerable amount of research existed showing the inadequacies of voiceprint identification methods. *See* review in FAIGMAN ET AL., *supra* note 5, ch. 37.

evidence—evidence of lack of validity of the science, unavailable at the time of the trial—and the court ordered a new trial. The Supreme Judicial Court, however, did not regard the 1979 NAS Report on voiceprints as casting doubt on the original expert testimony, “because [the NAS Report] does not repudiate the use of voice identification testimony.”⁸³ The court quoted from the Report:

The Committee concludes that the technical uncertainties concerning the present practice of voice identification are so great as to require that forensic applications be approached with great caution. The Committee takes no position for or against the forensic use of the aural-visual method of voice identification [the visual analysis of voice spectrograms], but recommends that if it is used in testimony, then the limitations of the method should be clearly and thoroughly explained to the fact finder, whether judge or jury.⁸⁴

Although these conclusions were cautiously offered, the FBI understood the report as a whole to be damning enough to lead it to cease offering voiceprint analysis, and a long, slow slide toward the disappearance of voiceprint expert testimony in the United States began. Moreover, if the proponent bears the burden of establishing validity (*Daubert*) or general acceptance in the scientific community (*Frye*), the NAS voiceprint report would seem to make the road to meeting the proponent’s burden exceedingly steep.

Voiceprint expert testimony was the expertise at issue in the very case by which the Alaska Supreme Court adopted *Daubert* as Alaska’s test of admissibility in 1999.⁸⁵ Before deciding on the applicable legal test, or applying that test to voiceprints, the high court remanded the case for the taking of further evidence which would be relevant to evaluation under *Daubert* as well as *Frye*.⁸⁶ Though the remand was made specifically to obtain *Daubert*-relevant evidence, the record that returned to the Alaska Supreme Court was devoid of any relevant studies or other data regarding the supposed scientific expertise being assessed.⁸⁷ The Alaska Supreme Court nevertheless announced its adoption of *Daubert*, reviewed the still-empty record under the new standard, and found it sufficient to affirm the lower court’s admission of the expert testimony under *Daubert*’s principles.⁸⁸

Remarkably then, despite adopting *Daubert* and its central requirement that “[p]roposed testimony must be supported by appropriate validation,”⁸⁹ no Alaska court from the bottom to the top had seen, or required the proponent to produce, any evidence on the question at hand. Voiceprint identification expert evidence was admitted almost without thought, even in Alaska’s inaugural *Daubert* decision, even though much more than the usual amount of research data existed on this forensic science, and even though voiceprint evidence elsewhere in the United States had fallen almost entirely out of use.

83. *Lykus*, 2005 WL 3804726, at *15.

84. *Id.* at *12.

85. *State v. Coon*, 974 P.2d 386, 394–96 (Alaska 1999).

86. *Id.* at 389.

87. *Id.* at 400–06.

88. *Id.* at 402 (adopting *Daubert* as Alaska’s interpretation of its Rule 702, and affirming the admission of voice spectrographic evidence under that test).

89. *Daubert v. Merrell Dow Pharms.*, 509 U.S. 579, 590 (1993).

D. Fire and Arson Investigation

Whether fires are deliberately set or accidental had long been a matter on which experts opined based on speculations concerning various arson “indicators” that had been collectively agreed upon by the fire investigation community, using reason and logic alone. Eventually these “indicators” were empirically tested by burning down structures to simulate set and accidental fires, and examining whether the “indicators” correlated as they should have. Nearly two dozen of the indicators were found to be invalid—that is, incapable of distinguishing set fires from accidental.⁹⁰ Before these discoveries were made, courts did not adequately scrutinize the proffered testimony, even under *Daubert*, and after the discoveries were made many experts remained ignorant of them.

In some cases, courts place heavy reliance on the proffered fire investigation expert’s “qualifications”—“experience” doing the task and sometimes the taking of brief courses on the subject—seeming to miss the central aspect of vetting expert testimony with which the *Daubert* line of cases is concerned, namely, the reliability (validity) of the content of the proffered testimony.⁹¹ These courts treat the doing of a job for some period of time or conducting some number of investigations as an assurance of validity, even if the witness has disregarded the best established scientific findings or relied on discredited older speculations of the field.⁹² Such courts not only do not scrutinize the content of the proffered testimony for validity, but sometimes allow testimony that the court knows to be contrary to the best knowledge and procedures possessed by the field.⁹³ Some courts have considered challenges to experts who disregarded or deviated from the conclusions and recommendations of NFPA 921,⁹⁴ but the trial court nevertheless admitted (and appellate courts upheld) the proffered expert’s testimony, reasoning that NFPA 921 constitutes a set of guidelines, which an examiner is free to employ or not, relying primarily on the examiner’s judgment and experience.⁹⁵ Unless an examiner can provide a convincing explanation for the disregard of, or departure from, the more scientifically grounded NFPA 921, the approach of these courts seems to undermine the purpose of *Daubert*. It is hard to see how ignoring the science, or discounting it in favor of one’s hunches or outdated beliefs, can assist a factfinder in reaching accurate conclusions.

An interesting aspect of fire and arson expert evidence is that it is used as much in civil litigation as in criminal.⁹⁶ Fire insurers vigorously challenge the causes of the fires that damage or destroy the property of their insured.⁹⁷ Fire investigators are their experts. It would be interesting to know how the quality of the expert testimony, advocacy by counsel, and rulings by judges differ for fire investigators in the civil and criminal arenas. Are there differences? What do they reveal about the legal system, the process, or the judiciary?

90. See generally 5 FAIGMAN ET AL., *supra* note 5, at 59–160 (supporting the unreliability of many “indicators”).

91. See *supra* Parts II.D and II.E (discussing *Daubert* line of cases).

92. See, e.g., *Thompson v. State Farm Fire and Cas. Co.*, 548 F. Supp. 2d 588, 596 (W.D. Tenn. 2008) (finding a witness was qualified largely based on his experience).

93. See FAIGMAN ET AL., *supra* note 5, ch. 38.

94. NAT’L FIRE PROT. ASS’N, NFPA 921, FIRE AND EXPLOSION INVESTIGATIONS (2008).

95. See, e.g., *Thompson*, 548 F. Supp. 2d at 596.

96. See FAIGMAN ET AL., *supra* note 6, ch. 38, pt. I (citing several civil cases).

97. *Id.*

E. Fingerprint Identification

In the area of fingerprint identification, there have been dozens of challenges to admissibility under *Daubert* (and a few under *Frye*), and not a single court has been able to cite any systematic empirical evidence supporting critical propositions underlying fingerprint identification claims.⁹⁸ All but a few of these courts nevertheless found the proffered testimony not only admissible but often worthy of high praise (“the very archetype of reliable expert testimony under [*Daubert*]”).⁹⁹ An extensive review of this body of cases concluded that they amounted to “a catalog of evasions.”¹⁰⁰ That catalog included shifting the burden of proof from the proponent of admission to the opponent; concluding that whatever satisfies British courts is sufficient for American courts (instead of applying American admissibility law); substituting trial processes in place of scientific processes; refusing to consider the challenge at a hearing; reliance on admission by earlier, pre-*Daubert* courts; reliance entirely on general acceptance; reliance on “flexibility” of criteria; bringing the standard down low enough so that the asserted expertise could meet it; and relegating all defects to issues of weight, not admissibility.¹⁰¹ Assuming a sound basis exists for admitting fingerprint expert testimony under *Daubert*, the easier course would have been to write an opinion reciting those findings. That the courts resorted instead to the gymnastics they did suggests they could not find sound bases for admission under *Daubert*, and therefore were compelled either to exclude or find other ways to justify admission.

F. Microscopic Hair Comparison

Of the forensic science errors associated with wrongful convictions, microscopic hair comparison is near the top of the list.¹⁰² Though courts work hard to avoid scrutinizing the proffered scientific evidence, one of the most thoroughgoing efforts to find a justification for admission despite the absence of “appropriate validation” is the Kentucky case of *Johnson v. Commonwealth*.¹⁰³ It illustrates the lengths to which some courts will go to evade the requirements of *Daubert*, and how general acceptance can be used as the magic carpet to soar over *Daubert*’s hurdles. The Kentucky Supreme Court had earlier adopted *Daubert* as that state’s test of admissibility of expert evidence.¹⁰⁴ A few years later, the same court had occasion to consider the question of whether microscopic hair identification expert testimony passed muster under Kentucky’s version of *Daubert*. The courts below had granted and affirmed admission though they had received no scientific evidence whatsoever on the challenged evidence.¹⁰⁵ The Kentucky Supreme Court acknowledged that the record cited no studies, and therefore contained

98. See 4 FAIGMAN ET AL., *supra* note 6, §§ 32:3–32:19 (reviewing the cases).

99. *United States v. Havvard*, 117 F. Supp. 2d 848, 855 (S.D. Ind. 2000).

100. 4 FAIGMAN ET AL., *supra* note 5, at § 32:3.

101. *Id.* §§ 32:3–32:13.

102. See Garrett & Neufeld, *supra* note 62, at 15 (suggesting that “microscopic hair analysis may provide reliable evidence on some characteristics of the individual from which the specimen was taken, but it may not be able to reliably match the specimen with a specific individual”).

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

104. *Mitchell v. Commonwealth*, 908 S.W.2d 100 (Ky. 1995).

105. *Johnson*, 12 S.W.3d at 261.

nothing that could support findings that the technique had been satisfactorily tested, or that methodologically competent studies existed (whether published and peer reviewed or otherwise), and therefore no data existed regarding error rates.¹⁰⁶ Nevertheless, the Kentucky Supreme Court thought the evidence was admissible thanks to the general acceptance factor.¹⁰⁷ The court held that general acceptance was satisfied, even though, as the court acknowledged, no earlier Kentucky case had ever held microscopic hair identification evidence admissible on the basis that it was generally accepted (or because of any other criterion), and made no findings of general acceptance:

Although we have never specifically addressed the scientific reliability of this method of hair analysis, we must assume that it at least satisfied the *Frye* test of general acceptance; for otherwise, the evidence would never have been admitted in the first place. The absence in our previous opinions of any in-depth analysis under the “general acceptance” test was probably due to the overwhelming acceptance of this procedure as a reliable scientific method for the past fifty years.¹⁰⁸

Thus, the mere fact that Kentucky courts had been admitting this kind of expert testimony for years was taken by the Kentucky Supreme Court to imply that they must have evaluated it under general acceptance and found it made the grade, otherwise it would not have been admitted. That is a lot of assuming. Those assumptions are unwarranted, not only because evidence might be admitted without challenge and therefore without occasion for a court to evaluate it, but because, as noted above, it is not true that *Frye*’s general acceptance test had been adopted or even employed in many states.¹⁰⁹ Indeed, a search of cases discloses that the earliest date that any Kentucky court had cited *Frye* for any proposition was 1983, while the precedential cases which the *Johnson* court assumed relied on *Frye* spanned the period 1950 through 1978.¹¹⁰

Thus, the supposedly heavy burden on proponents of expert evidence was met by a series of judicial assumptions drawn out of thin air. The Kentucky Supreme Court ingeniously turned the total absence of supportive evidence or law into everything that was needed to affirm admission of hair comparison evidence. Under a test that conditions admission on scientific proof of the validity of proffered expert evidence, no judge at any level ever looked at a single piece of scientific evidence regarding the expert evidence.¹¹¹

G. Toxicology and Epidemiology

In mass toxic tort cases, parties often seek to offer courts evidence from the fields of toxicology and epidemiology.¹¹² Many courts reject, or find insufficient, toxicological data generated by experiments using animals other than humans.¹¹³ To a roughly equal

106. *Id.*

107. *Id.* at 262.

108. *Id.*

109. *See supra* Part II.C.

110. Based on a Westlaw search by the author.

111. *See Johnson*, 12 S.W.3d at 258 (describing the record before the court).

112. *See FAIGMAN ET AL.*, *supra* note 6, chs. 22 & 23.

113. *Id.* at ch. 22, pt. 5.

and opposite degree, courts seem to welcome—even require—epidemiological data.¹¹⁴ These preferences reflect unschooled judicial intuitions about research methods.

Toxicology experiments using animal models have the advantage of being *experiments*.¹¹⁵ That is, they are randomized, controlled trials. Because of their design, nothing is better at ruling out extraneous influences and permitting strong inferences to be made about the possible effects of the substance in question. That is, they have strong *internal* validity. On the other hand, they have the disadvantage of using animals as research subjects. That is, questions arise about their *external* validity. We can draw sound inferences about causation in the animals, but we cannot be sure if those findings can be generalized to humans.

Epidemiological studies have the opposite strengths and weaknesses. Epidemiological studies have the disadvantage of being correlational.¹¹⁶ They seek out exposed and unexposed people and compare them to try to infer the effects of the substance at issue, but extraneous influences, even when efforts are made to adjust statistically for the influences, clutter the picture and lead to weaker inferences about cause and effect than are available from experiments. On the other hand, the research participants are humans. So, if we could draw sound inferences about cause and effect from these studies (which we cannot quite do), we would have a much easier time generalizing to other people.

Thus, one type of study suffers from weaknesses of internal validity and the other suffers from weaknesses of external validity. Judges tend to see only the obvious problem of external validity (animals versus humans). They have a hard time recognizing, much less appreciating, the problems of internal validity.

That problems of internal validity can be deadly problems is well illustrated by research on hormone replacement therapy (HRT) in post-menopausal women.¹¹⁷ Correlational (epidemiological) research led to the conclusion that HRT was a good idea.¹¹⁸ When women who were receiving estrogen replacement were compared to women who were not, the former group seemed healthier on a whole range of measures.¹¹⁹ But the possibility remained that their superior health resulted from all kinds of things those women did to remain healthy (exercise, diet, adequate rest, other medication, etc.), of which HRT was only one, and that their superior health profiles were in spite of the HRT, not because of it. Later, research of the true experimental kind, in which women were randomly assigned to receive HRT or a placebo, showed that HRT was a disastrously bad idea.¹²⁰ Harvard medical researcher Jerry Avorn concluded that decades of misguided estrogen replacement therapy had caused women tens of thousands of unnecessary breast cancers, heart attacks, and strokes.¹²¹ This tragic mistake resulted from a failure to adequately appreciate the weak internal validity, and therefore the

114. *Id.* at ch. 23, pt. 1.

115. *Id.* at ch. 5 (describing and contrasting experiments).

116. *See id.* chs. 23, 32 (describing strengths and weakness of experimental approaches).

117. JERRY AVORN, POWERFUL MEDICINES: THE BENEFITS, RISKS, AND COSTS OF PRESCRIPTION DRUGS 23–38 (2004).

118. *Id.* at 26.

119. *Id.* at 26–27.

120. *Id.* at 31.

121. *Id.*

tentativeness, of epidemiological research. The actual relationship between hormone replacement and health became painfully clear as soon as research using the better experimental design was completed.

A “fifth *Daubert* factor” has been applied to toxicology and epidemiology. That is whether the research was conducted in the normal course of study of a drug or other substance, or whether it was conducted in the course of or in anticipation of litigation. This “rule” inevitably tends to favor the experts proffered by defendant manufacturers over those enlisted by plaintiffs after suspected harm arises.¹²²

Another curiosity is the judicial sanctification of relative risks (RR) greater than 2.0.¹²³ This has become a powerful dividing line for courts evaluating epidemiological evidence. When the data show RRs equal to or greater than 2.0, the courthouse doors open wide to the studies and the expert witnesses armed with them. But an RR of 2.0 can reflect a tiny correlation and tiny absolute risk. An RR reflects only the numerator of the risk ratio. For example, if in an exposed group of 1,000,000 people, the number who contract the target disease is three, and in an unexposed group of 1,000,000 people, the number with that disease is one, then the RR = 3.0. But if the very same relationship were expressed as an absolute risk, or as a correlation, the value would be minuscule.

Finally, why should studies, or opinions based on them, be admissible only if the studies show differences significant at below some conventional level of probability for research (such as $p < .05$ or $p < .01$) when the ultimate decision in the case is effectively set at a much different level (preponderance of the evidence being akin to $p < .50$)?¹²⁴ Other evidence in a civil case, such as statements of observations by lay or expert witnesses, can be much more doubtful and still be admissible. Courts appear simply to have adopted the conventional standards of scientists, without considering the nature of the decision being made in the legal context. On the other hand, in criminal cases, when relevant but uncertain statements are offered, which fall far below whatever we mean by proof beyond a reasonable doubt, courts are quick to say that that is only one witness or one item of evidence, that a brick is not a wall, and that at the end of the day the factfinders will take it all into account and weigh it accordingly. This seems to be a radically different standard, which cuts sharply in favor of admission, in a context where the standard of proof would seem to move the fulcrum in the direction of greater (not less) caution regarding admissibility.¹²⁵

122. Compare this to the situation when dealing with forensic science. Of the little research that exists in forensic science, whatever of it could overcome the *Daubert* criteria and the “fit” filters would ultimately run aground on this fifth inside/outside litigation factor, because all research in forensic science, almost by definition, is conducted in anticipation of litigation. Moreover, the impact of that fact is felt at every step of the research process, just as the courts fear for toxic tort cases, but simply overlook when it comes to forensic science. D. Michael Risinger & Michael J. Saks, *A House With No Foundation*, 20 ISSUES SCI. & TECH. 35 (2003).

123. For a discussion of relative risk ratios, see FAIGMAN ET AL., *supra* note 6, ch. 23 (especially § 23:27).

124. In statistical hypothesis testing, p-values state the threshold probability of an erroneous rejection of the hypothesis of no effect of the independent variable on the dependent variable (the “null hypothesis”). When the calculated probability is below the p-value, statistical test is said to be “significant,” the null hypothesis is rejected, and the research concludes that the independent variable does have an effect on the dependent variable.

125. Assuming, of course, that decisions by courts on admissibility should be made in contemplation of the standards of proof that will later govern the verdict by the factfinder.

H. Economics Expert Witnesses

In many situations, individuals and businesses can suffer economic loss due to the illegal conduct of others.¹²⁶ Such claims can include breach of contract, fiduciary duty, or good faith dealing; antitrust and unfair or deceptive business practices; fraud and misrepresentation; tortious interference; infringement or theft of intellectual property or trade secrets; trademark infringement, false advertising, or similar deception under the Lanham Act; losses resulting from tortiously caused damage, injury, or death; employment discrimination or wrongful termination; and the subject of this symposium: securities fraud and associated damages. Economics expert evidence can sometimes assist in proving the causal connection between the defendant's conduct and the plaintiff's harm, and usually is informative in determining the amount of the plaintiff's resultant losses.

In the nineteenth century, when economics meant "political economy," numerous judges engaged directly with economists in discussions of economic policy. In the twentieth century, as economics became more technically proficient—and as a consequence developed more useful tools for analyzing causation and damages in specific cases—judges disengaged from their dialogues with economists. This is not as ironic as it might seem. Judges are less likely to understand technical, mathematical tools, and are likely to be less interested in debates that address esoteric economic theory. At the same time, judges seem to have retained some of their historic willingness to critically examine economics expert evidence in a way that judges generally do not do with other technical subject matter.¹²⁷

To be sure, economics expert testimony proffered on appropriate issues is usually admitted without judicial comment, but a review of the cases¹²⁸ reveals that even before *Daubert*, judges were willing to critically evaluate "whether economics expert witnesses [were] making faulty assumptions, [had] an insufficient factual basis, [were] applying faulty analyses, or simply [were] not able to add to the jury's ability to evaluate the damages in a case."¹²⁹ After *Daubert*, judicial scrutiny of economic experts has, if anything, grown more vigorous.

In admitting economics expertise, courts rely not so much on analysis of the underlying scientific theories and the empirical evidence on which those theories stand, but instead on the plausibility and apparent helpfulness of economics experts, and the fact that courts have done so for many years. In a typical case, economics experts begin with some intuitive notion of how to develop a reasonable damages estimate. The expert then gathers available data, fills gaps with explicit and plausible assumptions, and subjects all of this to computations of greater or lesser complexity. The courts make the assumption that, under many or most circumstances, a qualified expert using this approach can provide helpful guidance to a jury. When experts on economic issues are excluded, exclusion usually results from one or more of the following: inadequate input data, faulty assumptions, failure to provide a sound or testable model for analysis, or other methodological shortcomings of the analysis. The change following *Daubert*, if any, has

126. See generally FAIGMAN ET AL., *supra* note 5, ch. 43 (especially §§ 43:1–43:9).

127. *Id.*

128. *Id.*

129. *Id.* (note §§ 43:5–43:8).

been an increased judicial willingness to aggressively review proposed testimony for such flaws, at least in the area of damages.¹³⁰

Expert economics testimony in antitrust is a more complex mixture of economics and legal policy.¹³¹ Because in recent decades judges trying antitrust cases have relied less on per se rules and more on a broad rule of reason, they need economists more than ever—in regard to proof of a relevant market, market power, and likely economic effects, in addition to the calculation of damages. Expert economics witnesses have become all but essential in antitrust.¹³²

This does not mean judges admit such testimony as if free proof were the rule. They do police the experts to try to achieve validity, but they generally do this not by applying *Daubert* and excluding the experts, but rather by aggressive use of summary judgment (finding the evidence admissible but insufficient).¹³³ Antitrust economics is not what one would normally think of as science (propositions subject to, and subjected to, empirical falsification). Many propositions have not been tested, are untestable in the case at bar, and might be untestable, period. To that extent, economics expert witnesses shade into the category of non-science expertise that the Supreme Court dealt with in *Kumho Tire*, and courts have yet to develop good tests for such non-science expertise. However, that problem might be less pronounced in the use of economics experts, partly because judges simply have more experience with economics experts and a history of closer scrutiny of those experts.

IV. CONCLUSION

This Article has sought to provide a view of the broad landscape of judicial filtering of expert witnesses—over time, across rules, and involving different expert subject matters. It has not directly addressed economic experts in securities fraud class actions, which other papers in this symposium are better positioned to do, but rather has attempted to provide a context for thinking more deeply about what can be learned from those other papers.

Above all else, what seems clear is that decisions to admit or exclude are not simple or straightforward applications of an evidence rule used as a guide to filtering a body of purported knowledge offered to answer a particular factual question in a trial. Such decisions never have been. Indeed, in some categories of cases the rule and the body of knowledge appear to be almost irrelevant to the decision whether to admit or exclude. Yet in other categories of cases judges appear capable of thoughtfully scrutinizing the proffered expertise under the applicable rule and reaching informed and intelligent rulings. Our focus is not on the quirks of individual judges, but patterns that reflect the behavior of majorities or factions of judges.

These inconsistencies and contradictions might reflect innocent errors or misunderstandings of either the rule or the subject matter, or they might reflect that courts have agendas other than unmixed faithfulness to reaching the correct ruling on the proffer

130. *Id.*

131. *See generally* FAIGMAN ET AL., *supra* note 5, ch. 44.

132. *Id.* at ch. 43, pt. 1.

133. *Id.*

at bar under the applicable rules.¹³⁴ Our review of the history, doctrine, and cases has suggested numerous factors that might be at play.

We have seen that, at least since *Frye*, courts have realized that the expert is not the expertise—that is, that the validity of an asserted expertise and the qualifications of an asserted expert are separate questions.¹³⁵ Yet sometimes courts mistake the expert for the expertise—in the direction of overlooking questions of underlying validity when whatever is being offered is being offered by a witness with ample experience, decent credentials, and most especially when the witness is a government-employed expert whose job it is to present opinions on that very subject matter.

We also have seen that judges sometimes manipulate admissibility rules, either inadvertently or intentionally, producing different sets of outcomes than would emerge if a rule were consistently employed.¹³⁶ This is most clear in the choice whether to use a broad or narrow version of the *Frye* test. But it also can be seen in the use of *Daubert*, as when courts bend the rule or their application of it in whatever direction is needed in order to let evidence in or keep evidence out.¹³⁷

We have seen that *Frye* and *Daubert* usually should lead to the same result, but there are important and interesting areas of discordance where they should lead to different decisions.¹³⁸ These present more opportunities for evasion and manipulation, as judges in *Frye* jurisdictions import some elements of *Daubert*¹³⁹ and judges in *Daubert* jurisdictions rely overmuch on the general-acceptance prong of *Daubert* when a favored expertise fails to meet any of the other criteria.¹⁴⁰

We have seen that trial judges look for shortcuts—whether it is the marketplace test or *Frye*'s counting-of-noses—to try to learn what the consensus is in a field. Because

134. Sometimes judges are overtly result-oriented in making evidentiary rulings. Recall Judge Van Orsdel, *supra* note 23 and accompanying text, using the general acceptance test in *Frye* but refraining from using it in *Laney*. Here is another example: For some years I taught judges in the summer LL.M. program at the University of Virginia. One day a debate erupted about how to handle the admissibility of a novel type of expert evidence. Several of the approximately 30 students made arguments based clearly on whether they wanted to let the evidence in or keep it out: to let the evidence in, you must use *this* rule; if you use *that* rule you will have to keep the evidence out. These judges were explicitly choosing a rule to accomplish a pre-determined purpose, and saying so in front of their peers from around the country.

135. See *supra* Part II.B (asserting that the body of knowledge and the person testifying on that knowledge are separate and must be evaluated by the court as such).

136. See *supra* Part II.B (differentiating between applications of the *Frye* test when phrased in a broad versus narrow question).

137. Some of the clearest examples come from federal courts' struggles with challenges to the admission of fingerprint expert testimony. Two examples: In *United States v. Havvard*, 117 F. Supp. 2d 848, 849 (S.D. Ind. 2000), *aff'd*, 260 F.3d 597 (7th Cir. 2001), the court began by stating that its "decision may strike some as comparable to a breathless announcement that the sky is blue." Yet the judge could identify no empirical *Daubert* factors which the proffer satisfied, and chose to substitute trial process factors (e.g., that even if there were no scientific tests, the technique had survived decades of courtroom tests) for each of the Supreme Court's factors. Similarly, in *United States v. Cline*, 188 F. Supp. 2d 1287, 1294–95 (D. Kan. 2002), *aff'd*, 349 F.3d 1276 (10th Cir. 2003), unable to find that asserted fingerprint expertise satisfied the major *Daubert* factors, the court concluded instead that the field's shortcomings meant that alternative, less-rigorous criteria needed to be employed in order to facilitate admission. In other words, the hurdle needed to be lowered to allow the proffer to clear it.

138. See *supra* Part II.D (analyzing how the two tests evaluate different attributes of expert knowledge).

139. For example, Florida has been referred to as a *Frye-bert* state.

140. Which, recall, the Supreme Court declared in *Kumho Tire* to be impermissible.

Daubert is more demanding of judges, courts have looked elsewhere for shortcuts, such as following what they have done in the past in the face of challenged proffers, even though the past was under a different rule and even though a new rule, such as *Daubert*, means that “everything old is new again.”¹⁴¹ However inappropriate under the circumstances of a new rule which requires the judge to answer different questions, a shortcut such as looking to other and earlier cases is almost irresistible in a legal system which places so much reliance on past decisions. Shortcuts and inertia rule.

When faced with technical disputes outside of their own knowledge—which is axiomatic when we are talking about judges ruling on expert issues—judges understandably look for something to hold on to. Whether or not those things help to find proper answers to the questions the rule requires judges to ask, artificial support can be found. Conventionality: judges tend to rule in the direction of cultural conventions. If members of society overwhelmingly believe in a proposition, then notwithstanding the hard evidence adduced at a *Daubert* hearing, the court is inclined to find the proposition to be true. Follow the herd: few decision tasks are easier than doing whatever your peers have done, whether or not you (or they) understand the subject matter. Saying “me too” rarely requires much explanation. By contrast, few moves are harder than deciding in a direction contrary to that which many of your colleagues decided. That requires a lot of explanation.

We have sets of rules, or patterns of interpretation or application, that change from civil cases to criminal cases, from forensic science to epidemiology to economics in a contract breach to economics in an anti-trust case. Some differences might be properly attributable to the nature of the expertise or to the precise question the evidence is being proffered to answer. But when the inconsistencies can only be attributed to who the parties are or to the substantive legal issues, then more, and more candid, explanations are needed.

141. See, e.g., *United States v. Hines*, 55 F. Supp. 2d 62, 67 (D. Mass. 1999) (noting that the Supreme Court in *Daubert* and *Kumho Tire* “is plainly inviting a reexamination even of ‘generally accepted’ venerable, technical fields”); *United States v. Horn*, 185 F. Supp. 2d 530, 554 (D. Md. 2002) (“[E]verything old is new again.”).