

1-1-2016

Daubert Debunked: A History of Legal Retrogression and the Need to Reassess Scientific Admissibility

Barbara Billauer Pfeffer
Suffolk University Law School

Follow this and additional works at: <https://dc.suffolk.edu/jtaa-suffolk>



Part of the [Litigation Commons](#)

Recommended Citation

21 Suffolk J. Trial & App. Advoc. 1 (2015-2016)

This Article is brought to you for free and open access by Digital Collections @ Suffolk. It has been accepted for inclusion in Suffolk Journal of Trial and Appellate Advocacy by an authorized editor of Digital Collections @ Suffolk. For more information, please contact dct@suffolk.edu.

***DAUBERT* DEBUNKED: A HISTORY OF LEGAL RETROGRESSION AND THE NEED TO REASSESS “SCIENTIFIC ADMISSIBILITY”**

Barbara Pfeffer Billauer*

ABSTRACT:	1
I. THE COMMON WISDOM OF <i>DAUBERT</i> ’S IMPACT—DEBUNKED.....	3
A. Introduction: The Need for Science and Law to Talk the Same Language	3
B. Background: Daubert and Frye.....	6
C. Misunderstanding The RAND Report and Other Empirical Evidence.....	8
II. THE WHY AND HOW OF <i>FRYE</i> AND <i>DAUBERT</i> : SO WHAT’S THE ISSUE?	10
A. Through the Looking Glass: A Reanalysis of the RAND Study Data	12
B. The Pendulum Swings	17
C. Interpretation: What Does It All Mean?	21
D. From the Phoenix of Frye: Junk Law Emerges	24
E. What’s an Expert—and When Did It Change?.....	27
F. The Flaws of Frye	30
G. The Deficits and Damage Wrought by Daubert	40
H. The Ghost of Frye: Why Isn’t Daubert Being Followed?	42
III. HOW DID WE GET HERE AND HOW DO WE GET OUT?.....	43
A. Can Judges Understand Science?	47
B. Novel Science: The Rip Van Winkle Effect.....	49
C. Scientific Literacy in the Law—Another Approach.....	54
D. How the Mighty Have Fallen	56
E. Conclusion: BING-ing science and the Universal Translator or the Need to Re-Define Science and the Scientific Method in Law	56

ABSTRACT:

With “novel” scientific discoveries accelerating at an unrelenting

* Barbara Pfeffer Billauer is Research Professor of Scientific Statecraft at the Institute for World Politics, Washington D.C. and faculty member of Zefat Academic College of Law, Safed, Israel. She is currently a Ph.D. Candidate, Haifa University School of Law, and is Editor of the forthcoming UNESCO Judicial Book on Bioethics.

pace, the need for accessible and implementable standards for evaluating the legal admissibility of scientific evidence becomes more and more crucial. As science changes, legal standards for evaluating “novel” science must be plastic enough to respond to fast-moving changes. This, ostensibly, was the *Daubert* objective. Since it was decided in 1993, however, *Daubert*’s impact has been hotly contested—with plaintiffs and defendants each claiming the decision unfairly favors the other side. New approaches are constantly suggested to deal with the perceived impact, although there is no uniform consensus of exactly what that impact is. Without accurate and demonstrable evidence of *Daubert*’s effect, however, the current *ad hoc* approach of suggesting new gatekeeping methods is at best blind, and at worst, a partisan attempt to manipulate judicial decision-making.

This article reports on a retrospective meta-review of data regarding the impact of *Daubert* on admissibility of scientific evidence, for the first time (to my knowledge)—via a *simple* statistical perspective. This method confirms other assessments (using different methodologies) which conclude that allowing for short-term corrections post-*Daubert*, there is no difference in outcome whether *Frye* or *Daubert* tests are used. Given that entirely different standards were used in both cases—yet yield similar results—one must acknowledge the possibility that *both* tests are “scientifically” flawed. Further, empirical evidence and academic research agree that judges now substitute their own mechanisms for evaluating scientific evidence, in some cases applying neither *Daubert* nor the pre-existing *Frye* test. The reasons are unclear—perhaps due to an instinctual aversion to the muddled or flawed standards enunciated in both. Nevertheless, it does appear that judges are uncomfortable understanding several *Daubert* tests including the “falsification” criterion, which is the foundation stone of the *Daubert* analysis.

Via reanalysis of the RAND study of 2001 which examined effects of *Daubert*, this research further highlights the scientific/mathematical illiteracy of the current legal community (especially, as I demonstrate, compared to judges pre-*Frye*). This situation presents a major obstacle for devising a workable method for judicial gatekeeping of scientific evidence. Ultimately, however, lawyers and scientists must first agree what “science” is before better methods of evaluating proffered evidence can be devised. It is suggested that lawyers turn to scientists who developed the scientific method as it applies to the science of the courtroom—biology, chemistry and simple physics to enlighten themselves—and substitute these approaches for the flawed and faulty premises advocated by *Daubert* and *Frye*.

*Once, upon decisions bleary
 As one side cheered—the other teary
 Judges pondered, weak and weary,
 And seven robed men¹ succumbed to eerie
 Pseudo-Scientific thee-ery,
 Committing legal Hara-kiri.*

*Spouting Popper's view of science:
 Test by falsification—His is not reliance
 On objective verifi-ance,
 Observations and compliance
 With scientific methods and peer alliance.
 So now the time has come to raise Defiance!
 Quothe the Maven: NEVERMORE!*

I. THE COMMON WISDOM OF DAUBERT'S IMPACT—DEBUNKED

A. Introduction: The Need for Science and Law to Talk the Same Language

The story of *Lorenzo's Oil*² portrays a family searching for a remedy for their son's rare neurological illness. Clashing with "organized medicine," they find an oil which miraculously improves Lorenzo's condition. Controlled scientific studies, however, failed to confirm the oil's effectiveness in treating the disease.

The debate between empirical evidence and case reports rages on, impacting liability determination, allocation of research funding and choice of treatment. Novel biotechnologies bring novel questions, including what constitutes proper medical practice or effective care. Resolution of these matters turns on laboratory data and clinical trials in medicine, consensus of opinion in the bioethical community, and expert testimony in courts of law.

Most law courts turn to formulations set forth in the Supreme Court cases of *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,³ *General Electric Co. v. Joiner*,⁴ and the Federal Rules of Evidence⁵ to assess the propriety of

¹ Justices Rehnquist and Breyer dissented on the key portions.

² *Lorenzo's Oil: The Full Story* (BBC television documentary July 21, 2004), available at <http://news.bbc.co.uk/2/hi/health/3907559.stm>.

³ 509 U.S. 579 (1993).

⁴ 522 U.S. 136 (1997).

⁵ See FED. R. EVID. 702, 104 (outlining rules for testimony by expert witnesses and preliminary questions).

expert testimony. The predicates of their admissibility criteria turn on whether the proffered evidence is based on (1) “reliable” scientific research, (2) is conducted by the “scientific method,” and (3) is predicated on “science.”⁶ These determinations are entrusted to the judge, as “gatekeeper.”⁷ Thus, before a judge can allow a jury to consider proffered “scientific” evidence, he or she must understand what constitutes sound “scientific research.” Even before that determination is made, however, we must all agree on a definition of “science.” Should the legal definition of “science” (and its practice) as detailed in *Daubert* differ from its understanding by scientists, we should expect to find serious scientific errors creeping into legal resolutions—as indeed has occurred.⁸ It is no surprise, then, some studies indicate that gatekeepers have simply substituted their own methods for evaluating evidence, rather than relying on standards set forth in *Daubert*, which to them are incomprehensible.⁹

This problem can be illustrated by comparing a thermometer with a Celsius metric to convey the determination of “freezing point” in a culture tethered to the Fahrenheit system. Both are valid systems—but they cannot simply be transposed. Without a “common denominator” translating the “legal” (i.e., the Daubertian) definition of *science*¹⁰ and the scientific

⁶ See *Joiner*, 522 U.S. at 138 (holding no abuse of discretion when trial court concluded insufficient basis for expert opinion); *Daubert*, 509 U.S. at 590 (“In order to qualify as ‘scientific knowledge,’ an inference or assertion must be derived by the scientific method.”); FED. R. EVID. 702(a)-(b) (stating expert testimony must be “scientific . . . knowledge” based on sufficient facts or data).

⁷ See *Daubert*, 509 U.S. at 597 (discussing gatekeeping role of judge to screen scientific evidence).

⁸ See A. Leah Vickers, *Daubert, Critique and Interpretation: What Empirical Studies Tell Us About the Application of Daubert*, 40 U.S.F. L. REV. 109, 143 (2005) (“For example, scholars have taken issue with some judges’ interpretation of *Daubert* as requiring epidemiological data or specifically epidemiological data that demonstrates a doubling of risk.”); see also *Manko v. United States*, 636 F. Supp. 1419, 1434 (W.D. Mo. 1986) (memorializing this error).

⁹ See Vickers, *supra* note 8, at 143 (suggesting judges may not have sufficient understanding of reliability factors or how to apply them); LLOYD DIXON & BRIAN GILL, RAND INST. FOR CIVIL JUSTICE, CHANGES IN THE STANDARDS FOR ADMITTING EXPERT EVIDENCE IN FEDERAL CIVIL CASES SINCE THE DAUBERT DECISION 40 (2001) [hereinafter RAND STUDY], available at http://www.rand.org/content/dam/rand/pubs/monograph_reports/2005/MR1439.pdf (finding judges increasingly used non-*Daubert* reliability factors over time); see also *Daubert*, 509 U.S. at 600-01 (Rehnquist, J., dissenting) (“I do not doubt that Rule 702 confides to the judge some gatekeeping responsibility . . . 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.”).

¹⁰ See *Daubert*, 509 U.S. at 589-93. The Court in *Daubert* misunderstood “falsifiability,” as used in science, and commingled scientific concepts of “reliability” and “validity” into a hodgepodge-mishmash called “legal reliability.” See *id.* at 590 n.9 (“[T]he difference between accuracy, validity, and reliability may be such that each is distinct from the other by no more than a hen’s kick.”). Few cases clearly differentiate between the two. Judge Joan Gottschall’s evaluation of entrance exams used to hire Chicago firefighters is illustrative of the importance of the distinction. See *Lewis v. City of Chicago*, No. 98 C 5596, 2005 WL 693618, at *1-*12 (N.D.

definition of *science* (which, by the way, differs depending on the field of science involved),¹¹ confusion reigns. Many commentators distance *Daubert*'s impact by focusing on the supersedence of Rule 702 of the Federal Rules of Evidence.¹² Nevertheless, even a superficial reading of the rule indicates that its roots lie in *Daubert* and its shoots reflect the Supreme Court's reasoning.¹³ Thus, even the ambit of the Federal Rules has not curtailed ensuing judicial activism resulting from the confused state of affairs, in a vain attempt at compliance.¹⁴

III. Mar. 2005), *rev'd*, 528 F.3d 488 (7th Cir. 2008), *rev'd*, 560 U.S. 205 (2010). Differentiating between the tests' undisputed statistical reliability and its disputed validity to assure non-discrimination, Judge Gottschall noted that:

[E]ven if the 1995 Test could *reliably* measure what it was supposed to measure, it could not distinguish between those who were qualified for the position of CFD firefighter and those who were not. . . . [thus] it is not clear that the 1995 Test measures what it is supposed to measure . . . [which] undermines the 1995 Test's utility as a *valid* measure of candidates' relative cognitive skills. . . . Although the test scores are known, there is no actual evidence of a correlation between those test scores and job performance [i.e., its validity].

Id. at *1, *9, *10, *12 (emphasis added); *see also* Letter from Joan Gottschall, Judge of the U.S. District Court for the Northern District of Illinois, to author (Oct. 13, 2004) [hereinafter *October Letter from Judge Joan Gottschall*] (on file with author). In her note, Judge Gottschall expressed appreciation for being given an understanding of scientific terms and noted that her new-found understanding between the difference of scientific validity and reliability assisted her in rendering the decision. *See October Letter from Judge Joan Gottschall, supra* ("I'm enclosing a copy of the opinion I was beginning to work on immediately after the seminar. I think you'll see immediately how handicapped I would have been without understanding the concepts of reliability and validity."); *see also* Letter from Judge Gottschall, Judge of the U.S. District Court for the Northern District of Illinois, to author (Aug. 30, 2005) (on file with the author) (expressing similar sentiments).

¹¹ *See* THOMAS KUHN, *THE STRUCTURE OF SCIENTIFIC REVOLUTIONS* 50-51 (4th ed. 2012) ("An investigator who hoped to learn something about what scientists took the atomic theory to be asked a distinguished physicist and an eminent chemist whether a single atom of helium was or was not a molecule. . . . For the chemist the atom of helium was a molecule because it behaved like one with respect to the kinetic theory of gases. For the physicist . . . the helium atom was not a molecule because it displayed no molecular spectrum. Presumably both men were talking about the same particle, but they were viewing it through their own research training and practice. Their experience in problem-solving told them what a molecule must be.").

¹² *See* Michael D. Green & Joseph Sanders, *Admissibility Versus Sufficiency: Controlling the Quality of Expert Witness Testimony in the United States* (Wake Forest Univ. Legal Studies & Univ. of Houston Law Ctr., Paper No. 2016468, 2015), *available at* <http://ssrn.com/abstract=2016468> (arguing courts are moving away from *Daubert*-factor analysis and toward sufficiency analysis, reflected in Rule 702).

¹³ *See* FED. R. EVID. 702(c) ("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 . . . (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.") (emphasis added).

¹⁴ *See* Vickers, *supra* note 8, at 133 (highlighting pattern of judicial behavior regarding evidential admissibility); discussion *infra* Part I.B. (explaining RAND Study data and judges'

The ultimate objective of *Daubert* (and ensuing rules, discussions and reviews), is to provide judges with transparent and easily accessible tests to vet scientific evidence.¹⁵ Before we can refine these tests, a clear assessment of *Daubert*'s effect is warranted. Since *Daubert* is grounded on establishing that the proffered evidence is "scientific," we return to where we started: We must assure that we are all using the same definition of "science."

As I have previously put forth, the basic definition of "science" in the legal world is predicated on misunderstanding certain scientific concepts and inapt use of philosophical views of science.¹⁶ Accordingly, I argue that the (*Daubert*) foundation is flawed. If *Daubert* is indeed influencing judicial decision-making, attempts at rectification must first include a redefinition of "science" such that lawyers, judges and experts are on the same page—which is often not the case.

B. Background: Daubert and Frye

In 1993, the case of *Daubert* pronounced that experts may only testify on evidence considered "scientific."¹⁷ The *Daubert* decision replaced the earlier paradigm enunciated in the 1923 case of *Frye v. United States*¹⁸ which relied on the "general consensus of the scientific community" to dictate admissibility.¹⁹ Both cases were ostensibly triggered by a proffer of "novel" scientific evidence, which the judges felt the then-existing standards of evaluation were insufficient.²⁰

The effects of *Daubert* are hotly contested: Some claim *Daubert* has overly excluded evidence; others claim *Daubert* relaxed standards of

conflicting uses of scientific data in court); see also Sophia I. Gatowski *et al.*, *Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-Daubert World*, 25 LAW & HUM. BEHAV. 433, 443 (2001) (finding majority of judges surveyed overwhelmingly supported the "gatekeeping" role as defined in *Daubert*).

¹⁵ See *Daubert*, 509 U.S. 579, 597 (1993) (explaining purpose of scientific inquiry).

¹⁶ See Barbara P. Billauer, *Admissibility of Scientific Evidence Under Daubert: The Fatal Flaws of Falsifiability and Falsification* (Dec. 10, 2015) [hereinafter Billauer, *Admissibility of Scientific Evidence under Daubert*], available at <http://ssrn.com/abstract=2701737> ("The Daubert mantra demands that judges, acting as gatekeepers, prevent para, pseudo or 'bad' science from infiltrating the courtroom. To do so, the Judges must first determine what 'science' is? And then, what 'good science' is?").

¹⁷ See *Daubert*, 509 U.S. at 597.

¹⁸ 293 F. 1013 (D.C. Cir. 1923).

¹⁹ *Id.* at 1014.

²⁰ See *Daubert*, 509 U.S. at 584 (considering expert testimony based on animal, pharmacological, and "reanalysis" of previously published epidemiological studies); *Frye*, 293 F. at 1014 (introducing type of systolic blood pressure deception test).

admissibility.²¹ Both sides are adamant. This discourse is, of course, predicated on the belief that *Daubert* effectuated a significant change in evidentiary admissibility—either for good or for bad, depending on the bias of the commentator. The debate rages on.²² A proliferation of diametrically opposite opinions claiming that *Daubert* is notoriously anti-plaintiff,²³ or pro-plaintiff—at least in civil cases—have surfaced.²⁴ Recently, the duel began another round, this time accompanied by new solutions to address the apparent imbalances.²⁵ These solutions are attacked by those claiming the legal status quo adequately addresses the situation.²⁶

It is hard to understand how the viewpoints could be so disparate if the empirical evidence is so clear. Hence, before legal commentators take their gloves off for another fencing match, I suggest that it behooves us to re-check the original data: What does the data actually say? What exactly did *Daubert* accomplish? Have legal commentators, relying on verbal conclusions rather than hard data—been misled, and is this entire tempest—illusory?

²¹ Compare Green & Sanders, *supra* note 12 (claiming *Daubert* decision “has led to the exclusion of experts in hundreds, perhaps thousands of cases”), with Cassandra H. Welch, Note, *Flexible Standards, Deferential Review: Daubert’s Legacy of Confusion*, 29 HARV. J.L. & PUB. POL’Y 1085, 1086 (2006) (arguing flexible approach to admissibility has led to confusing standards by which courts evaluate evidence).

²² Compare Jeffry D. Cutler, *Implications of Strict Scrutiny of Scientific Evidence: Does Daubert Deal a Death Blow to Toxic Tort Plaintiffs?*, 10 J. ENVTL. L. & LITIG. 189, 191-92 (1995) (arguing liberalization of standard for admissibility makes recovery less likely for typical toxic tort plaintiff), with Paul M. Barrett, *Justices Rule Against Business in Evidence Case—Restrictive Standard for Use of Scientific Testimony in Trials Is Struck Down*, WALL ST. J., June 29, 1993, at A3 (characterizing *Daubert* as pro-plaintiff decision).

²³ See RAND STUDY, *supra* note 9, at xvi (finding that challenges to expert evidence became increasingly fatal to cases). According to Brian Gill, “[t]he 2001 RAND study found that 90% of the courtroom uses of *Frye* and *Daubert* was anti-plaintiff.” LLOYD DIXON & BRIAN GILL, RAND CORP., RESEARCH BRIEF: CHANGES IN THE STANDARDS FOR ADMITTING EXPERT EVIDENCE (2002) [hereinafter RAND RESEARCH BRIEF], available at http://www.rand.org/pubs/research_briefs/RB9037 (describing 2001 RAND Study).

²⁴ See Vickers, *supra* note 8, at 128-37 (“Despite criticisms that *Daubert* tightened admissibility standards, the decision itself appeared to liberalize the standard by demoting general acceptance to one of five factors to be considered. . . . Nevertheless, the actual effect of the decision has been largely to the contrary. . . . As a result of *Daubert* challenges to admissibility, rates of exclusion increased significantly after 1993 in civil cases.”).

²⁵ See Green & Sanders, *supra* note 12, at 22 (proposing that scientific evidence without expert testimony is insufficient to create reasonable inference of causation).

²⁶ See Aaron D. Twerski & Lior Sapir, *Sufficiency of the Evidence Does Not Meet Daubert Standards: A Critique of the Green-Sanders Proposal*, 23 WIDENER L.J. 641, 641 (2014) (“[W]hat courts have been doing by purporting to follow the dictates of *Daubert* is deciding issues of admissibility with little regard for the *Daubert* criteria, but rather based on the sufficiency of the evidence to infer a causal connection and the harm alleged.”).

C. Misunderstanding The RAND Report and Other Empirical Evidence

Three studies have meticulously gathered and analyzed empirical evidence about the effects of *Daubert* through 1998.²⁷ Other work has discussed these results drawing various apt—and inapt—conclusions.²⁸

One impressive study compared judicial opinions between 1991 and 1998 and reported that “one-third of the judges in 1998 said they admitted expert evidence less often than they did before *Daubert*. . . . Judges in 1998 permitted 59% of cases to proceed to trial without limitation on the evidence compared to a 75% rate taken in an earlier survey.”²⁹ One commentator noted that in the Third Circuit exclusion rates spiked to “70% between mid-1995 and mid-1996, up from 53% for the two years preceding *Daubert*.”³⁰

This data seems to solidly sustain the claim that following *Daubert*, judges both more carefully reviewed scientific evidence and more frequently excluded scientific evidence.³¹ On the surface, one might say

²⁷ See RAND STUDY, *supra* note 9, at 12; Carol Krafka et al., *Judge and Attorney Experiences, Practices and Concerns Regarding Expert Testimony in Federal Civil Trials*, 8 PSYCHOL. PUB. POL’Y & L. 309, 311-12 (2002) [hereinafter *FJC Study*] (describing method of survey wherein federal court judges were polled about experience with expert evidence); Jennifer L. Groscup et al., *The Effects of Daubert on the Admissibility of Expert Testimony in State and Federal Criminal Cases*, 8 PSYCHOL. PUB. POL’Y & L. 339, 342-43 (2002) [hereinafter *Groscup Study*] (studying over 2,000 appellate court cases over span of eleven years).

²⁸ See Christopher R. Grivas & Debra A. Komar, *Kumho, Daubert, and the Nature of Scientific Inquiry: Implications for Forensic Anthropology*, 53 J. OF FORENSIC SCI. 771, 771-76 (2008) (examining impact *Daubert* and other cases had on forensic anthropology); Joe S. Cecil, *Ten Years of Judicial Gatekeeping Under Daubert*, 95 AM. J. OF PUB. HEALTH 574, 575 (2005) (“A recent analysis by the RAND Corporation of a sample of 399 published and unpublished federal district court decisions appearing in the Westlaw database over a 20-year period indicates the extent to which courts have shifted toward excluding proffered scientific and technical evidence.”); Joseph Sanders, Shari S. Diamond & Neil Vidmar, *Legal Perceptions of Science and Expert Knowledge*, 8 PSYCHOL. PUB. POL’Y & L. 139, 141 n.13 (2002) (noting early on, plaintiffs and defendants attempted to spin *Daubert* in their direction, but ultimately “in practice the *Daubert* test has been more restrictive than *Frye*.”); Sophia Gatowski, *supra* note 14, at 443 (discussing her survey results of understanding of *Daubert* criteria by 400 state court trial judges).

²⁹ *FJC Study*, *supra* note 27, at 330; RONALD ROESCH, PATRICIA A. ZAFPT & STEPHEN D. HART, *FORENSIC PSYCHOLOGY AND LAW* 308 (2009) (discussing data of the FJC Study).

³⁰ Vickers, *supra* note 8, at 126-27 (citing RAND STUDY, *supra* note 9, at xvi).

³¹ See sources cited *supra* notes 27-28 (supporting the assertion that expert evidence was excluded more frequently following *Daubert*). Interpreting the FJC study to support the claim that *Daubert* favored defendants falls prey to the infirmities of *Daubert*, i.e., that carefully constructed standards are not set forth for interpretation of data leading to misleading, or even erroneous conclusions. See *FJC Study*, *supra* note 27, at 330. This issue also exemplifies the objectives of *Joiner*—that the conclusion must flow from the methodology (i.e., the study design). See *id.* Thus, the conclusions generated from the FJC results must relate to the study question, and hence we must ask: *What was the FJC study designed to test?* If it was to determine whether there was a change in judicial approach immediately after the time period of 1991-1993, a time when “junk-science” overran and drowned the courts, the study was *properly* designed to

this means *Daubert* favors defendants.³² If this is the case, then the matter is open and shut; *Daubert* did what it was supposed to do—and in so doing favored defendants.

Or perhaps it isn't so simple. Perhaps these studies and their interpretations exemplify exactly what is wrong with the current situation—inadequacy of basic scientific and mathematical education of today's lawyers (and judges) prevents *any* ability to critically assess or understand scientific or mathematical evidence, thereby questioning their ability to present scientific testimony to the courts in any manner capable of being evaluated.³³

As I will show, the conclusion of the RAND data (relied on by those claiming *Daubert* favors defendants), is not supported by the data, and is in fact wrong.³⁴ Dead wrong. Unless, that is, all we want to measure are two snapshots in time—1991-1993 and 1995-1997. But there is a whole lot more going on than those numbers represent. A careful review of the underlying data—at least the data which is presented—demonstrates that we are entirely off track, and that all the hundreds of law reviews, articles and reports devoted to bettering *Daubert*, furthering *Daubert*, and curtailing *Daubert*, I suggest, have been for naught.³⁵

This rigorous review is long overdue, and given the battles of the

determine that fact, and obtain *reliable* results. It showed, compared to the time period immediately before and after *Daubert*, that there was a difference—one favoring the defendants (or if you will, correcting the aberrant all-welcoming attitude prior to *Daubert* where just about anyone professing just about anything was allowed to testify). As to why this is a valid question and an important finding, *see infra* section II.A., discussing the RAND Study data and its misleading conclusions. However, to generalize from the FJC study that as a general premise *Daubert* favors defendants, oversteps the bounds of the study and illustrates the dangers of extrapolating backwards beyond the study limits; a practice sure to generate invalid and erroneous conclusions. By contrast, the RAND study was designed to address precisely that question, i.e., the effect of *Daubert* overall, beginning from a baseline time period of 1980-1989 and not merely relating to post-1991 changes. *See* RAND STUDY, *supra* note 9 at xiii. Notwithstanding the authors' verbal conclusion, the data showed it did not. *See* discussion *infra* section II.A.

³² Thus, whether *Daubert* influenced the judiciary to favor defendants after 1991 does not mean that *Daubert* favors defendants as a general proposition, as will be discussed in great depth in the next section. *See* discussion *infra* section II.A. Nor does it answer the question—for how long after 1991 did *Daubert* influence judges to favor defendants? The assumption that it continued throughout the study (or until the present time) is patently invalid, although it is entertained by virtually all legal commentators.

³³ *See* Margaret B. Kovera & Bradley D. McAuliff, *The Effects of Peer Review and Evidence Quality on Judge Evaluations of Psychological Science: Are Judges Effective Gatekeepers?*, 85 J. APPLIED PSYCHOL. 574, 575 (2000) (“[R]esearch on methodological reasoning skills suggests that judges may not be able to assess scientific validity accurately without additional training.”).

³⁴ *See* discussion *infra* section II.A. (demonstrating flaws in interpreting RAND graph data).

³⁵ *See* discussion *infra* section II.A.-C. (arguing RAND Study interpretations are flawed, and outcomes under *Daubert* and *Frye* are incredibly similar).

legal experts on how to redress the perceived injustices of *Daubert*, appears to be sorely needed. While the legal community has been remiss in this delay, at least we can console ourselves; in the last year at least one book and one significant law review article generated information (discussed in the following sections) that sheds new light on matters which further explains the anomalies between the data and the conclusions previously reached.³⁶

However, there is another imperative for this research: Over two decades have passed since *Daubert* was decided. In these twenty years we have seen an explosion of toxic tort cases including those against manufacturers of Zoloft, Paxil, Viagra, Thalidomide—even the ubiquitous Tylenol, to name but a few.³⁷ The current case management tool of Multi-District Litigation (MDL) helps foster greater uniformity in rulings³⁸—thereby reducing conflicting opinions, and curtailing Supreme Court review. As time goes on, more complex chemicals, pharmaceuticals and bionic implants are being developed—with the potential to cause both good and harm, cure disease and instigate more lawsuits. As an administrative tool, MDL is undoubtedly beneficial. But an unanticipated side-effect of fewer divergent holdings—the predicate for *Daubert* and its progeny (*Joiner* and *Kumho Tire Co. v. Carmichael*)³⁹ rulings in the first place—is that we are less likely to see Supreme Court intervention: In this light, determining whether the *Daubert* paradigm and its reasoning are valid, effective, and partisan, must fall to academia.

II. THE WHY AND HOW OF *FRYE* AND *DAUBERT*: SO WHAT'S THE ISSUE?

The *Daubert* case concerned the admissibility of epidemiological evidence to demonstrate that an increased risk of birth defects was caused

³⁶ See discussion *infra* section II.D. (reviewing Fisher's book, cited *infra* note 109); discussion *infra* section II.F. (analyzing backstory of *Frye*, as detailed in Lepore article, cited *infra* note 45).

³⁷ See Catherine R. Borden et al., *Centripetal Forces: Multidistrict Litigation and Its Parts*, 75 LA. L. REV. 425, 433 n.23 (2014). "These are obviously large proceedings . . . including Asbestos, Diet Drugs, Prempro, Niotal, Welding Fumes, Vioxx, and Seroquel proceedings as well as the following: *In re Phenylpropanolamine (PPA) Products Liability Litigation*, 69 Pa. D. & C.4th 320 (19th, 2003); *In re Silica Products Liability Litigation*, 398 F. Supp. 2d 563 (S.D. Tex. 2005) - MDL No. 926; *In re Ortho Evra Products Liability Litigation*, 422 F. Supp. 2d 1379 (J.P.M.L. 2006); *In re Aredia & Zometa Products Liability Litigation*, MDL No. 1760, 2008 U.S. Dist. LEXIS 43433 (M.D. Tenn., 2008)." *Id.*

³⁸ See *id.* at 432 ("[F]our disposition codes . . . account for 96% of the cases before the [multi district] Panel. In other words, a relatively small number of cases on the Panel's docket are not centralized . . .").

³⁹ 526 U.S. 137 (1999).

by the anti-nausea medication, Bendectin.⁴⁰ By contrast, in *Frye*, the use of systolic blood pressure was proffered as a “novel” measure of veracity of a witness, the expert witness claiming to be the first “to have perfected a device or means of measuring and recording the blood pressure to such an extent that with the aid of this device could detect deception.”⁴¹ The *Frye* court rejected the claim and excluded the evidence, a clear legal (and scientific) no-brainer.⁴²

On its face, the “science” behind the expert’s opinion in *Frye* was sorely lacking. But rather than dissecting the science (as was the practice of previous courts),⁴³ the Court resorted to pablumatic dicta (i.e., “reliance on the consensus in the scientific community”).⁴⁴ Thus, even as *Daubert* is heralded as the “new mantra” of legal admissibility, the ghost of *Frye* still hangs heavy on the courts: *Frye* is still good law in eight jurisdictions at last count (although the number is constantly changing) along with the District of Columbia, for good reason.⁴⁵ It gives Judges the comfort of

⁴⁰ See *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 579 (1993).

⁴¹ *Frye v. United States*, 293 F. 1013, 1013-14 (1923). The issue still plagues the courts. See NAT’L RESEARCH COUNCIL, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 88-89 (2009) (citing commentary to Federal Rules of Evidence 700 and 702). The controversy over lie detector tests is a modern offshoot of *Frye*. *Id.* Other variants of physiological measures claimed to represent changes in psychological states such as measurement of sound wave changes in a person’s voice when alleged to be lying, on the theory that constriction of the muscles of the larynx is a manifestation of “tension” associated with telling untruths, likely will face similar scrutiny. See Jeffrey J. Walczyk et al., *Advancing Lie Detection by Inducing Cognitive Load on Liars: A Review of Relevant Theories and Techniques Guided by Lessons from Polygraph-Based Approaches*, FRONTIERS IN PSYCHOLOGY (Feb. 1, 2013), <http://dx.doi.org/10.3389/fpsyg.2013.00014> (discussing new “truth-telling” measurement approaches prompted by reservations raised in National Research Council report). “The technology of voice analysis has been developed over the last 30 years mainly to improve methods of investigation. . . . [Lie detector glasses] track and analy[ze] the mechanics of vocal patterns to identify emotional states such as excitement, stress, uncertainty and deception. The software can run on variety of computer systems, including laptops for portable use wherever security is a concern.” Mike Hanlon, *Lie Detector Glasses Tell You if Someone is Telling the Truth*, GIZMAG (Jan. 9 2004), <http://www.gizmag.com/go/2530>.

⁴² See *Frye*, 293 F. at 1014 (excluding systolic blood pressure test).

⁴³ See discussion *infra* section II.E. (discussing change in expert qualifications over time).

⁴⁴ See 1 DAVID L. FAIGMAN ET AL., MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY §§ 1-3.0 (2d ed. 2001). *Frye* does not require that judges “have any facility with scientific methods to make the admissibility decision,” but only mandates that they develop “some basis for knowing what [it is that] most scientists believe.” *Id.* at § 1:7.

⁴⁵ See Jill Lepore, *On Evidence: Proving Frye as a Matter of Law, Science, and History*, 124 YALE L.J. 1092, 1140-42 (2015) (articulating influence of *Frye* on rules of evidence); Martin S. Kaufman, *The Status of Daubert in State Courts*, ATL. LEGAL FOUND. 1-3 (2006), <http://www.atlanticlegal.org/daubertreport.pdf> (listing states adopting and/or rejecting *Daubert*). Along with the District of Columbia, states still following *Frye* include: California, Illinois, Maryland, Minnesota, New Jersey (although it is only applied to criminal cases), New York, Pennsylvania, and Washington. See, e.g., *People v. Leahy*, 882 P.2d 321, 327-28 (Cal. 1994) (holding *Kelly/Frye* approach is appropriate and survives *Daubert* in California); *U.S. v. Jenkins*,

being able to rely on the science of the day—that is, science as accepted by the majority of the scientific community—without resorting to independent analysis. By pronouncing the “general acceptance test,” judges are able to reject questionable theories of experts without articulating a reason.

By comparison, *Daubert* did not come without benefit. It unshackled judges from being bound by rigid adherence to the scientific status quo. However, it also imposed burdens—i.e., understanding the principles of science and the scientific method—for which Judges were, by and large, unprepared.⁴⁶ The new paradigm resulted in an overwhelming plethora of law review articles⁴⁷ and, an immediate—but short lived—sea-change in rulings.⁴⁸ For the most part, this was based on the court’s difficulty in distinguishing “the real stuff” from “pseudo-pscience” [sic] or advocatorial misuse of irrelevant science.⁴⁹ Nevertheless, the concept “novel science” as commonly understood in science,⁵⁰ was a red herring in the legal arena, as most issues presented to courts (and certainly the one raised in *Daubert*) are ones with long histories—both in science and law. And it is this excuse that bears examination.

A. Through the Looking Glass: A Reanalysis of the RAND Study Data

In the years following *Daubert*, as mentioned earlier, studies

877 A.2d 1013, 1021 (D.C., 2005) (confirming *Frye* is the applicable standard); In re Commitment of Simons, 821 N.E.2d 1184, 1188 (Ill. 2004) (observing admission of expert testimony is governed by standard expressed in *Frye*); *Burral v. State*, 724 A.2d 65, 80 (Md. 1999) (holding Maryland has not abandoned *Frye/Reed* in favor of standards set forth in FRE 702); *Goeb v. Tharaldson*, 615 N.W.2d 800, 810 (Minn. 2000) (applying *Frye-Mack* standard to determine admissibility of novel scientific evidence); *State v. Harvey*, 699 A.2d 596, 621 (N.J. 1997) (stating *Daubert* has applied to toxic tort cases; test in criminal cases remains general acceptance); *People v. Wesley*, 633 N.E.2d 451, 453-54 (N.Y. 1994) (applying *Frye* test to affirm admissibility of DNA profiling evidence); *Grady v. Frito-Lay, Inc.*, 839 A.2d 1038, 1044 (Pa. 2003) (concluding *Frye* rule will continue to be applied in Pennsylvania); *State v. Copeland*, 922 P.2d 1304, 1314 (Wash. 1996) (holding *Frye* and evidence rules co-exist as law of Washington). *But see* *Smith v. State*, 880 A.2d 288, 304 n.12 (Md. 2005) (“Maryland has not rejected the *Daubert* standard, leaving to case-by-case development whether and to what extent *Daubert* may apply here.”).

⁴⁶ See NAT’L RESEARCH COUNCIL, *supra* note 41, at 90-97 (describing judiciary responses to admissibility of forensic evidence).

⁴⁷ See Twerski, *supra* note 26, at 641. Twerski and Sapir found 3,929 results in the Westlaw database for “*Daubert v. Merrell Dow Pharmaceuticals*” with the search limited to “Secondary Sources, Law Reviews and Journals.” *Id.*

⁴⁸ See *id.* Some claim no sea-change occurred: “A review of case law after *Daubert* shows that the rejection of expert testimony is the exception rather than the rule. *Daubert* did not work a ‘sea change over federal evidence law’ . . .” FED. R. EVID. 702 advisory committee’s note.

⁴⁹ See *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 595 (1993).

⁵⁰ See KUHN, *supra* note 11, *passim* (discussing philosophy of science as it applies to novelties and new discoveries).

evaluated its effects and articles discoursed on their findings.⁵¹ Of the empirical studies, arguably the RAND study was the most data-packed, and according to its authors, “shows . . . judges addressed the issue of reliability more often—and found evidence unreliable more often—after the *Daubert* decision and up through mid-1997. In subsequent years, those same proportions fell.”⁵² The reversal probably came about because plaintiffs and defendants began to tailor the evidence that they proposed, the authors claim, to comport with the *Daubert* requirements.⁵³

Certainly, the claim that following *Daubert*, reliability was both addressed and found unreliable more often than it was pre-*Daubert*, is accurate—but it is horribly misleading.⁵⁴ When viewed in the harsh light of history—over a longer duration rather than as a “snapshot in time,” *Daubert* has made no difference! Looking ten years out on both sides of the “*Daubert* event,” statistically speaking, nothing changed. It is this fact that I will demonstrate here and therefore, ask—

1. On what is the apparent perception of change based—i.e., how could we have gotten things so wrong?
2. What accounted for the radical pendulum shift immediately following *Daubert*?
3. If the two standards (*Daubert* and *Frye*) are so different—how do we explain the similarity (indeed the almost identical nature) of effects over a twenty year period?
4. Are *Daubert* and *Frye* valid and feasible methods to evaluate scientific evidence?
5. Or are both methods so flawed that we should start all over in designing a paradigm for judicial gatekeeping?

The initial data sets in the RAND report were depicted in graphical form with eight data points.⁵⁵ On its face, the first graph indicates an abrupt increase in exclusion of scientific evidence in the four years following the

⁵¹ See sources cited *supra* note 27 and discussion Part I.C. (discussing FJC, RAND, and Groscup studies). The FJC Study reported limiting or excluding expert testimony in 41% of 1998 cases, versus 25% of the cases referenced in the 1991 survey. See *FJC Study*, *supra* note 27, at 322. In contrast to the RAND and FJC Studies, the Groscup Study found no change in rates of admission at the trial and appellate court following *Daubert*. See *Groscup Study*, *supra* note 27, at 363. This finding is said to be the result of differences between the criminal and civil settings, and may be complicated by the fact that investigators only reviewed cases on appeal. See *id.* at 364-65.

⁵² RAND RESEARCH BRIEF, *supra* note 23.

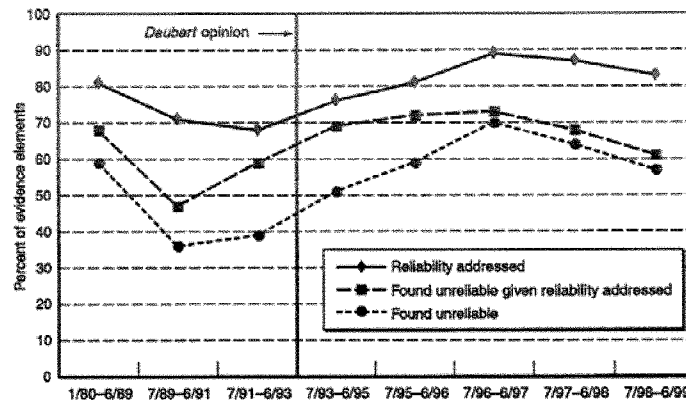
⁵³ See *id.*

⁵⁴ See *FJC Study*, *supra* note 27, at 330 (surveying judges and finding expert evidence admissible less often); RAND STUDY, *supra* note 9, at xvi (noting Third Circuit increase in exclusion after *Daubert*); discussion *infra* (demonstrating flaws in interpreting RAND graph data).

⁵⁵ See *infra* Figures 1, 5. Reprinted with permission from the RAND Corporation.

Daubert decision, before tapering off.⁵⁶

Figure 1. Frequency with which reliability addressed and evidence found unreliable (case type, substantive area of evidence, and appellate circuit held constant).



RAND Figure 4.1. Reproduced with permission from the RAND Corporation. See RAND STUDY, *supra* note 9, at 28.⁵⁷

A careful look at the data points, however, indicates they were broken down into uneven time periods: 1980-89, 1989-91, 1991-93, 1993-95, 1995-96, 1996-97, 1997-98 and 1998-99, combining categorical data (the first four sets) with more continuous data (the last four).⁵⁸ This method of parsing data (in uneven time segments) is unconventional, if not totally unacceptable—and would likely be rejected.

The graph can be recreated to account for this flawed approach—assuming that within the eight discrete time periods there were no changes.⁵⁹ Even with this correction, the RAND method of depicting the data artificially skews the changes in admissibility over the years to look much sharper as can be seen by the recreated graph below.⁶⁰

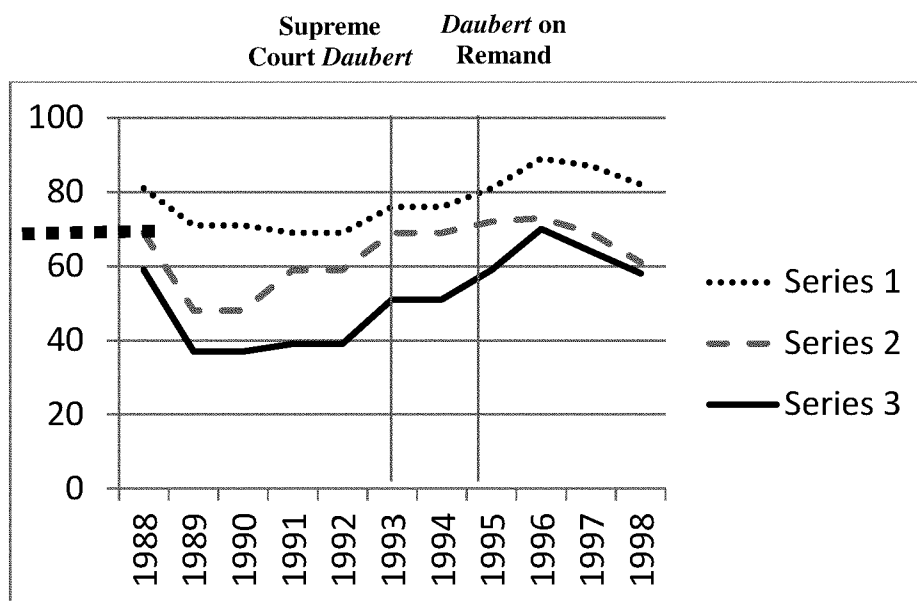
⁵⁶ See *infra* Figure 1.

⁵⁷ It should be noted that RAND Figure 4.1, reprinted above, also appears in the RAND Study as Figure S.1 on page at xv, but with a different caption: “Trends in challenges to reliability of expert evidence (case type, substantive area of evidence, and appellate circuit held constant).” See RAND STUDY, *supra* note 9, at xv, 28.

⁵⁸ See *supra* Figure 1.

⁵⁹ See *infra* Figure 2 (reformulating RAND chart).

⁶⁰ Compare Figure 2 with Figure 1. Categorical data sets obscure (or smooth out) changes over time. Utilizing two methods in one graph gives different weights to the data sets. See *infra* Figure 2.

Figure 2. Reformulated chart using data on a continuous basis.

Series 1 (round dotted line) refers to evidence addressed for reliability; Series 2 (dashed line) refers to evidence found unreliable given reliability assessment; and Series 3 (straight line) refers to evidence found unreliable for any reason, assuming the evidence remained the same, between 1980 and 1989.

Extrapolating backward to 1980 (using the square dotted line), the changes in admissibility within the two decades does not appear as extreme as might be assumed from looking at the RAND Graph, except for the sharp reduction in evidence found unreliable after 1997 (series 2).

When better formatted, the changes in admissibility immediately prior and following *Daubert* do reflect an increased incidence of rejection—but hardly as abrupt as depicted in the RAND graphs.⁶¹ However, when viewed over a twenty year time frame—as opposed to the four years pre- and post-*Daubert*—we see that, comparing the years 1988 with 1998—the percent of evidence excluded actually decreased.⁶² According to the RAND data provided, the results between 1980 through 1988 seemed to remain constant with an increase in evidence admitted increasing sharply after 1988 until shortly before *Daubert*.⁶³ Thus,

⁶¹ Compare Figure 2 with Figure 1 (reflecting more gradual change than RAND data).

⁶² See *supra* Figure 2 (reformulating RAND data).

⁶³ See *supra* Figure 1.

between 1980 and 1988, evidence was excluded approximately 70% of the time, while by 1998, it was excluded in about 60% of instances!⁶⁴

The authors' claim that more evidence was excluded following *Daubert* "because plaintiffs and defendants began to tailor the evidence that they proposed," simply does not hold up.⁶⁵ The data actually demonstrates that in the decade 1980 to 1989, only 30-40% of all evidence was considered reliable.⁶⁶ Over the next few years—until 1993—the percentage of evidence deemed reliable started creeping up.⁶⁷ By 1991, that proportion increased to 52%—and by the time *Daubert* reached the Supreme Court in 1993, 65% of evidence submitted was deemed reliable.⁶⁸ In other words, in the short four-year time frame between 1989-1993, the percentage of evidence considered admissible almost doubled, as can be seen from the eight data points depicted in Figure 1 from the RAND report.⁶⁹ What this means is that up to 1989, more probably than not, scientific evidence would be *excluded*, while as of 1993, more probably than not, scientific evidence would be *admissible*.

What remains unexplained is the sharp rise in the four-year period immediately following *Daubert*, when the pendulum swung and more evidence was excluded—until things returned to equilibrium after 1997.⁷⁰ One possibility for the return to "normalcy" is that the influence of *Daubert* on remand in the 9th Circuit⁷¹ actually evaluated the epidemiological evidence using the approach newly formulated by the Supreme Court and rejected it, thereby calming judicial activism.⁷²

⁶⁴ See *id.*

⁶⁵ See RAND RESEARCH BRIEF, *supra* note 23; Vickers, *supra* note 8, at 127 n.26 ("The [RAND] authors interpret this data to reflect changing behavior by parties in response to the tightening of standards, specifically inferring that 'parties proposing evidence either did not propose or withdrew evidence not meeting the new standards, or better tailored evidence they did propose to fit the new standards.' They also hypothesize that parties challenging evidence may have been so encouraged by past success that they cast a wider net in terms of what evidence they challenged, and hence the proportion of successful challenges may have dropped.").

⁶⁶ See *supra* Figure 2 (reformulating RAND data).

⁶⁷ See *id.*

⁶⁸ See *id.*

⁶⁹ See *supra* Figure 1.

⁷⁰ See *id.*

⁷¹ *Daubert v. Merrell Dow Pharm., Inc.*, 43 F.3d 1311 (9th Cir. 1995) (on remand).

⁷² See Vickers, *supra* note 8, at 126-28 ("The success rate of these challenges increased as well. It appears that judges were scrutinizing reliability more carefully, applying stricter standards and, as a result, excluding an increasing proportion of evidence. . . . [T]he researchers found trends in successful challenges based on relevance and qualifications that paralleled that of reliability, leading them to conclude that judges are scrutinizing testimony more carefully with regard to all criteria. Interestingly, this trend changed in 1997, when the percentage of challenges and their success rate started to decline. Post-1997, the study shows a gradual decrease in the percentage of challenged evidence found unreliable.").

B. The Pendulum Swings

As to the sharp increase following *Daubert*, “this can be easily explained by the results of a recent study by Professors Cheng and Yoon . . . [which] affirm the hypothesis that *Daubert*’s profound impacts have resulted less from *Daubert*’s actual doctrinal test and more from the fact that the decision increased awareness of, and concern about, junk science.”⁷³ In fact, Cheng and Yoon produce empirical evidence supporting my observation and interpretation of the RAND data—that whether one proceeds under *Daubert* or *Frye*—the results are essentially the same.⁷⁴

Most *Daubert* discussion and disagreement concerns admissibility of scientific evidence in products liability and toxic tort cases, although scientific and technical experts are used in a myriad of other lawsuits.⁷⁵ Nevertheless, much of the *Daubert* controversy that finds its way to law journals and the courts concerns toxic torts and products liability, and a look at the exclusion statistics by case type is instructive.

Numerous graphs are presented in the RAND Report relating to subject matter (e.g., health, engineering and technology, and physical sciences), but these are not illuminating when it comes to admissibility frequency. Eventually, at page 29, the Report does furnish the information of interest—but for the first time in the report, the data is not in graph format.⁷⁶ Rather, it is presented in tabular form, condensed into five data points.⁷⁷

This presentation makes it extremely difficult to compare rulings in products liability and toxic tort cases with cases overall, until it is

⁷³ *Id.* at 135 (“Cheng and Yoon conclude that their study lends support to the theory that the actual doctrinal test enumerated in *Daubert* is substantially less important than the awareness the decision generated. They assert: The results of this study are consistent with the theory that the power of the Supreme Court’s *Daubert* decision was not so much in its formal doctrinal test, but rather in its ability to create greater awareness about the problems of junk science . . . [C]ourts apply some generalized level of scrutiny when considering the reliability of scientific evidence, regardless of the governing standard. If accepted, this thesis suggests that debates about the practical merits and drawbacks of adopting a *Frye* versus a *Daubert* standard are largely superfluous.”); see also Edward K. Cheng & Albert H. Yoon, *Does Frye or Daubert Matter? A Study of Scientific Admissibility Standards*, 91 VA. L. REV. 471, 472 (2005) (analyzing *Daubert*’s impact on state courts).

⁷⁴ See Cheng & Yoon, *supra* note 73, at 511.

⁷⁵ See Borden, *supra* note 37, at 431 (“There are almost as many cases in MDL No. 926 (*In re Silicone Gel Breast Implant Products Liability Litigation*) in the database as there are cases from all other non-product liability types of proceedings (air disaster, antitrust, and so on) combined.”).

⁷⁶ See RAND STUDY, *supra* note 9, at 29.

⁷⁷ See *id.* at 29 tbl. 4.2 (reproduced as Figure 3, *infra*).

reformulated in a comparable format.⁷⁸ When this is done, comparison with overall data (presented in Figures 1 and 2) is easier and yields strikingly dissimilar results (see Figure 4).⁷⁹ Immediately following *Daubert*, the percentage of scientific evidence deemed reliable and admissible *increased!*⁸⁰

Figure 3. Proportion of challenged evidence in which reliability addressed and found unreliable, by case type (predicted with substantive area of evidence and geographic region held constant).

Case Type	Opinion Date					
	N	1/80-6/89	7/89-6/93	7/93-6/95	7/95-6/97	7/97-6/99
A. Reliability Addressed						
Product liability and toxic torts	187	86	80	78	90	92
Other BI and PD	93	47	35	61**	64	81**
Business transactions	111	59	16	34	19	65**
Other	210	62	49	58	71**	54
B. Evidence Found Unreliable						
Product liability and toxic torts	187	61	52	25	65	54
Other BI and PD	93	39	51	49	66	74
Business transactions	111	2	2	32**	9	33**
Other	210	31**	6	18*	23**	20**

NOTE: Reference period used to calculate statistical significance of changes is shaded. A single asterisk (*) indicates that the difference from the reference period is statistically significant at 10 percent; a double asterisk (**) indicates that the difference from the reference period is statistically significant at 5 percent.

RAND Table 4.2. Reproduced with permission from the RAND Corporation. See RAND STUDY, supra note 9, at 29.

This trend continued before eventually reaching an admissibility level slightly more than the 1980-89 level.⁸¹ Nevertheless, the bundling of data in the table is troubling and obscures key information. For example, 1991-93, a time period where a large overall change occurred, is not broken out separately—those years are amalgamated with 1989-91, softening out the curve and possibly obscuring important information—such as the alarming increase in “junk” science claimed to be admitted during those years.⁸²

⁷⁸ See *id.* (highlighting results reflecting reliability of evidence in various types of litigation). Statistical significance appears not to have been reached in any of the product liability data, although given the 50% change in the years 93-95, that is difficult to believe. See *id.* However, should these results indicate that statistical significance has not been reached, the study—at least as to products liability—would be unreliable, and hence inadmissible.

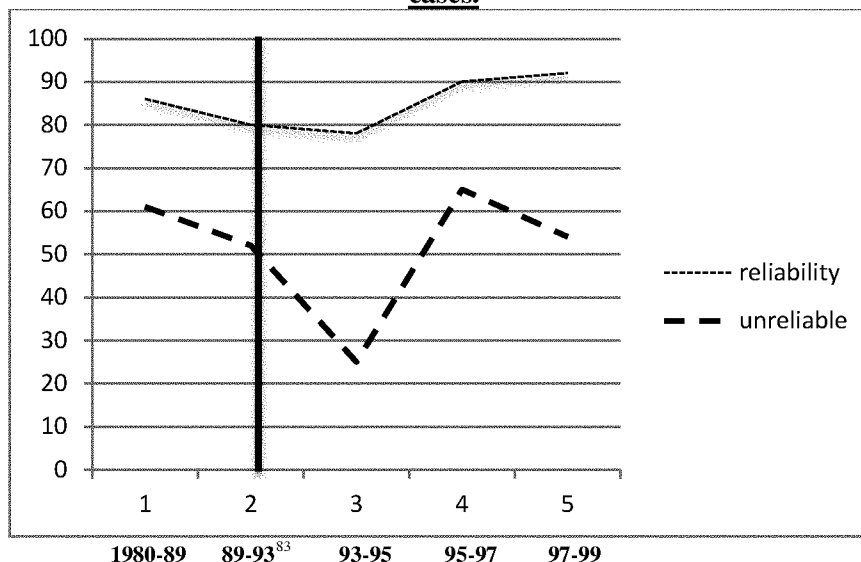
⁷⁹ Compare *supra* Figures 1-2 with *infra* Figure 4. Since the data was bundled into five data points, instead of the eight used in the earlier representations, the correspondence is not exact. See *id.*

⁸⁰ See *infra* Figure 4.

⁸¹ See *supra* Figure 3.

⁸² See *supra* Figure 3 and *infra* section II.D. (discussing emergence of “junk science”)

Figure 4. The *Daubert* decision and the effect on products liability cases.



The straight line represents the percentage of cases where reliability was assessed, and the dash line represents the percentage of cases where evidence was excluded as unreliable.

However, overall the results are still startling: Compared to 1989-93 where 52% of scientific evidence in toxic tort cases was deemed unreliable and excluded, after *Daubert* only 25% of challenged evidence was deemed unreliable—i.e., the percentage of scientific evidence admitted in the five years after *Daubert* in products liability cases nearly doubled before correcting itself.⁸⁴

By 1999, the percentage of challenged evidence deemed unreliable had risen to 54%, a little lower than in the years between 1980-99.⁸⁵ Comparing products liability rulings (above) with all rulings (below) we find more interesting anomalies.⁸⁶ However, it bears repeating that the products liability data is broken down into five categorical data points of uneven durations—smoothing out parts of the curve, making other sections appear steeper than they otherwise might be, and making overall

⁸³ The data by type of claim (Figure 3/RAND Table 4.2) is not categorized with the same years as the earlier and later data.

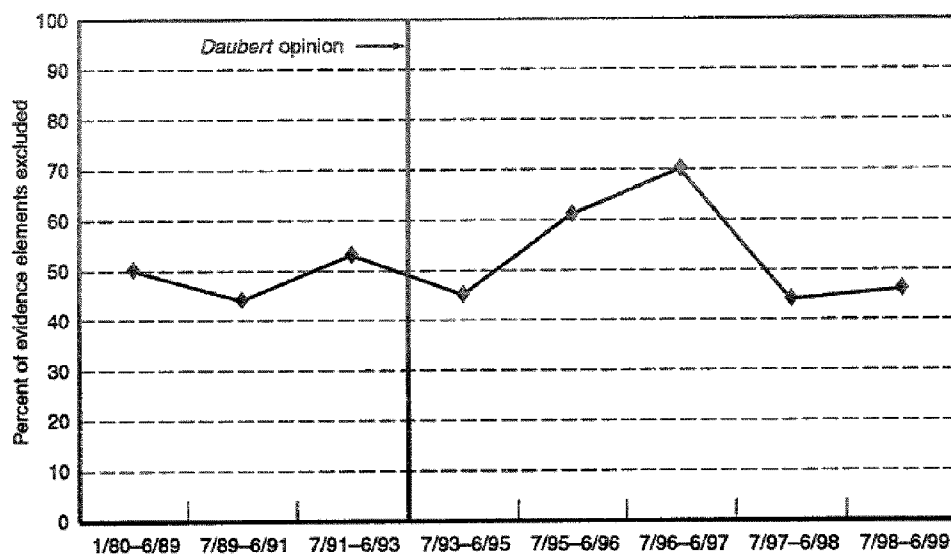
⁸⁴ See *supra* Figure 4.

⁸⁵ See *id.*

⁸⁶ Compare *supra* Figures 3 and 4 with *infra* Figure 5.

comparison of results less precise and reliable, compared to the eight data points of other graphs.⁸⁷

Figure 5. Percentage of evidence elements excluded (case type, substantive area of evidence, and appellate circuit held constant).



RAND Figure 7.1. Reproduced with permission from the RAND Corporation. See RAND STUDY, *supra* note 9, at 55.

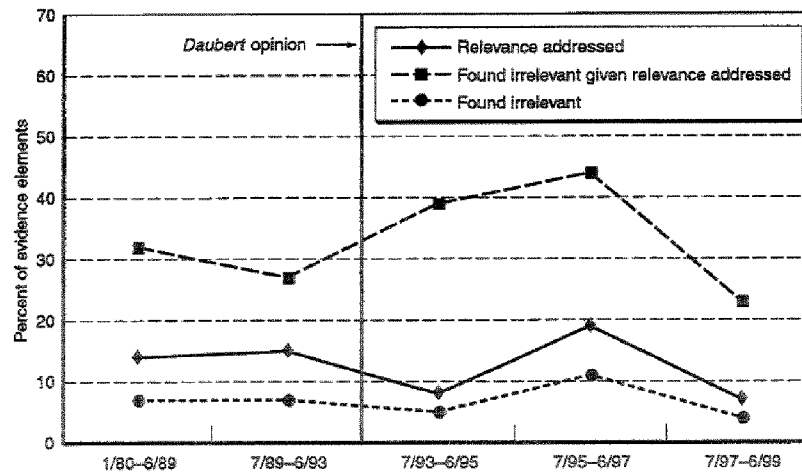
The relevancy statistics, too, are plagued with use of five data points, instead of the eight used to evaluate reliability.⁸⁸ Nevertheless, the overall results appear to be the same: Slightly more evidence was deemed admissible and less was excluded after 1997—that is, after *Daubert*—compared to the years 1980-1989.⁸⁹

⁸⁷ See *supra* Figures 3 and 4 (using five categorical data points of uneven durations).

⁸⁸ Compare *infra* Figure 6 with *supra* Figure 5.

⁸⁹ See *infra* Figure 6.

Figure 6. Frequency with which relevance addressed and evidence found irrelevant (case type, substantive area of evidence, and appellate circuit held constant).



RAND Figure 6.2. Reproduced with permission from the RAND Corporation. See RAND STUDY, *supra* note 9, at 50.

C. Interpretation: What Does It All Mean?

As noted earlier, *Daubert* and *Frye* take radically different approaches regarding admissibility of “novel” science, yet the judicial outcomes, at least over the long term, are incredibly similar.⁹⁰ One explanation might be that during relevant time periods judges were using neither *Daubert* nor *Frye* to make their decisions, a conclusion validated by the RAND study and various commentators.⁹¹ We might consider it a form of “non-differentiated” bias,⁹² a concept borrowed from epidemiology.⁹³

⁹⁰ See discussion *supra* Part I.B. (discussing background of *Daubert* and *Frye*); see also Cheng & Yoon, *supra* note 73, at 471 (reaching same conclusion via different methodology).

⁹¹ See Vickers, *supra* note 8, at 146-47; RAND STUDY, *supra* note 9, at xiii.

⁹² See *Bias: Introduction*, BOSTON UNIV. SCH. OF PUB. HEALTH (Mar. 30, 2015), http://sphweb.bumc.bu.edu/otlt/MPH-Modules/EP/EP713_Bias/index.html. “Bias refers to systematic errors in any type of epidemiologic study that result in an incorrect estimate of the association between exposures and outcomes.” *Id.* “Nondifferential misclassification means that the frequency of errors is approximately the same in the groups being compared. Misclassification of exposure status is more of a problem than misclassification of the outcome. A study may be biased by misclassification of either exposure status, or outcome status, or both.” *Information Bias (Observation Bias)*, BOSTON UNIV. SCH. OF PUB. HEALTH (Mar. 30, 2015), http://sphweb.bumc.bu.edu/otlt/MPH-Modules/EP/EP713_Bias/EP713_Bias4.html.

⁹³ See KENNETH J. ROTHMAN, *EPIDEMIOLOGY: AN INTRODUCTION* 98 (2nd ed. 2002) (In “exposure misclassification, the misclassification is nondifferential if it is unrelated to the

Since *Daubert* is not being followed—at least not uniformly or across the board—obviously, it is useless and renders moot all the legal articles “tweaking” *Daubert* to rectify its ostensible pro-plaintiff or pro-defendant leanings. The finding also brings up three immediate questions:

1. What accounted for the short-term peri-*Daubert* Changes?
2. Why isn’t *Daubert* being followed today?
3. Why wasn’t *Frye* followed before *Daubert*?

To recapitulate, in the short-term post-*Daubert* period, the overall percentage of evidence deemed *inadmissible* increased.⁹⁴ The increase continued for about three to four years, rising sharply until 1997, before plunging again to the levels which existed a decade earlier.⁹⁵ However, for the three-year time frame prior to *Daubert*, the percentage of evidence deemed *admissible* sharply rose.⁹⁶ While more evidence was evaluated for admissibility after *Daubert*, by 1997 roughly the same percentage of evidence was deemed admissible.⁹⁷ The percentage of cases where scientific evidence was deemed admissible between 1980 and 1989 was in the area of 80%.⁹⁸ In fact, comparing the years before 1989 to the period post-1997, the percentage of evidence deemed admissible actually increased by about 10%.⁹⁹

The RAND authors interpret this data in a vacuum, claiming that the post-*Daubert* effects were due to changes in the way evidence was tailored by plaintiffs and defendants to suit the Supreme Court’s new directives.¹⁰⁰ Their interpretation is disputed—and is in fact contradicted by their own findings.¹⁰¹

I (and others) argue that the spike in evidentiary rejection in the years preceding 1997 occurred because *Daubert* sounded a clarion cry that evidence was being too hastily and inappropriately admitted in the period

occurrence or presence of disease; if the misclassification of exposure is different for those with and without disease, it is differential. Similarly, misclassification of disease [outcome] is nondifferential if it is unrelated to the exposure; otherwise, it is differential.”).

⁹⁴ See discussion *supra* Part II.B. and Figures 1-2 (illustrating increase in inadmissible evidence four years following *Daubert*).

⁹⁵ See discussion *supra* Part II.B. and Figures 1-2 (illustrating shift). The percentage of evidence deemed admissible, given reliability assessment actually increased, compared to the 1980-89 time frame. *Id.*

⁹⁶ See *supra* Figures 1-2 (indicating sharp increase in the years between 1990-93).

⁹⁷ See *id.* (returning to 1993 levels by 1997).

⁹⁸ See *id.*

⁹⁹ See *id.*

¹⁰⁰ See RAND RESEARCH BRIEF, *supra* note 23.

¹⁰¹ See Gatowski, *supra* note 14, at 433. Gatowski found that state court judges didn’t even understand many of the *Daubert* tests. *Id.* It appears judges are less than clear on the meaning of falsifiability. *Id.* at 444. In telephone interviews with written follow-up, only 6% of 400 state trial court judges gave answers that indicated a clear understanding of the notion. *Id.*

immediately prior—rather than any implementation of the *Daubert* tests.¹⁰² In effect, then, *Daubert* merely effectuated a short-term course correction to address an anomalous situation—before matters returned to the old status quo.¹⁰³ Vickers claims that rather than the influence of tests enumerated or suggested in *Daubert*, the short-term clamp-down was due to increased cultural awareness of the importance of doing so:

Daubert has had a profound effect on the admissibility of evidence but not via the means that most critics would guess. In fact, *Daubert*'s impacts appear to be the result not of the doctrinal test set forth in the decision, but rather of a cultural phenomenon either sparked by the decision, or to which the decision has contributed.¹⁰⁴

Cheng and Yoon assert that *Daubert*'s influence was not from its doctrinal reform, but from its educative function: “To the extent the decision had a real effect on admissibility, it did so primarily by informing judges that they should function as gatekeepers to ensure that bad science does not make its way into the courtroom.”¹⁰⁵

But, if indeed *Daubert* sounded a sort of clarion call for judges to “heed safety of scientific evidence,” the question becomes—why? Why, suddenly, in 1993 did judges need to be reminded to “keep bad science out of the courtroom”?¹⁰⁶ While Vickers, Cheng and Yoon claim that education or judicial awareness was the key, they do not explain why suddenly keys were needed.¹⁰⁷

¹⁰² See, e.g., Cheng & Yoon, *supra* note 73, at 503 (arguing influence of *Daubert* is not doctrinal, but creating awareness of junk science); Vickers, *supra* note 8, at 144 (observing impact of *Daubert* appears to be cultural shift toward scientific evidence).

¹⁰³ See discussion and Figures 1-2 *supra* Part II.B. (noting sharp rise in excluded evidence immediately after *Daubert* and returning to base-line equilibrium after 1997, assuming the years 1980-89 reflect baseline).

¹⁰⁴ Vickers, *supra* note 8, at 110. “It appears that once judges started acting as more active gate keepers, they more carefully examined relevance, qualifications, and other considerations for admitting evidence, in addition to reliability.” *Id.* at 126.

¹⁰⁵ See *id.* (citing Cheng & Yoon, *supra* note 73, at 503).

¹⁰⁶ As to why post-1997 evidentiary rulings revert to the pre-1989 levels, see *supra* note 72 (discussing trends Pre- and Post-1997). I suggest that after a short period of judicial crackdown (1993-1997) plaintiffs’ attorneys stopped introducing patently improper evidence—a practice that in the years leading up to *Daubert*, was judicially sanctioned, causing more bad science to be proffered, and more to be admitted. See discussion *infra* Part II.D. (discussing emergence of “junk science”). The vicious pre-*Daubert* circle would continue until judges got the message that this practice would not be tolerated by the *Daubert* holding. See *id.* and *infra* note 128 (discussing meeting with Chief Justice Rehnquist and concern about “junk science” in courtroom).

¹⁰⁷ See Cheng & Yoon, *supra* note 73, at 503 (asserting *Daubert* created awareness of junk science); Vickers, *supra* note 8, at 147 (noting cultural shift toward greater skepticism of

One intuitive answer is that in the years immediately prior to *Daubert* an unusual amount of “bad science” was admitted.¹⁰⁸ Thus, only if judges had been admonished not to let that happen—would we see such a drastic change in results immediately following the *Daubert* decision. This hypothesis will be investigated in the following section.

D. From the Phoenix of Frye: Junk Law Emerges

Some years before *Daubert*, American courts began to be plagued with an epidemic of toxic tort cases: Asbestos and Dalkon Shield, to name the most notorious.¹⁰⁹ To be sure, drug companies and manufacturers had demonstrated a history of abuse.¹¹⁰ Partly, what inflamed the asbestos litigation were reckless statements by company managers and reckless conduct of the early manufacturers.¹¹¹ And so, the toxic tort litigation industry “took off.” The public had been sensitized to cavalier pronouncements by the “them that has” and were primed for revenge, and the plaintiff’s bar took advantage of this societal state of mind.

Enter the Silicone Breast Implantation cases.¹¹² While no legitimate medical science ever substantiated any nexus between breast implants and connective tissue disorders, by the early 1990s a juggernaut was about to explode:

- 1) Several million anxious recipients of silicone devices, at least one million of them with breast implants; 2) Six thousand plaintiffs registered in state courts and four thousand in federal courts, each one convinced she was the unwitting victim of a defective product; 3) Virtually every plastic surgeon in the nation co-named in one or more implant suits; 4) Hundreds of plaintiff attorneys positioned

scientific evidence and educational function of decision).

¹⁰⁸ See discussion *infra* II.D. (recounting emergence of “junk science”).

¹⁰⁹ See generally JACK C. FISHER, SILICONE ON TRIAL: BREAST IMPLANTS AND THE POLITICS OF RISK 205 (2015) (recounting controversial history of silicone breast implants); BARBARA P. BILLAUER, *Preparing for Workplace Litigation*, in HANDBOOK OF OCCUPATIONAL SAFETY AND HEALTH 687, 692 (Lawrence Slote 2d. ed. 1998) (mentioning asbestos cases).

¹¹⁰ See BILLAUER, *Preparing for Workplace Litigation*, *supra* note 109, at 692.

¹¹¹ See DONNA BYRNE, MEALEY’S LITIGATION REPORTER, ASBESTOS 982 (1984). On September 12, 1966, E.A. Martin, the Director of Purchases for the Bendix Company (which used asbestos in its products) wrote a letter to one of the principals at Johns-Mansville which was widely circulated and the predicate for many punitive damages awards. *Id.* Martin wrote: “My answer to the problem is: If you have enjoyed a good life working with asbestos products, why not die from it? There has to be some cause.” *Id.*

¹¹² See FISHER, *supra* note 109, at 205 (recounting controversial history of silicone breast implants).

to receive a 30-40 percent share of any award they might extract from a jury or a settlement conference; 5) Four major corporations and soon a fifth facing mass tort peril. . . . [By] [s]eptember 1993 the four major implant manufacturers . . . jointly announced that they had collectively set aside 4.75 *billion* dollars to settle claims filed over the next thirty years.¹¹³

But 1993 also brought good news for science and industry. At least one defense lawyer named David Bernick understood the science and had the savvy to convey it to a jury, and once he got involved the tide turned.¹¹⁴ Even before this, notable researchers had taken notice, and set the stage by creating a maelstrom.¹¹⁵ Peter Huber wrote “Galileo’s Revenge” and coined the term “junk science.”¹¹⁶ Marcia Angell, the editor of *The New England Journal of Medicine* would publish *Science on Trial*, where among other recommendations, she advocated raising the standards for testimony.¹¹⁷ Steve Milloy would write *Junk Science Judo*.¹¹⁸ In 1997 Peter Huber, now joined with Kenneth Foster, again reiterated his position in a new book (avidly supporting the *Daubert* requirements).¹¹⁹ But not until 2015 did Jack Fisher compile the detailed history of the Silicone Breast Implant litigation, laying bare the key factors that brought down the industry, and in the process, cataloging the explosion and proliferation of judicially-approved “junk science.”¹²⁰

¹¹³ *Id.*

¹¹⁴ See *id.* at 210 (discussing Bernick’s success litigating highly complex scientific issues).

¹¹⁵ See PETER W. HUBER, *LIABILITY: THE LEGAL REVOLUTION AND ITS CONSEQUENCES* 220-32 (1988) (arguing for replacement of most tort liability with contractual principles).

¹¹⁶ See PETER W. HUBER, *GALILEO’S REVENGE: JUNK SCIENCE IN THE COURTROOM* 2-3 (1993) (focusing on how “junk science” has invaded courts).

¹¹⁷ See MARCIA ANGELL, *SCIENCE ON TRIAL: THE CLASH OF MEDICAL EVIDENCE AND LAW IN THE BREAST IMPLANT CASE* 111-32 (1997) (warning of widening gap between scientific reality and what is presented in court).

¹¹⁸ See STEVEN MILLOY, *JUNK SCIENCE JUDO: SELF-DEFENSE AGAINST HEALTH SCARES AND SCAMS*, 1 (2001). The Fox News “Junk Science commentator” who runs a web site, junkscience.com, dedicated to “debunking” what Milloy labels “faulty scientific data and analysis.” See JUNK SCIENCE, <http://junkscience.com/about> (last visited Oct. 29, 2015).

¹¹⁹ See KENNETH R. FOSTER & PETER W. HUBER, *JUDGING SCIENCE: SCIENTIFIC KNOWLEDGE AND THE FEDERAL COURTS* 12, 17 (1997). The author points out that there are numerous issues with Foster’s and Huber’s understanding of “forensic scientific evidence,” indicating a lack of appreciation of the nuances of either law, science-education or biology/chemistry, i.e., the sciences of the courtroom. See Billauer, *Admissibility of Scientific Evidence under Daubert*, *supra* note 16.

¹²⁰ See FISHER, *supra* note 109, at 205 (recounting controversial history of silicone breast implants).

Among Fisher's key points is that in the few years leading up to 1993, admission of testimony was based on an attitude of judicial *laissez-faire*; expert testimony from just about anyone the plaintiffs chose to designate as such was admitted—even persons without relevant training, credential or experience.¹²¹

The importance of experts as an influence on juries is summed up by Fisher:

Attorneys for both plaintiff and a defendant understand the influence that any expert witness can exert on the opinions of jurors, whether or not the testimony is legitimate. Juries are easily confused whenever distinctions between authentic and false expertise are cloudy. The typical juror is particularly vulnerable to hindsight bias, meaning a tendency to judge prior incidents according to a known outcome. . . . in cases of toxic exposure the easiest decision is to link the reported effect with the alleged cause, especially when a phony expert is brought forward to promote the notion. All valid scientific evidence to the contrary is more than likely rejected. Only a judge is in a position to control hindsight bias by ensuring the highest standard of evidence offered at trial.¹²²

Following the Supreme Court's decision in *Daubert*, Judge Robert Jones, faced with *Hall v. Baxter Healthcare*,¹²³ became the first judge to appoint a panel of experts to advise the court in breast implant cases.¹²⁴ These included a polymer chemist, an immunologist, a rheumatologist and an epidemiologist.¹²⁵ By this time, there was a vast array of literature to review—and the expert panel concluded that “any theory supporting a claim of alleging linkage between breast implants and any autoimmune disease was at best an untested hypothesis.”¹²⁶ Accordingly, many of the plaintiff's stable of experts were now declared judicial *persona non grata* and their testimony was now ruled inadmissible in some jurisdictions, but

¹²¹ See *id.* at 222.

¹²² *Id.* at 222 (citing Debra L. Worthington et al., *Hindsight Bias, Daubert and the Silicone Breast Implant Litigation: Making the Case for Court-appointed Experts in Complex Medical and Scientific Litigation*, 8 PSYCH. PUB. POL'Y & L. 154 (2000)).

¹²³ 947 F. Supp. 1387 (D. Or. 1996).

¹²⁴ See FISHER, *supra* note 109, at 236-37.

¹²⁵ See *id.*

¹²⁶ *Id.* at 225.

the aggressive practice of using unqualified experts continued.¹²⁷ By 1993 the Supreme Court got the message: Something needed to be done.¹²⁸

E. What's an Expert—and When Did It Change?

Given the importance of vetting expert witnesses, one wonders what the RAND study data shows regarding *Daubert*'s effects on admissibility of scientific experts based on qualifications during the critical time period of the Silicone Breast Implant cases—and specifically, were there any changes in the level of qualifications of experts? The answer to the second question—is nothing. In other words, while both questions were investigated, we are never given a clear answer to the second.

To be sure, the RAND data did provide overall information, noting that following *Daubert*, courts increased consideration of expert qualifications.¹²⁹ Interestingly, the data also showed that the percentage of experts deemed unqualified to testify for any reason remained almost constant from 1980 through 1999, actually *decreasing* a bit from 5% to about 3%!¹³⁰ As to experts deemed *unqualified* to testify based on qualifications—this percentage *decreased* from 20% in 1980 to 13% twenty years later!¹³¹

¹²⁷ See *id.* (describing effects of expert panel on plaintiff's expert witnesses).

¹²⁸ See Personal meeting with William Rehnquist, Chief Justice of the Supreme Court of the United States, and Dr. Norman Bailey, President, Inst. for Global Econ. Growth, Supreme Court Chambers, Washington, D.C. (Mar. 2003). In the wake of *Daubert*, Chief Justice Rehnquist was extremely concerned about junk science in the courts and the general ignorance of science amongst judges, prosecutors and lawyers in general. *Id.* He was supportive of programs designed to instruct the judiciary in the scientific method. *Id.* The purpose of the meeting was to introduce him to a novel method of teaching science to the judiciary, which was the purpose of the Foundation for Law and Science Centers, Inc. (FLASC), of which Barbara Billauer was founder and president, Dr. Norman Bailey was treasurer and vice president, and of which Judge Paul Michel and Professor Leon Lederman were Chairmen at the time. *Id.* The discussion centered on the format of the new method of teaching judges which was devised by the author and which was subsequently patented. See *id.*; Method & Sys. for Providing Interactive Legal Training, U.S. Patent No. 20070048720 (issued Mar. 1, 2007), available at <http://patents.justia.com/patent/20070048720>.

¹²⁹ See RAND STUDY, *supra* note 9, at 50-52 (displaying data graphically in Figures 6.1-4). The data set inconsistency and incongruity manifests here as well. See *id.* Every data presented in graph form has five points; the decade prior to 1989 comprises one set, a four year interval from 1989-1993 is the second data point, and thereafter three two-year interval data points. *Id.* As stated before, this method of categorization weights certain data points more heavily than others, smooths the curves and obscures critical data. See discussion *supra* Part II.B. and note 60.

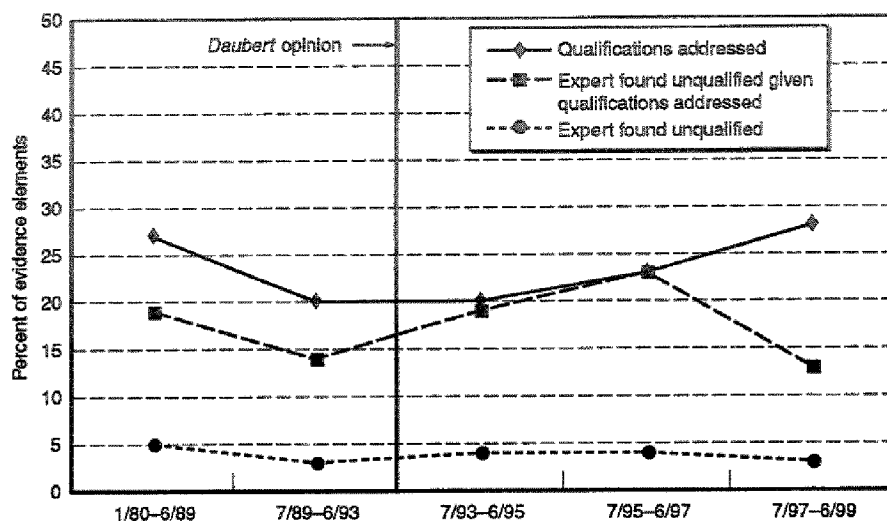
¹³⁰ See *infra* Figure 7 (reproducing Figure 6.3 from RAND STUDY, *supra* note 9, at 51).

¹³¹ See *id.* The use of the phrase “given qualifications addressed” is not clear. See *id.* In typical statistical language it could mean, assuming that all qualifications were the same—and thus qualifications were not the deciding factor; alternatively, it could mean that the data

Of most interest in light of the Silicon Breast Implant cases was whether *Daubert* resulted in changes in level of training and type of expert credentials. Significant by its omission, the RAND study does not tell us.

The nature of expert qualifications—and any change the eight discreet time periods investigated with regard to other questions—is never presented in graphical format, although it is produced in tabular format—sort of.¹³² However, to discern changes that might relate to the Silicon Breast Implant cases, we would need comparisons of segmented time periods, such as 1980-1984, 1985-1989, 1989-1993, 1993-1997, and following 1997 (parenthetically these are also the time periods reflecting the greatest changes when other issues were noted). However, these data were bifurcated into two groups, pre and post 1993, and it is impossible to parse out the critical information.¹³³

Figure 7. Frequency with which expert qualifications addressed and expert found unqualified (case type, substantive area of evidence, and appellate court held constant).



RAND Figure 6.3. Reproduced with permission from the RAND Corporation. See RAND STUDY, *supra* note 9, at 71.

represents disqualification based on qualifications—the exact opposite. See *id.* In fact, given the tabular data produced, it would seem to mean that qualifications were *not* the determining factor represented in the middle line. See *infra* Figure 8 (reproducing Table A.6 from RAND STUDY, *supra* note 9, at 72).

¹³² See *infra* Figure 8 (tabulating qualifications of experts in two groups, 1980-93 and 1993-99).

¹³³ See *id.*

Figure 8. Characteristics of experts (percent of expert-opinion pairs identified in 399 opinions).

Characteristics of Experts (percent of expert-opinion pairs identified in 399 opinions)			
	1/80–6/93 (N=206)	7/93–6/99 (N=363)	Entire Sample Period (N=569)
A. Degree Attained (as described in opinion)^a			
Doctorate	23	27	25
MD	16	16	16
JD or LLB	9	3	5
Masters	3	4	4
Other	5	9	8
Not discussed	50	46	47
B. Gender			
Female	6	6	6
Male	87	89	88
Could not be ascertained from name or pronouns	7	5	6
Total	100	100	100

^aAn expert may have more than one degree, so percentages in this panel sum to more than 100.

RAND Table A.6. Reproduced with permission from the RAND Corporation. See RAND STUDY, supra note 9, at 72.

In sum, the irregularities, lack of consistency or transparency, and confusing use of terminology in the RAND report is most apparent when considering the change in expert credentials—but as to why the sudden shift in the way the data is presented or calibrated, no scientific explanation based on valid and reliable evidence or controlled studies is given, and we are left to our own devices to wonder why.¹³⁴

As to why there was a tapering off of admissible evidence based on reliability and relevancy following 1997, we simply do not know for sure.¹³⁵ All we can tell is that generally speaking, after 1997, at least in toxic torts and products liability cases, the percentage of scientific evidence ruled admissible under *Frye* is roughly the same under *Daubert* as it was under *Frye* before 1989.¹³⁶ And even though the two cases enunciated

¹³⁴ See discussion *supra* Parts I.C., II.A.-C. (analyzing flaws in RAND data). In fact, the recent publication of Fisher's book in 2015 highlights the importance of this omission. See FISHER, *supra* note 109, at 132, 150, 191-92, 194. While the delay in having Fisher's comprehensive and insightful characterization and chronology of events incident to the Silicone Breast Implant Litigation is unfortunate, it appears that in evaluating the impact of scientific evidence, delay in reporting may not be unusual, as is illustrated by Jeanne Guillemin's exquisitely researched investigation on the Anthrax "attacks" of 2001, published ten years later. See generally JEANNE GUILLEMIN, AMERICAN ANTHRAX: FEAR, CRIME, AND THE INVESTIGATION OF THE NATION'S DEADLIEST BIO TERROR ATTACK 1 (2011) (revisiting America's anthrax biohazard scare post-9/11).

¹³⁵ See *supra* Figures 1-7. Whether this could be a reflection of the politics of the federal judges appointed by a Republican Administration is an open question.

¹³⁶ See *id.* (illustrating similar outcomes in admissibility over long term).

vastly different paradigms for judicial decision making, the differences had no effect on outcome.¹³⁷ The obvious question is—how could that be?

F. The Flaws of Frye

The predicate for both cases according to the opinions turns on how to handle “novel” scientific evidence.¹³⁸ Certainly, the *Frye* case put an issue before the judges that had never been raised before, and one that still vexes courts: To what extent are physiological changes representative of a person’s state of mind, emotional, or psychological state—i.e., is there any physiological method to measure when a person is lying?¹³⁹

The scientific validity of lie detector tests is still questioned and questionable, and, as stated earlier, simple reliance on pre-existing tests of evidentiary propriety would have sufficiently validated the court’s ruling.¹⁴⁰ Nevertheless, the appellate court felt compelled to create a new paradigm for admissibility of scientific evidence: The general consensus test.¹⁴¹

To support their conclusion, the court held:

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 the 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

¹³⁷ See Vickers, *supra* note 8, at 140 (“If Cheng & Yoon are correct, it would seem that judges in all jurisdictions have adopted this role and that the *Daubert* decision is largely irrelevant in explaining what is happening in the courtroom.”).

¹³⁸ See *Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923) (discussing novel systolic blood pressure lie detecting machine); *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579 (1993) (concerning novel epidemiological evidence to demonstrate birth defects).

¹³⁹ See generally *Frye*, 293 F. at 1014 (“We think 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.”).

¹⁴⁰ See discussion *supra* Part II. and *infra* Part III.E. (discussing practice of previous courts to assess scientific validity). To even begin to establish its validity would require multiple tests—none of which had not been done at the time when the test results were attempted to be admitted. See generally Walczyk, *supra* note 41 (reviewing theories and techniques of polygraph-based approaches to lie detection). Even today, adequate evidence of validity and reliability of lie-detector tests is wanting. See *id.* Cultural or personality differences can affect interpretation, e.g., pathological liars can recreate a new “truth” and believe in it to the extent that to them it becomes “real” and undifferentiable from a “lie.” See *id.* New devices will likely face similar problems. See *id.*

¹⁴¹ See *Frye*, 293 F. at 1014.

is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.¹⁴²

The *Frye* court's categorization of the issue as vetting "novel" science (i.e., when has the evidence crossed over from being "experimental" to "demonstrable"), is, at least, accurate and honest.¹⁴³ But the then-standard mechanisms of evidentiary admissibility (relevance, materiality and competency) would have worked quite well—as it did for the trial court judge—and garnered the same result, rendering the District Court's dicta superfluous, troublesome, and until now, incomprehensible.¹⁴⁴

Two recent (and exquisitely researched) articles, one in the *Yale Law Journal*,¹⁴⁵ and a slightly earlier piece in the *Journal of the American Academy of Psychiatry and Law*,¹⁴⁶ revealed a slew of hitherto unknown background facts that probably influenced the appellate court to look for as much "cover" as possible in rendering their verdict.¹⁴⁷ Thus, rather than relying on their own judgment, the District Court of Appeals "blamed" the "consensus in the scientific community," for their refusal to admit testimony based on a "novel" scientific device—saddling us, at least until *Daubert*—with this bright line rubric.¹⁴⁸

While the *Frye* holding is well-known, the facts have remained obscure. Indeed, we do not even know exactly what triggered the appellate court's gratuitous and unnecessary holding, notwithstanding some stellar sleuth work in uncovering the doings and goings-on of the trial court. These, however, do shed some light on the situation:

¹⁴² *Id.*

¹⁴³ *See id.*

¹⁴⁴ *See generally* Lepore, *supra* note 45, at 1135 (illustrating trial court Judge's understanding of scientific credibility); Kenneth J. Weiss et al., *Frye's Backstory: A Tale of Murder, a Retracted Confession, and Scientific Hubris*, 42 J. AM. ACAD. PSYCHIATRY L. 226, 226-33 (2014) (discussing trial court's evaluation of expert).

¹⁴⁵ *See* Lepore, *supra* note 45, at 1139 n.263 (discussing claim by expert witness regarding validity of systolic blood pressure in determining deception).

¹⁴⁶ *See* Weiss, *supra* note 144, at 226-33 (discussing facts about defendant and psychologist in *Frye*).

¹⁴⁷ *See* Lepore, *supra* note 45, at 1097 (reconstructing psychologist's role in *Frye*); Weiss, *supra* note 144, at 226-33 (discussing facts behind the *Frye* case); *see also* J. E. Starr, "A Still-Life Watercolor": *Frye v. United States*, 27 J. FORENSIC SCI. 684, 684-94 (1982) (recounting facts of *Frye* litigation); Jim Fisher, *The Polygraph and the Frye Case*, JIM FISHER BLOG (Jan. 7, 2008), <http://jimfisher.edinboro.edu/forensics/frye.html> (detailing correspondence between expert witness and police officer involved in *Frye* case). The author notes the letter exists in the Bancroft Library at the University of California, Berkeley. *See* Fisher, *supra*, at n.6.

¹⁴⁸ *See Frye*, 293 F. at 1014 ("Somewhere in this twilight zone the evidential force of the principle must be recognized . . . 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.").

The *Frye* case concerned the murder of Robert Wade Brown, a well-to-do black doctor and president of the National Life Insurance Company.¹⁴⁹ Dr. Brown was shot dead at point blank range in the front hall of his Washington house before an eye-witness.¹⁵⁰ The murder was both a mystery and a racial incident, carefully and relentlessly covered by the press.¹⁵¹ A year after Dr. Brown's death, James Alphonso Frye was charged with killing Brown and indicted for first-degree murder.¹⁵² All that would save him from certain conviction (and the death penalty) was a compelling alibi—which, with the help of his lawyers, he managed to concoct.¹⁵³ Now, all that remained was to convince the jury of his credibility.

Mr. Frye's lawyers happened to be students of one Dr. William Marston, a lawyer-turned-psychologist, who "claims to have perfected a device or means of measuring and recording the blood pressure to such an extent that with the aid of this device he can detect deception."¹⁵⁴ After administering the primitive lie detector test to Frye in prison, and on the basis of the results of his new—and barely tested—device, Dr. Marston concluded that James Frye was telling the truth—that he did not kill Dr. Brown, and he was prepared to testify to that effect.¹⁵⁵

The trial judge, Walter I. McCoy, was hardly science-phobic.¹⁵⁶ A graduate of Harvard College and Law School, he parsed straight through to the infirmities of Dr. Marston's invention, was not intimidated by Marston's glittering credentials or kowtowed by Marston's association with Dean Wigmore, of evidentiary fame, who had consulted with Marston on his "novel" discovery.¹⁵⁷ In fact, McCoy had no trouble excluding the evidence—without the need for the "*Frye*" test.¹⁵⁸

¹⁴⁹ See Weiss, *supra* note 144, at 226.

¹⁵⁰ See *id.*

¹⁵¹ See *id.*

¹⁵² See *id.* at 227.

¹⁵³ See *id.*

¹⁵⁴ Lepore, *supra* note 45, at 1138.

¹⁵⁵ *Id.* at 1123-27 (recounting actual events leading to Marston taking stand at Frye's trial).

¹⁵⁶ See Walter Irving McCoy, BIOGRAPHICAL DIRECTORY OF THE U.S. CONGRESS, <http://bioguide.congress.gov/scripts/biodisplay.pl?index=M000376>. Prior to Law School, McCoy attended Phillips Exeter Academy and Princeton. *Id.* In addition to a law degree from Harvard, Justice McCoy also earned a Master's degree. *Id.*

¹⁵⁷ See Lepore, *supra* note 45, at 1100-22 (chronicling life of Marston). The proffered expert, Dr. William Marston, was much bally-ho'ed by the forensic-science elite, as operating "the only psycho-legal research laboratory in the United States." *Id.* at 1137. According to the American University Courier, in October 1922, "William Moulton Marston, psychologist and lawyer, has been appointed to the chair of Psychology at the American University . . . and arrangements have been made to open . . . what will probably be the only psycho-legal research laboratory in the United States . . ." *Id.* at 1137 n.250. According to Marston, Wigmore made a practical suggestion on improving the deception test in court: To amplify the blood pressure

As Judge McCoy plunged right in to evaluate the expert and the proffered evidence, he was not impressed:

McCoy had reviewed at least one study with care, the study Marston had published in Wigmore's *Journal of Criminal Law and Criminology*, in which Marston reported the results of deception tests he'd conducted on twenty criminal defendants and, as McCoy saw at a glance, the investigation was wildly unscientific: the cases were handpicked; there was no control group; and the blood pressure test itself might have affected subsequent events.¹⁵⁹

In his repartee with attorney Mattingly, who was representing Frye, Judge McCoy was wittingly eloquent in demonstrating his complete understanding of the nature of science, the scientific method and the competing evidentiary issues at hand:

THE COURT. . . . As far as that test is concerned, Dr. Marston will admit that it was not scientific as far as his instrument was concerned, because, as he understands, as a scientist, he has to exclude everything except the constants before he can make a deduction. If there are a lot of variables, all he can say is that on the whole this is probably so. When it is developed to the perfection of the telephone and the telegraph and wireless and a few other things we will consider it. I shall be dead by that time, probably, and it will bother some other judge, not me.¹⁶⁰

At this point there is colloquy about admitting novel science, to which the judge admits that merely because something is new is not grounds for exclusion (presaging *Daubert*, in a sense):

Mr. MATTINGLY. That is always the way with anything new.

reading so the jurors could see the fluctuations as they were happening. See Weiss, *supra* note 144, at 230. Although it could be done, Marston declined: "I believed that the jury should not be permitted to form their own opinions of a witness's blood pressure record. Interpretation of the record should be made only by experts, psychologists with legal training and experience in lie detecting." *Id.*

¹⁵⁸ See Weiss, *supra* note 144, at 226-33.

¹⁵⁹ Lepore, *supra* note 45, at 1135.

¹⁶⁰ *Id.* at 1132 (quoting Transcript of Record at 14, *Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923) (on file with National Archives, RG 276, Briefs 3986, Box 380)).

The COURT. I suppose it depends upon whether you are before a conservative judge or a young one who is willing to take chances. I have gotten too old and too much inured to certain general principles in regard to the trial of cases to depart from them rashly. Of course anything may happen. It may be that cases will be tried in the absence of defendants with a mere record of whether he is telling the truth about certain things brought in by an expert; I do not know, but so far the jury looks at the witnesses, hears what they have to say, compares their statements with other statements, and so forth, and then does what human beings out of Court do when they determine whether or not a man is telling the truth.¹⁶¹

Mattingly then goes on to discuss some instances where new science was admitted:

The COURT. Well, I will give you this distinction. Fifty years ago if anybody had said that the human voice spoken in Washington could be heard in Chicago he would have been thought crazy. Since that time we all know that such is the fact, and we do not bring experimental matters into court, but when it is established that scientific development has reached such a point as to become a matter of common knowledge as to its results we allow the results to be shown in court.¹⁶²

Judge McCoy raises several bases for excluding Marston's testimony—none based on general acceptance of the *scientific community*.¹⁶³ He talks about controlled studies; he cites the *ceterisparibus* rule of eliminating all other causes but the one under consideration; he talks about regard for “general principles in regard to trial of cases”; he distinguishes between science still in the experimental stages, and science which “has become a matter of common knowledge” (i.e., generally accepted in the *lay community*).¹⁶⁴ Here, he is alluding to the telephone and the fact that the common man is well-acquainted with the device and its

¹⁶¹ *Id.* at 1134 (quoting Transcript of Record, *supra* note 160).

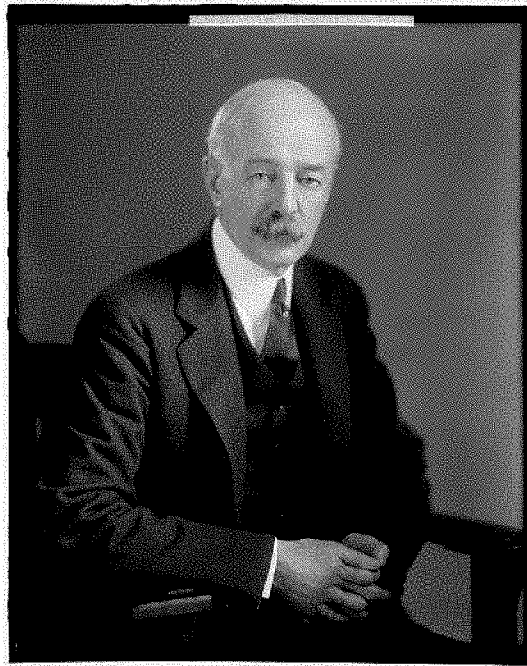
¹⁶² *Id.* at 1129 (quoting Transcript of Record, *supra* note 160).

¹⁶³ *See id.* at 1127-36 (quoting exchange between Judge McCoy and attorney Mattingly challenging Marston's testimony).

¹⁶⁴ *Id.* at 1129 (quoting Transcript of Record, *supra* note 160); Weiss, *supra* note 144, at 228 (discussing McCoy's reasons for excluding Marston's testimony).

reliability—if not the mechanism of how it works.¹⁶⁵ But most important, Judge McCoy recognizes a secondary issue—whether the device is encroaching on the province of the jury.¹⁶⁶ And he clearly notes his refusal to take from the jury their role of discerning credibility of witnesses.¹⁶⁷ The issue, then, McCoy properly notes, is not only one of whether Marston's contraption was reliable—but whether it served a useful purpose which was not equally well (or perhaps better) performed by the jury—the ascertainment of credibility of a witness.¹⁶⁸

Following the opinion, the jury convicted Frye, albeit of the lesser charge of second degree murder, sparing him the death sentence.¹⁶⁹ Some commentators assert that the entire colloquy regarding admissibility of testimony designed to show Frye's innocence was heard in the jury's presence and affected their verdict.¹⁷⁰



Chief Justice Walter I McCoy¹⁷¹

¹⁶⁵ See Lepore, *supra* note 45, at 1132 (quoting Transcript of Record, *supra* note 161).

¹⁶⁶ See *id.* at 1133-34 (reconfirming jury's role in court proceedings).

¹⁶⁷ See *id.* at 1133.

¹⁶⁸ See *id.* (explaining rationale for refusing admission of evidence).

¹⁶⁹ See *Frye v. United States*, 293 F. 1013, 1013 (D.C. Cir. 1923) (restating trial court conviction and basis for appeal); see also Weiss, *supra* note 144, at 227 (chronicling events after Frye's arrest and conviction).

¹⁷⁰ See Weiss, *supra* note 144, at 227-28 (explaining influence lie detector test has on jury).

¹⁷¹ Chief Justice Walter I. McCoy, LIBRARY OF CONGRESS PRINTS & PHOTOGRAPHS DIV.,

The ruling made headlines.¹⁷² Newspaper headlines focused on the “quick and ignominious death” of the deception apparatus: “Court Rules Out Lie-Finding Device,” “Invention Met Its Death on First Trial,” and “Quick Death to ‘Sphygmomanometer,’” for example.¹⁷³ “Immediately after the trial, however, the District of Columbia Court of Appeals agreed to rule on the admissibility of the sphygmomanometer. It appears that the fate of the lie detector sold more newspapers than stories of the underlying crime.”¹⁷⁴

The District of Columbia Court of Appeals ruled on James Frye’s appeal in 1923.¹⁷⁵ “In what some consider a maddeningly terse two-page opinion,” Associate Justice Van Orsdel first articulates the principles behind using systolic blood pressure to monitor truth-telling, before shackling us with the “general consensus in the scientific community test.”¹⁷⁶ What is astounding is that while acknowledging the issue of limiting expert testimony to matters outside the ken of the jury, Van Orsdel did not seize on that to affirm Judge McCoy’s ruling; instead, casting the general consensus hex.¹⁷⁷ The question remains why.

One possibility is the role of Wigmore.¹⁷⁸ The interest of Wigmore in the case is well documented, as was Wigmore’s interest in psychology and its use in the courtroom.¹⁷⁹ For Wigmore, “[n]o science [sic] was more important to the law of evidence, than psychology, and no aspect of psychological research was more important to judicial proof than the study of testimony.”¹⁸⁰ Wigmore was ready to throw precedent to the winds and

<http://www.loc.gov/pictures/resource/hec.20992>.

¹⁷² See Weiss, *supra* note 144, at 228 (describing newspapers’ sensational reaction to court’s treatment of deception detector).

¹⁷³ See *id.* at n.9-13 (listing newspaper headlines).

¹⁷⁴ See *id.* at 228.

¹⁷⁵ *Frye v. United States*, 293 F. 1013, 1013 (D.C. Cir. 1923) (considering single issue of whether exclusion of expert testimony regarding “novel” deception test was proper).

¹⁷⁶ See *Frye*, 293 F. at 1013-14; Weiss, *supra* note 144, at 228 (observing brevity of opinion).

¹⁷⁷ See *Frye*, 293 F. at 1014; Lepore, *supra* note 45, at 1014. “When the question involved does not lie within the range of common experience or common knowledge, but requires special experience or special knowledge, then the opinions of witnesses skilled in that particular science, art, or trade to which the question relates are admissible in evidence.” *Frye*, 293 F. at 1014. Clearly, the issue of witness credibility is within the realm and ken of a jury, and in fact, it is within their exclusive purview, even if the means by which they do their job is not “scientific.” See, e.g., SEYMOUR DWIGHT THOMPSON, 3 A TREATISE ON THE LAW OF TRIALS IN ACTIONS CIVIL AND CRIMINAL 2135 (2d ed. 1912) (“The jury are the exclusive judges of the credibility of the witnesses . . .”).

¹⁷⁸ See Lepore, *supra* note 45, at 1117 (discussing influence of Wigmore).

¹⁷⁹ See *id.*

¹⁸⁰ See *id.* In *Principles of Judicial Proof*, Wigmore’s discussion of testimonial evidence runs more than four hundred pages. See JOHN HENRY WIGMORE, *PRINCIPLES OF JUDICIAL PROOF: AS GIVEN BY LOGIC, PSYCHOLOGY, AND GENERAL EXPERIENCE, AND ILLUSTRATED IN*

abandon conventional “scientific” methods in order to allow new tests to be admitted.¹⁸¹ And Wigmore’s scathing ability to lance into an adversary was not to be taken lightly, especially when it involved psychological testing and evidence.¹⁸² It is known that Marston approached Wigmore to judge his findings, and that Wigmore consented, but it is not known precisely the role Wigmore played in the case or whether his opinions held any sway with the judge or appellate judges.¹⁸³

Probably more significant is the amount of newspaper coverage the ruling of Judge McCoy garnered—for not only was this a case involving scientific evidence, but one involving race relations.¹⁸⁴ Van Orsdel surely suspected his opinion would receive similar—if not greater coverage—and scrutiny.

The *Frye* court’s search for “cover” under these circumstances is understandable, if regrettable, for in creating the “general consensus” paradigm, the District Court exposed their naiveté about the manner of doing science. They also cursed us with an unfortunate legacy; scientists are just as loathe to abandon the status quo as, say, academics. As the physicist Max Planck stated: “A new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar

JUDICIAL TRIALS, 312-743, 585-91 (1913). Wigmore’s remarks, reprinted from his 1909 review of *On the Witness Stand*, are noteworthy. See Lepore, *supra* note 45, at 1105.

¹⁸¹ See Lepore, *supra* note 45, at 1124-25. Interestingly, it has been noted that “If such tests are ever adopted, it is probable that the jury system will have to be abandoned unless education will have advanced so far that twelve men picked at random will adequately absorb blood pressures, time reactions, and intelligence quotients, and combine the mass into a just verdict. In other words, the jury might also be subjected to an intelligence test.” Zechariah Chafee, Jr., *The Progress of the Law, 1919-1921 Evidence*, 35 HARV. L. REV. 302, 309 (1922).

¹⁸² See John H. Wigmore, *Professor Münsterberg and the Psychology of Testimony*, 3 ILL. L. REV. 399, 400-34 (1909) (reporting on early impact of experimental psychology on American judiciary). Wigmore’s attack on Felix Frankfurter regarding his opinion in the Sacco and Vanzetti case is memorialized in a letter from Justice Brandeis to Frankfurter, dated April 27, 1927:

Your telegram reached me at Court. Your letter with the Wigmore blast & your answer did not come to me until late this afternoon. I talked personally to Holmes J., & he is fully prepared for any rabid attack by Wigmore on any one, through W’s attacks on him. W’s attack on the Senate in the April Illinois Law Review & other performances are evidence of an unbalanced mind.

Letter from Louis D. Brandeis to Felix Frankfurter (April 27, 1927), in 5 LETTERS OF LOUIS D. BRANDEIS, 282-83 (Melvin I. Urofsky & David W. Levy eds., 1978).

¹⁸³ See Lepore, *supra* note 45, at 1124-25 (chronicling Marston’s efforts and correspondence with Wigmore).

¹⁸⁴ See *id.* at 1141 (discussing historical context of *Frye*).

with it.”¹⁸⁵ “New scientific ideas never spring from a communal body, however organized, but rather from the head of an individually inspired researcher who struggles with his problems in lonely thought and unites all his thought on one single point which is his whole world for the moment.”¹⁸⁶

Planck’s sentiment has been paraphrased as “science advances, one funeral at a time,” a viewpoint ratified by Thomas Kuhn.¹⁸⁷ In fact, the halls of science are littered with the bodies and even suicides of scientists who took their own lives¹⁸⁸ or died, poor and unrecognized, because their critical discoveries were ignored, trashed, or ridiculed by “the general consensus of the scientific community.”¹⁸⁹ In documenting the history of modern astrophysics and black holes, Marcia Bartusiak notes case after case of scientists who produced scientific work which is now generally accepted, but which was ridiculed, attacked or marginalized when first presented.¹⁹⁰ Bartusiak neatly sums up her thesis as follows: “All truth passes through three stages: First, it is ridiculed; second, it is violently

¹⁸⁵ MAX PLANCK, *SCIENTIFIC AUTOBIOGRAPHY & OTHER PAPERS* 33-34 (Frank Gaynor trans., Philosophical Library 1949).

¹⁸⁶ Max Planck, Address on the 25th Anniversary of the Kaiser-Wilhelm Gesellschaft (Jan. 1936).

¹⁸⁷ KUHN, *supra* note 11, at 5, 7, 159-72 (discussing scientific advancement in human history).

¹⁸⁸ See, e.g., Peter Moore, Op-Ed., *The Great Victorian Weather Wars*, N.Y. TIMES, Aug. 7, 2015, at SR7, available at http://www.nytimes.com/2015/08/09/opinion/sunday/the-great-victorian-climate-debate.html?_r=0 (documenting case of meteorologist Robert Fitzroy, who made great strides and took his life in 1865); Lawrence P. Lessing, *Edwin H. Armstrong: American Inventor*, ENCYCLOPEDIA BRITANNICA, <http://www.britannica.com/biography/Edwin-H-Armstrong> (discussing life and suicide of Edwin Armstrong, discoverer of FM radio) (last visited Dec. 3, 2015).

¹⁸⁹ See RUTH F. HUME, *GREAT MEN OF MEDICINE* 98-202 (1960). Hume discusses the experiences of luminaries such as Edward Jenner, discoverer of the smallpox vaccine, Rene Laennec, discoverer of the stethoscope, and Joseph Lister, who introduced antiseptic surgery, who fall into this category. See *id.* Galileo suffered similarly, as did Copernicus. See MARIO BIAGIOLI, *GALILEO, COURTIER: THE PRACTICE OF SCIENCE IN THE CULTURE OF ABSOLUTISM* 77, 90, 99 (1994) (highlighting Galileo’s struggles); Letter from Cardinal Bellarmine to Foscarini (Apr. 12, 1615), available at http://elearning.gilman.edu/pluginfile.php/58826/mod_resource/content/1/Bellarmino%20Attack%20on%20the%20Copernican%20Theory.pdf (last visited Dec. 3, 2015). For a description of the chauvinistic attitudes of physicists aiming to preserve the status quo (i.e., “the general consensus”), see generally LEE SMOLIN, *THE TROUBLE WITH PHYSICS* 265 (2006) (“In our attempts to make unbiased evaluations of our peers’ work, we professors tend almost reflexively to reward those who agree with us and penalize those who disagree.”). Smolin quotes particle physicist JoAnne Hewett: “I find the arrogance of some string theorists astounding, even by physicists’ standards . . .” and notes that the arrogance Dr. Hewett describes has been prominent in the community of string theorists for some time. See *id.* at 268-69. Smolin also notes the impact of other agendas besides science in arriving at and preserving a consensus. See *id.* at 345.

¹⁹⁰ See MARCIA BARTUSIAK, *BLACK HOLE: HOW AN IDEA ABANDONED BY NEWTONIANS, HATED BY EINSTEIN, AND GAMBLER ON BY HAWKING BECAME LOVED* ix-x (2015).

opposed; and third, it is accepted as being self-evident.’ The concept of the black hole fully experienced each and every phase.”¹⁹¹ An example is the grief given to Stephen Hawking when he first announced results of mathematical experiments that “provided the missing pieces of black hole thermodynamics”:

Needless to say, this idea [at first] did not enthrall his fellow physicists. Relativist Werner Israel says that it “aroused strong opposition almost as soon as it was in print. . . . Skepticism was prolonged and virtually unanimous.” When Hawking first announced his result . . . it was greeted with total disbelief. At the end of the talk the chairman . . . claimed it was all nonsense. “Sorry, Stephen,” he said, “but this is absolute rubbish.”¹⁹²

Subsequently, Hawking was proven right.¹⁹³

In sum, the “general acceptance” test is, scientifically speaking, a futile means of evaluating whether “novel” science is “scientific” or not. Plainly speaking, it is just wrong—although it may be the best we have in some legal circumstances.

Fortunately, the concept of general acceptance in the scientific community appears to have been eschewed by other judges, at least before *Daubert*.¹⁹⁴ Counter-intuitively, this was not the case afterwards: “The RAND Study found that before *Daubert*, ‘general acceptance’ was not commonly used as a factor for admissibility . . . The surprising finding,

¹⁹¹ *Id.* (quoting nineteenth-century philosopher Arthur Schopenhauer).

¹⁹² *Id.* at 169.

¹⁹³ *See id.* Similar experiences happened to Peter Higgs (of the Higg’s boson) as well:

The idea was initially met with suspicion and even ridicule in certain circles. His first paper was rejected by a journal, and some peers accused him and some colleagues of failing to grasp the basic principles of physics. . . . [According to Higgs,] “[n]obody else took what I was doing seriously, so nobody would want to work with me.” . . . Upon publication of his work on the particle in 1964, he and his colleagues were widely dismissed as young pretenders, with some even suggesting they should abandon their research or risk.

Kunal Dutta, *Working on the ‘God Particle’ Saved My Life, Says Peter Higgs*, INDEP. (Feb. 21, 2014); *see also* National Post, *From Ridicule to Rave: Higgs & His Boson*, CANADA.COM (July 5, 2012), <http://www.canada.com/story.html?id=7e84471e-082b-4459-b510-d905f31fe01c> (“When Peter Higgs first proposed that an invisible field strewn across space gave mass to the building blocks of the universe, the theory was ridiculed by some of the most respected minds of the time.”).

¹⁹⁴ *See* Vickers, *supra* note 8, at 128 (explaining that before *Daubert*, “general acceptance” was not commonly used).

however, was that after *Daubert*, a lack of general acceptance was as much a barrier to admission as before, and possibly a greater one.”¹⁹⁵

This bizarre finding is troublesome and a ready explanation is unavailable. However, I would venture that perhaps when faced with proffers of “novel” science which is intuitively “unscientific,” although they cannot articulate why in technical terms, modern judges, like those in *Frye*, resort to the comfort of resting on the laurels of the established scientific community on which to hang their judicial decisions.

G. The Deficits and Damage Wrought by Daubert

This does not, however, explain the rejection of *Frye*’s “general consensus” test by the *Daubert* judges (at a time when the test was apparently not even in judicial vogue).

In *Daubert*, the dilemma was predicated on a clash between epidemiological studies and biological data regarding the pre-natal effects of the anti-nausea drug Bendectin.¹⁹⁶

After reviewing the evidence, the *Daubert* court first held that:

Given the vast body of epidemiological data concerning Bendectin, the court held, expert opinion which is not based on epidemiological evidence is not admissible to establish causation. Thus, the animal cell studies, live animal studies, and chemical structure analyses on which petitioners had relied could not raise by themselves a reasonably disputable jury issue regarding causation. Petitioners’ epidemiological analyses, based as they were on recalculations of data in previously published studies that had found no causal link between the drug and birth defects, were ruled to be inadmissible because they had not been published or subjected to peer review.¹⁹⁷

¹⁹⁵ *Id.* (summarizing RAND Study findings). When it is used, however, “general acceptance was usually sufficient for admissibility. . . . Using regression analysis, the [RAND] authors found that general acceptance pre-*Daubert* was not a good indicator of whether evidence would be found reliable or not, but general acceptance post-*Daubert* was, in fact, a good indicator of reliability. According to the authors of the study, there is ‘[n]o indication it became easier for novel evidence to be admitted.’” *Id.*

¹⁹⁶ See *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 582-584 (1993) (discussing alleged birth defects caused by prenatal ingestion of Bendectin).

¹⁹⁷ *Id.* at 583-84 (citing District Court decision, *Daubert v. Merrell Dow Pharm., Inc.*, 727 F. Supp. 570, 575 (S.D. Cal. 1989)).

But not content to limit themselves to the evidence before them, the *Daubert* court then set out a new paradigm on which to evaluate all scientific evidence,¹⁹⁸ produced additional constructs for judicial consideration and overruled the *Frye* test.¹⁹⁹ Ostensibly, the basis for this new approach was predicated on the quagmire of novel concepts and complicated scientific evidence with which they were faced. The scientific evidence involved, however, was anything but novel (or to use the *Frye* court's terminology, "experimental" as opposed to "demonstrable"). Nor was it the first time the Supreme Court had confronted such evidence—the first introduction of epidemiological evidence was made more than a *century* earlier!²⁰⁰ Thus, it cannot in any way be said that the evidence confronting the court was "novel."²⁰¹

Nevertheless, spitting out what clearly was *obiter dicta*,²⁰² the *Daubert* court introduced a series of new constructs to substitute for the now debunked "general consensus" test.²⁰³ Indicating that the *Frye* test

¹⁹⁸ See *id.* at 589-90 (articulating new standard by which courts will judge scientific evidence).

The subject of an expert's testimony must be "scientific . . . knowledge." The adjective "scientific" implies grounding in the methods and procedures of science. Similarly, the word "knowledge" connotes more than subjective belief or unsupported speculation. The term "applies to any body of known facts or to any body of ideas inferred from such facts or accepted as truths on good grounds." [quoting Webster's Dictionary] . . . But, in order to qualify as "scientific knowledge," an inference or assertion must be derived by the scientific method. Proposed testimony must be supported by appropriate validation—i.e., "good grounds," based on what is known. In short, the requirement that an expert's testimony pertain to "scientific knowledge" establishes a standard of evidentiary reliability.

Id.

¹⁹⁹ See Vickers, *supra* note 8, at 135. It must be reiterated that whether a case is heard under *Frye* or *Daubert*, the outcome would likely be the same. See *id.* This is confirmed by Cheng and Yoon's study, showing "that a state's choice of evidentiary standards does not have a statistically significant effect on removal rates in that state . . ." *Id.* (citing Cheng & Yoon, *supra* note 73, at 503). From this data, Cheng and Yoon "inferred that a state's adoption of *Frye* or *Daubert* makes no difference in practice." *Id.*

²⁰⁰ *United States v. Forty Barrels & Twenty Kegs of Coca-Cola*, 241 U.S. 265, 271 (1916) (considering whether caffeine added to food was deleterious ingredient).

²⁰¹ See *id.* Perhaps what the Supreme Court was referring to was not the novelty of the method—but the interpretation of the results—a scientifically inappropriate use of the concept of "novel science."

²⁰² See *Originalism v. Minimalism*, 36 CATO POL'Y REP. 9, 9 (2014) ("Sunstein proposes that judges should generally avoid broad rules and abstract theories and attempt to focus their attention only on what is necessary to resolve particular disputes.").

²⁰³ See Vickers, *supra* note 8, at 144. "While [Daubert] helped revolutionize how judges across the country in all jurisdictions perceive scientific evidence and their role in assessing it, it also provided a standard with little usable content." *Id.*

was no longer valid,²⁰⁴ although not totally eviscerating the “general consensus test,” the *Daubert* court enunciated that evidence must be grounded in science—which they defined as evidence which is (a) “falsifiable,”²⁰⁵ (b) uses the scientific method, (c) considers error rates, and (d) possibly peer review.²⁰⁶

H. The Ghost of Frye: Why Isn't Daubert Being Followed?

In evaluating the extent to which these (non-exclusive and non-binding) parameters enumerated in *Daubert* were followed, it appears that mostly they were not,²⁰⁷ and that judges’ “bench philosophy of science” seemed to reflect the rhetoric, rather than the substance, of *Daubert*.²⁰⁸ It therefore worth remembering in this context that not all states have incorporated *Daubert*; some still rely on *Frye*, perhaps for the reason that while judges are sympathetic with the philosophy of *Daubert*, they are more comfortable understanding its precepts of *Frye*.²⁰⁹

²⁰⁴ See *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 585 (1993) (“[S]ince 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. . . . Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define.”) (quoting *Frye v. United States*, 293 F. 1013, 1014 (1923)).

²⁰⁵ See Cecil, *supra* note 28, at 575. Cecil commented that “Few [state] judges, for example, were able to define the concept of ‘falsifiability,’ which was one of the factors mentioned in *Daubert*.” *Id.* Gatowski demonstrated that “only 5% of the respondents demonstrating a clear understanding of ‘falsifiability,’ and only 4% demonstrating a clear understanding of error rate. Although there was little consensus about the relative importance of the guidelines, judges attributed more weight to general acceptance as an admissibility criterion.” Gatowski, *supra* note 14, at 444; see also Veronica B. Dahir et al., *Judicial Application of Daubert to Psychological Syndrome and Profile Evidence: A Research Note*, 11 PSYCHOL. PUB. POL’Y & L. 62, 64 (2005) (discussing guidelines set out in opinion and origins of “falsifiability”).

²⁰⁶ See *Daubert*, 509 U.S. at 593-94; see also Gatowski, *supra* note 14, at 444-48 (discussing study results regarding judges’ opinions of falsifiability, error rate, peer review, and general acceptance). Whether “peer review” is really distinct from general consensus is disputable. Further, the peer review process has been criticized as being motivated by the scientific establishment who staffs the journals and reviews the papers to keep the status quo firmly in place and leads to an indefensible practice of rejecting new or novel scientific discoveries. See generally 287 JAMA 2739-2898 (June 5, 2002) (focusing on peer review quality). Nevertheless, Sophia Gatowski’s study found that 71% of judges understood the utility of the peer review process, so perhaps it is useable merely because Judges understand it. See Gatowski, *supra* note 14, at 447.

²⁰⁷ See Vickers, *supra* note 8, at 143 (suggesting judges may not have understanding of reliability factors or how to apply them); see also RAND Study, *supra* note 9, at 40 (finding over time, judges used non-*Daubert* reliability factors more often).

²⁰⁸ See Gatowski, *supra* note 14, at 453 (implying judges may never understand practical value of *Daubert* unless provided sufficient scientific education).

²⁰⁹ See Lepore, *supra* note 45, at 1140; sources cited *supra* note 46 (listing eight states where *Frye* is still good law). “Fifteen states, along with the District of Columbia, continue to rely on

Because The RAND study and Gatowski's work empirically demonstrate that post-*Daubert*, judges did not use the *Daubert* tests to make their decisions,²¹⁰ several questions arise:

1. If judges did not use the *Daubert* tests post-*Daubert*, why not?
2. What accounted for the shift in admissibility in the post-*Daubert* years?
3. On what basis did judges make their decisions up until 1997, and
4. On what basis are courts making their decisions now?

The importance of these questions bears on two additional issues: (1) if we don't know on what basis judges determine admissibility, appellate review becomes problematic; (2) if the reason judges are not following *Daubert* is want of understanding, then it behooves us to set out a training program that is accessible and which they will be able to follow—if indeed this is possible, a question Vickers raises.²¹¹

III. HOW DID WE GET HERE AND HOW DO WE GET OUT?

Before proceeding to simplify and legitimize a workable system to enable gatekeepers to weed out noxious “pseudo-pscientific” [sic] interlopers and voodoo science, it is worthwhile to examine how the current situation came about. Certainly once *Frye* was no longer valid, the court was hard pressed to come up with an alternative:

When agreement about what constitutes scientific knowledge can range so widely that even long-held ideas are challenged, it is not easy to come up with a workable alternative to the *Frye* test, which requires the judge to be an arbiter of the views of practicing scientists. Trying to decide which expert is reasoning properly seems a rather difficult task for a court, when even scientists often

the *Frye* standard. . . . *Frye* is not only alive, but it is the plurality rule in state courts, which are the venue for the vast majority of litigation.” Lepore, *supra* note 45, at 110 (quoting David E. Bernstein, *Frye, Frye, Again: The Past, Present and Future of the General Acceptance Test*, 41 *Jurimetrics* 385 (2001)); Kaufman, *supra* note 45, at 1-3 (2006) (listing states adopting *Daubert*). This also manifests a continuing trend since the decision was issued in 1993 and the Federal Rules were amended in 2000. See Lepore, *supra* note 45, at 1140.

²¹⁰ See Gatowski, *supra* note 14, at 453; RAND Study, *supra* note 9, at 40.

²¹¹ See Vickers, *supra* note 8, at 110 (“Thus, criticisms aimed at the text itself, specifically at the criteria for reliability . . . would seem to miss the mark. The better questions to ask are why the decision failed to provide criteria that judges actually find useful in drawing these distinctions and what criteria judges are, in fact, using to distinguish between admissible and inadmissible evidence. . . . It also begs the serious question of why the decision did not create a more useful doctrinal test.”).

disagree on how to do it.²¹²

Without any grounding in science themselves, the *Daubert* majority turned to the world of philosophy, perhaps a more comfortable intellectual venue, to find an anchor.²¹³ Hence *Daubert* derives its definition of “science” (upon which the holding hinges) from Karl Popper, the psychologist turned philosopher who relentlessly argued that “science” considers matters which are “falsifiable” and subjects hypotheses to the test of “falsification.”²¹⁴ In other words, Popper claims that scientists proffer hypotheses and then go about trying to disprove them.²¹⁵ Luckily for the *Daubert* court, this view is supported by a narrow cadre of scientists, physicists mainly, most of whom were engaged in the world of quantum theory and cosmology.²¹⁶ It is not surprising that the American Association for the Advancement of Science (AAAS) submitted an amicus brief supporting the Popperian view (and cited by the court), as the president at the time the brief was written was an eminent Nobel Prize winning theoretical physicist.²¹⁷

While Popper’s definition is, indeed, embraced by quantum physicists²¹⁸ (a subject totally irrelevant to the courtroom), it has been rejected²¹⁹ by the formulators and codifiers of the scientific method

²¹² Jan Beyea & Daniel Berger, *Scientific Misconceptions Among Daubert Gatekeepers: The Need for Reform of Expert Review Procedures*, DUKE L.J. 327, 355 (2001).

²¹³ See *Daubert v. Merrell Dow Pharm.*, 509 U.S. 579, 593 (1993) (citing Karl Popper’s definition of “falsifiability”).

²¹⁴ See KARL POPPER, CONJECTURES AND REFUTATIONS: THE GROWTH OF SCIENTIFIC KNOWLEDGE 37 (1989) [hereinafter POPPER, CONJECTURES AND REFUTATIONS]; KARL POPPER, THE LOGIC OF SCIENTIFIC DISCOVERY 16, 315 (2002).

²¹⁵ See POPPER, CONJECTURES AND REFUTATIONS, *supra* note 214, at 37 (“[T]he criterion of the scientific status of a theory is its falsifiability, or refutability, or testability”).

²¹⁶ See Billauer, *Admissibility of Scientific Evidence under Daubert*, *supra* note 16. The author argues Popper’s views are irrelevant to the science of courtroom, i.e., biology, physiology and medicine, chemistry and simple (Newtonian) physics, which are totally different from modern physics in approach and practice. See *id.*; see also GEORGE GORE, THE ART OF SCIENTIFIC DISCOVERY: OR, THE GENERAL CONDITIONS AND METHODS OF RESEARCH IN PHYSICS AND CHEMISTRY 14 (1878) (stating biology is more complicated than either physics or chemistry).

²¹⁷ See Lepore, *supra* note 45, at 1146. In fact, one commentator suggested that the amicus brief has become an evidentiary backdoor, a way for “expert testimony” to be introduced into the record without having to satisfy the Federal Rules of Evidence standards for expertise. *Id.*

²¹⁸ The AAAS president at the time was Leon Lederman, the Nobel Laureate physicist who discovered the 6th quark and is a devout champion of the philosophy of Karl Popper. See Discussion with Leon Lederman, Director Emeritus, Fermi National Accelerator Laboratory (Summer 1996). In a conversation regarding the definition and meaning of “cosmology,” Professor Lederman expressed to me his opinion made without reservation that the definition of science turns on “falsifiability.” *Id.*

²¹⁹ The concept was brilliantly mocked in literature of the day. See SINCLAIR LEWIS, ARROWSMITH 1 (1924). *Arrowsmith* tells the story of a young doctor glorifying the ideal of pure research, hard data and controlled experiments—with the objective of falsifying results (“the

research as it applies to medicine and physiology,²²⁰ chemistry,²²¹ and Newtonian physicists²²² (i.e., the scientific subjects of interest in the courtroom). Further, Popper's approach is incompatible with basic formulations of causation as contemplated in law, a subject I tackle in a companion paper.²²³ These facts alone should render *Daubert* an inherently invalid method upon which to evaluate proffered scientific evidence.

Layered upon these problems are a hodge-podge of confusing or inartfully explained tests, and the co-mingling the distinct scientific concepts of validity (testing designed to prove or verify the hypothesis in question) and reliability (results which are capable of being reproduced by independent investigators, without regard to its veracity in terms of proving the hypothesis in question). In science, both tests must be satisfied to pass constraints of scientific rigor. In *Daubert* these tests are amalgamated into a single precept of legal reliability.²²⁴

As to why these tests have not been implemented, commentators and surveys report that judges feel unnerved by the gatekeeper role

destroyers") as Lewis called them, "men [who] had more fun destroying other people's theories than creating their own." *Id.* at 42, 117. Ultimately young Arrowsmith gives up his practice, leaves his wife and walks out on his son to work as a recluse on research of his choosing even with little likelihood of success "whether it should be the structure of the atom, or a disproof of the results of Drs. Wickett and Arrowsmith." *Id.* at p. 427.

²²⁰ See CLAUDE BERNARD, AN INTRODUCTION TO THE STUDY OF EXPERIMENTAL MEDICINE 177 (Henry C. Greene trans., Dover Publications (1957)) ("Negative facts, no matter how numerous they may be, can never destroy a single positive fact. That is why pure and simple negation is not criticism, and this method should be absolutely rejected in science, because science is never built up by negation."). Claude Bernard's objective was to establish the use of the scientific method in medicine and proof by experimental verification. *See id.*

²²¹ See ROGER PENROSE ET AL., THE LARGE, THE SMALL AND THE HUMAN MIND 166 (Malcolm Longair ed., 1999). As Nancy Cartwright says, "[e]ssentially all sciences except physics are special sciences. That means that their laws hold at best only *ceteris paribus*. They hold only so long as nothing from outside the domain of the theory in question interferes." *Id.* Kuhn notes that physicists and chemists do not even "see" the same things the same way, even something as elemental as a molecule. *See* KUHN, *supra* note 11, at 50-51.

²²² See PLANCK, *supra* note 185, at 131 (noting difference between classical physics and "quantum" or modern physics); *see also* JAMES K. CHANDLER ET AL., QUESTIONS OF EVIDENCE: PROOF, PRACTICE, AND PERSUASION ACROSS THE DISCIPLINES 492 (1994) (noting that some philosophers of science believe that "the laws of physics lie."). This approach would clearly be an anathema to *Daubert* which insists on vigorous compliance with the scientific method, presumably meaning by controlled experiment.

²²³ See Billauer, *Admissibility of Scientific Evidence under Daubert*, *supra* note 16 (demonstrating Popper's views are incompatible with joint or multi-causation, substantial causation, etc.). Popper eschews the idea of causation or determinism. *See id.*; *see also* J. L. HEILBRON, DILEMMAS OF AN UPRIGHT MAN: MAX PLANCK AS SPOKESMAN FOR GERMAN SCIENCE 66 (1986) ("The assumption of an absolute determinism is the essential foundation of every scientific inquiry.").

²²⁴ See *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 590, 590 n.9 (1993) (articulating how scientific reliability in the law will be determined based only upon scientific validity).

imposed by *Daubert*.²²⁵ The RAND “findings also illuminate concerns that judges may not be competent to perform the task of distinguishing between good and bad science. . . . [But] perhaps the most common complaint about *Daubert* is that it forces judges to become amateur scientists, a role they are not well-prepared for and should not be asked to play.”²²⁶ Additional data suggests that judges had a hard time or were ill-prepared to understand the *Daubert* criteria.²²⁷ Some investigators found that judges did not utilize the *Daubert* tests for want of understanding.²²⁸ The dissent of Chief Justice Rehnquist²²⁹ stated his concern outright—that judges might have difficulty comprehending the concept of falsifiability, a sentiment corroborated by Gatowski’s empirical research.²³⁰ For whatever reason, “the data would seem to imply that *Daubert* has failed to provide a framework that judges, at least in recent years, find useful in distinguishing between good and bad science, or admissible and inadmissible evidence.”²³¹

This may well sum up the current state of affairs—judges just do not have the training or ability to discern the difference between good and bad science.²³² If that is indeed the case, one wonders about the ability of the *Daubert* judges to even conjure the tests in the first place.

²²⁵ See Gatowski, *supra* note 14, at 433.

²²⁶ *Id.*

²²⁷ See RAND Study, *supra* note 9, at 6. Practically speaking, the RAND data suggests that whether a case is heard under *Frye* or *Daubert*, the outcome would likely be the same. See *id.*; Vickers, *supra* note 8, at 194. This conclusion is confirmed by Cheng and Yoon’s study that showed “that a state’s choice of evidentiary standards does not have a statistically significant effect on removal rates in that state,” from which Cheng and Yoon inferred that a “state’s adoption of *Frye* or *Daubert* makes no difference in practice.” Vickers, *supra* note 8, at 195. Others, including Andrew Jurs and Scott DeVito, disagree. See Andrew Jurs & Scott DeVito, *The Stricter Standard: An Empirical Assessment of Daubert’s Effect on Civil Defendants*, 62 Cath. U. L. Rev. 675, 680-681 (2013) (concluding *Daubert* is stricter standard for expert admissibility).

²²⁸ See Gatowski, *supra* note 14, at 433.

²²⁹ See *Daubert*, 509 U.S. at 598-601 (Rehnquist, J., dissenting) (warning dangers of using falsifiability as definer of “science”). “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.” *Id.* at 600.

²³⁰ See Vickers, *supra* note 8, at 114 (discussing Chief Justice Rehnquist’s dissent); Gatowski, *supra* note 14, at 433 (evaluating results from survey of state court judges). Gatowski’s research indicates that judges had difficulty in understanding these considerations, though and while it seems that more attention was paid to the evidence, it cannot be said that it was evaluated based on the *Daubert* parameters. See Gatowski, *supra* note 14, at 433. Thus, by comparison to the *Frye* test where 82% of judges surveyed demonstrated a clear understanding of “general acceptance,” Gatowski’s survey of 400 state court judges demonstrated that “only 4% of judges demonstrated a clear understanding of falsifiability and of error rates.” *Id.* at 447-49.

²³¹ Vickers, *supra* note 8, at 142.

²³² See Neil Vidmar & Shari S. Diamond, *Juries and Expert Evidence*, 66 BROOK. L. REV. 1121, 1172 (2001) (discussing Gatowski survey). Gatowski found 73% of judges had no experience with epidemiological data. *Id.* In addition, 96% of judges reported that they had not

An alternative explanation is possible: It may be that the methods and approaches of *Daubert* (and as we have seen in *Frye*) are just plain wrong. I posit judges somehow instinctively know this (again, just like *Frye*) and hence refuse to adhere to the *Daubert* tests. Lacking an alternative, but recognizing vigilance is necessary they turn to the *Frye* test—again, as a cover. If this is the case, as some commentators suggest, a different educational approach than is provided in *Daubert* or conventional training programs is necessary.²³³

But before leaving *Daubert* and RAND, it is necessary to ask the same question we asked of the *Frye* Appellate Court: What motivated the *Daubert* judges to create a new paradigm and reject *Frye*? In other words, why did they do it—especially if the science involved was by no means “novel”? And since *Daubert* encumbers judges with acquiring a much greater understanding of the science behind the evidence than was asked of judges under *Frye*, when judges have little training in scientific methods, reasoning, and research, a second question must be asked: Was their decision wise?

That the *Frye* test is mired in the status quo, with little hope of integrating new scientific advances into the developing matrix of case law is clear. Thus, in response to the first question, I suggest that given the speed with which science is advancing, and the increasing awareness that the hallowed sanctums of “science” and its established scientists move slowly in welcoming newcomers, it is not surprising that we need a newer, more adaptive and adaptable legal paradigm to address innovation, technical breakthroughs and maverick scientists.

The second question is a bit harder. Certainly, the *Daubert* court searched for a different paradigm to help them sort out their dilemma. But the key question is why did the Justices need help at all? After all, complex science is not new to the courts, as will be seen in the following section.

A. Can Judges Understand Science?

As mentioned earlier, the first epidemiological case was presented to the Supreme Court in 1916.²³⁴ The first brief containing scientific

received instruction about general scientific methods and principles (even though the majority had received some CLE training). *Id.*

²³³ See Vickers, *supra* note 8, at 137 (“Cheng and Yoon then recommend that attempts at changing or improving the way courts treat scientific evidence could be more effective if ‘advocates for rigorous use of scientific evidence shifted their focus away from tinkering with doctrinal tests and instead toward “softer” solutions that increase the judiciary’s understanding of scientific concepts and processes.’”).

²³⁴ See *United States v. Forty Barrels & Twenty Kegs of Coca-Cola*, 241 U.S. 265 (1916)

evidence, *Muller v. Oregon*,²³⁵ was submitted to the Supreme Court in 1908!²³⁶ Nicknamed “the Brandeis Brief,”²³⁷ the brief was 104 pages, of which 102 were entirely devoted to scientific evidence.²³⁸ The Supreme Court at the time, including Melville Fuller, David Brewer (who wrote the opinion),²³⁹ Joseph McKenna, William Day, William Moody, John Harlan, Edwin White, and Oliver Wendell Holmes²⁴⁰ all finished their schooling during the mid to late 1800s—at a time the germ theory of disease had not yet gained general acceptance—apparently had no problems understanding the science.²⁴¹ It must be noted that while at least some of the Justices attended top flight colleges (e.g., Harvard and Yale), their courses of study did not concentrate on the sciences.²⁴²

(considering whether caffeine added to food was deleterious ingredient).

²³⁵ 208 U.S. 412 (1908).

²³⁶ See Brief for the State of Oregon, *Muller v. Oregon*, 208 U.S. 412 (1908) (No. 107), 1908 WL 27605.

²³⁷ See *id.* It was written by Louis Brandeis and his sister-in-law, Josephine, a sociologist (aided by her sister, Pauline), and concerned the effects on women of unhealthy working conditions. See Florence Kelley, *THE SELECTED LETTERS OF FLORENCE KELLEY, 1869-1931* 493-94 (Kathryn K. Sklar & Beverly W. Palmer eds., 2009).

²³⁸ See Ruth Bader Ginsburg, *Muller v. Oregon: One Hundred Years Later*, 45 WILLAMETTE L. REV. 359, 356 (“A British epidemiological study showed a higher prevalence of disease among laundry workers than among other women.”).

²³⁹ The brief was mentioned in the opinion in an unusual and explicitly admiring footnote by Justice Brewer. See *Muller*, 208 U.S. at 419 n.† (“In the brief filed by Mr. Louis D. Brandeis for the defendant in error is a very copious collection of all these matters, an epitome of which is found in the margin.”).

²⁴⁰ The philosophy of Justice Oliver Wendell Holmes, although he did not write the opinion, may have a bearing on how judges should interpret science today:

The life of the law has not been logic: it has been experience. The felt necessities of the time, the prevalent moral and political theories, intuitions of public policy, avowed or unconscious, even the prejudices which judges share with their fellow-men, have had a good deal more to do than syllogism in determining the rules by which men should be governed. The law embodies the story of a nation's development through many centuries, and it cannot be dealt with as if it contained only the axioms and corollaries of a book of mathematics.

OLIVER WENDELL HOLMES, JR., *THE COMMON LAW* 1 (1881). The true basis of the decision was often an “inarticulate major premise,” however. A judge was obliged to choose between contending legal arguments, each posed in absolute terms, and the true basis of his decision was sometimes drawn from outside the law, when precedents were lacking or were evenly divided.

²⁴¹ See Ginsburg, *supra* note 238, at 364-65 (discussing Court's evaluation of scientific arguments in case). What would have happened if *Frye* had been appealed to the Supreme Court is an interesting question. I propose the Court, which at that time included two of the Justices who heard the “Brandeis Brief,” including Holmes, and also Brandeis himself, would have ratified the holding—but, not being frightened of either science or Wigmore, they would have struck down the “general consensus” basis which has been plaguing courts ever since.

²⁴² See Barbara P. Billauer, Aaron Aaronsohn and the Zionist Conversion of Justice Louis Brandeis, *POLICY STUDIES ORG.* 1, 7 (2015), available at <http://ssrn.com/abstract=2414425>. The

Science-phobia in the courts may be the root cause of the problems generated by both *Daubert* and *Frye*. It appears that in the first quarter of the twentieth century, judges had little problem understanding science—even science which might flummox today’s judges.²⁴³ The excerpts in the following section are presented in extensive quotations, as to summarize them would constitute an injustice to the level of judicial understanding at the time.

B. Novel Science: The Rip Van Winkle Effect

Somehow, the ability of judges to understand science went into deep sleep—around the time toxic tort cases began to proliferate. By the mid-1980s, it appears judges began to fall asleep at the prospect of scientific evidence being admitted.²⁴⁴ By 1991, they were in suspended animation and refused to awaken.²⁴⁵ But lest anyone think judges are incapable of learning science, a brief foray into life before *Frye* is instructive. It appears that before *Frye*, judges were quite capable to master not only current science, but also the intricacies of novel science.

The issue of harmful consequences of chemical additives to food was first addressed legislatively in the United States in 1906.²⁴⁶ Less than a decade later, the first cases presented themselves to American courts, and the judiciary proved they were up to the mettle of understanding the scientific minutiae of harm.²⁴⁷ The 1911 opinion of District Judge Sanford in *United States v. Forty Barrels & Twenty Kegs of Coca-Cola*²⁴⁸ is presented first by way of example. In that case, the court had to consider whether Coca-Cola, designated a food product, was “adulterated,” in that it contained “an added ingredient, caffeine,” which was alleged to be a

concept of scientist-legal scholars was not unknown at the time. *See id.* Roscoe Pound, who began teaching at Harvard Law School in 1910 and later became Dean in 1916, had a Ph.D. in botany and had a wide array of friends who were practical and theoretical agronomists. *See* N. E. H. HULL, ROSCOE POUND AND KARL LLEWELLYN: SEARCHING FOR AN AMERICAN JURISPRUDENCE 41 (1st ed. 1998). By contrast, Pound never finished law school—he read for the bar at his father’s firm, after completing only one year of law school and not doing very well. *See id.* at 8.

²⁴³ *See* discussion *infra* Part III.B. (comparing scientific understanding of judges throughout twentieth century).

²⁴⁴ *See* discussion *supra* Part II.D. (discussing emergence of “junk science” in 1980s).

²⁴⁵ *See id.* (recounting emergence of silicone breast implant litigation in early 1990s).

²⁴⁶ *See* Adam Burrows, *Palette of Our Palates: A Brief History of Food Coloring and Its Regulation*, 8 COMPREHENSIVE REV. IN FOOD SCI. AND FOOD SAFETY 394, 396 (2009), available at <http://onlinelibrary.wiley.com/doi/10.1111/j.1541-4337.2009.00089.x/epdf>.

²⁴⁷ *See* sources cited *infra* note 254 (noting wave of litigation following passage of 1906 Food and Drug Act).

²⁴⁸ 191 F. 431 (E.D. Tenn. 1911), *aff’d*, 215 F. 535 (6th Cir. 1914), *rev’d*, 241 U.S. 265 (1916).

“poisonous and deleterious ingredient that may render such food product injurious to health,” and in so doing, was forced to consider the effects of different doses of a substance allegedly causing harm—a rather sophisticated concept that troubles even today’s judges.²⁴⁹

As the court in *United States v. Lexington Mill & Elevator Co.*²⁵⁰ noted:

The trial judge decided that if the added substance was qualitatively poisonous, although in fact added in such minute quantity as to be noninjurious to health, it still fell under the ban of the statute; and the distinction is sought to be drawn between substances admittedly poisonous when administered in considerable quantities but which serve some beneficial purpose when administered in small amounts, and those substances which it is claimed never can benefit and which in large doses must injure. This distinction is refined. To apply it must presuppose that science has exhausted the entire field of investigation as to the effect upon the human body of these various substances; that nothing remains to be learned.²⁵¹

While certainly a “novel” issue, the court nevertheless proceeded to unapologetically and fearlessly address the issue—without the need for intervention or reliance on any opinions outside those presented to the court, such as general consensus of scientists.²⁵²

The passage of the 1906 Food and Drug Act resulted in the government’s vigorous attack on adulteration requiring a fairly sophisticated level of scientific understanding—both of food processing and of the results of chemical contamination.²⁵³ The success of the government in both criminal and civil cases attested to their proficiency.²⁵⁴

²⁴⁹ See *id.* at 431.

²⁵⁰ 202 F. 615 (8th Cir. 1913).

²⁵¹ *Id.* at 619.

²⁵² See *id.* (analyzing complex issues of “adulteration” and effects).

²⁵³ See Wesley E. Forte, *The Food and Drug Administration and the Economic Adulteration of Foods*, 41 IND. L.J. 346, 346-49 (1966) (discussing history of government in response to food adulteration).

²⁵⁴ See, e.g., *Union Dairy Co. v. United States*, 250 F. 231 (7th Cir. 1918) (discussing milk diluted by water); *Frank v. United States*, 192 F. 864 (6th Cir. 1911) (discussing pepper diluted by corn); *United States v. Frank*, 189 F. 195 (S.D. Ohio 1911) (discussing lemon extract diluted by alcohol and water); *United States v. South Hero Creamery Ass’n, White & Gates* 1142 (D. Vt. 1925) (discussing butter with less than 80 per cent milk-fat); *United States v. Atlantic Macaroni Co., White & Gates* 793 (E.D.N.Y. 1917) (discussing macaroni dyed yellow to conceal inferiority); *United States v. German American Specialty Co., White & Gates* 459 (S.D.N.Y.

Nor did juries seem to have a problem understanding the rather technical and scientific evidence.²⁵⁵

And even scientific evidence at the cutting edge—requiring an understanding of the impact of chemicals at the cellular level—did not flummox courts. For example, in a decision that would put today’s judges to shame, the court in *Lexington Mill & Elevator Co. v. United States*, decided at the turn of the last century, dealt with the issue of whether bleached flour is considered harmful.²⁵⁶ The decision is astounding in demonstrating the level of mastery of science.

Before reading the decision, it might be informative to note that the nature of hemoglobin was first established by E. F. Hoppe-Seyler in 1857.²⁵⁷ The reaction between nitrite and hemoglobin was first studied only about a decade before the decision,²⁵⁸ and the quantitative relationship enunciated in 1911, two years before the appellate decision.²⁵⁹

1913) (discussing eggs diluted by skim milk); *United States v. Libby, McNeill & Libby*, 210 F. 148 (4th Cir. 1913) (discussing condensed skimmed milk diluted by sugar); *United States v. Sixty Barrels of Wine*, 225 F. 846 (W.D. Mo. 1915) (considering claret wine diluted by pomace wine); *William Henning & Co. v. United States*, 193 F. 52 (5th Cir. 1912) (evaluating tomato catsup diluted by pumpkin); *United States v. One Hundred Barrels of Vinegar*, 188 F. 471 (D. Minn. 1911) (discussing cider vinegar diluted by distilled vinegar); *United States v. 420 Sacks of Flour*, 180 F. 518 (E.D. La. 1910) (addressing flour bleached to conceal inferiority).

²⁵⁵ See, e.g., *United States v. 3998 Cases of Canned Tomatoes, White & Gates* 1213 (D. Del. 1928) (revealing that jury failed to find that excess water had been added to canned tomatoes); *United States v. 4% Cases of Crème De Menthe, White & Gates* 1191 (E.D. Mo. 1926) (failing to find caffeine had been substituted in part for crème de menthe flavor non-alcoholic cocktail); *United States v. Barnet Creamery Ass’n, White & Gates* 1149 (D. Vt. 1925) (failing to find butter deficient in butterfat); *United States v. 37 One Pound Packages of Colors, White & Gates* 1165 (E.D. Pa. 1925) (failing to find that food colors had been diluted by paste); *United States v. Marmarelli, White & Gates* 1122 (S.D.N.Y. 1924) (failing to find that defendants had diluted olive oil with cottonseed oil); *United States v. Potter, White & Gates* 409 (E.D. N.C. 1912) (failing to find that excess water was used in canning oysters); *United States v. Heide, White & Gates* 325 (S.D.N.Y. 1911) (failing to find that 5% glucose reduced the quality of almond paste).

²⁵⁶ See *United States v. Lexington Mill & Elevator Co.*, 202 F. 615, 621-22 (8th Cir. 1913), *rev’d* 202 F. 615 (8th Cir. 1913), *aff’d* 232 U.S. 399 (1914).

²⁵⁷ See Andras Gedeon, *SCIENCE AND TECHNOLOGY IN MEDICINE: AN ILLUSTRATED ACCOUNT BASED ON NINETY NINE LANDMARK PUBLICATIONS FROM FIVE CENTURIES* 268 (2007) (outlining ninety nine most significant discoveries in medical field).

²⁵⁸ See Daniel B. Kim-Shapiro et al., *The Reaction Between Nitrite and Hemoglobin: The Role of Nitrite in Hemoglobin-Mediated Hypoxic Vasodilation*, 99 J. OF INORGANIC BIOCHEMISTRY 237-46 (2005) (discussing historical studies and discovery of hemoglobin and heme-compounds).

²⁵⁹ See Leon A. Greenberg et al., *The Reaction of Hemoglobin with Nitrite*, J. OF BIOLOGICAL CHEMISTRY 665, 665 (1943) (“Although nitrite has been extensively used as an agent for the formation of methemoglobin, the quantitative relations of the reaction appear to be uncertain; three widely different values have been reported for the amount of methemoglobin formed to nitrite utilized. The first statement of a quantitative relationship was made by Barcroft and Müller . . . in 1911.”); see also Joseph Barcroft & Franz Müller, *The Relation of Blood-flow to Metabolism in the Submaxillary Gland*, 44 J. PHYSIOL. 259, 259 (1912) (stating quantitative

The Eighth Circuit court opinion is brought down here in substantial measure, to demonstrate the extreme sophistication of knowledge and understanding of the judiciary at the time:

We are not unmindful of the contention that the evidence conclusively shows that our subjected to the bleaching process is injurious to health in some degree, even if its injurious effect is so slight as to be incapable of observation, and that hence the instruction we have found to be error was error without prejudice. This contention is founded upon expert testimony as to the result from the taking of nitrites into the human system. It is said that nitrites taken into the human body act upon the coloring matter of the red corpuscles of the blood so as to change the hemoglobin of the blood into methemoglobin. In the language of one of the chief chemical experts of the government this effect is thus described:

‘In the blood stream there are red corpuscles, invisible to the naked eye, which contain a red coloring substance known as hemoglobin, when not combined with oxygen, and when combined with oxygen forming a dissociable compound, oxyhemoglobin. In respiration, the hemoglobin contained in the red corpuscles of the venous blood is brought into the lungs, where it having an affinity for the oxygen, which is one of the gaseous constituents of the air, combines with the oxygen to form oxyhemoglobin. This oxyhemoglobin contained in the red blood corpuscles is then conveyed, through the arterial system, to the various parts of the body, and of the terminals of the arterial system, passing through a mass of tissue, it gives up its oxygen, to oxidize the tissues, or materials that may be in solution there, to form carbon dioxide and to form water, and this oxyhemoglobin is thereby reduced to the condition of hemoglobin which is returned by the venous system to the lungs, to be again oxygenated. That is where the hemoglobin will again combine with oxygen to form

relationship). Even basic research on the compound was still being performed ten years after the case was decided. See James M. Neill, *Studies on the Oxidation-Reduction of Hemoglobin and Methemoglobin: The Changes Induced by Pneumococci and by Sterile Animal Tissue*, J. EXPERIMENTAL MED. 299, 299-313 (1925), <http://jem.rupress.org/content/41/2/299.full.pdf> (furthering studies on the reaction of hemoglobin with nitrite).

oxyhemoglobin, and a given quantity of hemoglobin may serve to carry a given quantity of oxygen to the system. Now, however, if any of this hemoglobin is converted into methemoglobin, which is a compound of oxygen with hemoglobin, in which the oxygen is more firmly combined than in the case of oxyhemoglobin, although the quantity of oxygen is the same, the oxygen is so firmly attached—combined with the hemoglobin—that the vital processes are not sufficiently strong to separate the oxygen from the hemoglobin, nor to use the oxygen to oxidize the tissue and tissue material, to sustain life, and, consequently, it passes through the circulation to the arterial system and the venous system, and continues this cycle until, finally, it is destroyed by the liver. Therefore, a certain quantity of the hemoglobin is rendered inefficient. It no longer functionates as a carrier of oxygen to the system, serves, or acts, as a foreign body in the blood circulation, and, therefore, must be removed. As I have said before, an extra strain is placed upon the liver, in order to remove it, and an extra strain is placed upon the red blood marrow, in adults, to regenerate the corpuscles, and to replace the corpuscles of the hemoglobin that have been rendered inactive by the action of nitrite, and the formation of methemoglobin.

It is also said that the continued presence of nitrites in the system does not develop and tolerance on the part of the body or means of neutralizing its normal action. On the other hand, it was proved that no injurious effect had ever been observed from the use of bleached flour, although such flour had been largely used; that nitrites in some or greater amounts are frequently present in potable water, bacon, ham, fruits, and certain vegetables, and even in the saliva of both adults and children, and no evil result has been detected; that urea usually present in saliva is, when taken into the stomach, a neutralizer of nitrites, and is a method by which nature averts harm from minute quantities of nitrites so constantly taken into the system. In this conflict of evidence it was essentially a matter for the jury to find the fact under proper instructions. **Expert testimony is but evidence. In case of dispute the controversy cannot be settled by the judicial knowledge**

of the court. It cannot be held that the evidence was so conclusive in favor of the government as to warrant the court in withdrawing this issue from the jury. . . .

. . . .

The mixture referred to in the first subdivision must be held to include a chemical compound as well as a mechanical mixture. **While this does not accord with the scientific definition of a mixture, yet in common acceptation mixtures and compounds are not discriminated. The evil intended to be remedied by the statute is not limited to a mechanical mixture, but is just as potent when the chemical union results from the two substances with the deleterious effect intended to be prevented by the act. . . .**

There was evidence that bleached flour did not improve with age in the manner characteristic of unbleached flour, nor did it, as the claimant contended, suddenly take on the condition of properly aged flour which had not been subjected to the bleaching process. That in dough made from bleached flour the gluten never attained the toughness found in dough from unbleached and properly aged flour, and that this toughness was a valuable property in the making of bread. In other words, that as an ultimate result of the mixing of the flour with nitrogen peroxide gas the bread-making quality had been injuriously affected. We are not concerned with the opposing testimony. It was for the jury to determine the fact, and the court did not err in refusing to peremptorily instruct for the claimant so far as the claim of adulteration was based on the first subdivision before quoted.²⁶⁰

C. Scientific Literacy in the Law—Another Approach

Why courts relied on *Frye* or created the *Daubert* tests when other standards of evidentiary acceptability were available remains a mystery. A good ten plus years before *Daubert*, courts set out standards both for evidence and for appellate review of such evidence as it pertained to health and safety in workplaces where toxic diseases were commonplace, which were commonsensical and easily applied: In *American Textile*

²⁶⁰ *United States v. Lexington Mill & Elevator Co.*, 202 F. at 621-22 (emphasis added) (citations omitted).

Manufacturers Institute v. Donovan,²⁶¹ all the court required—at least in an administrative setting—was “the best available evidence” for standard setting and “substantial evidence” for review.²⁶² The court noted that:

[P]etitioners urge not only that OSHA must show that a standard addresses a significant risk of material health impairment, but also that OSHA must demonstrate that the reduction in risk of material health impairment is significant in light of the costs of attaining that reduction. Respondents on the other hand contend that the Act requires OSHA to promulgate standards that eliminate or reduce such risks “to the extent such protection is technologically and economically feasible.”²⁶³

OSHA provides:

Agents under this subsection, shall set the standard which most adequately assures, to the extent feasible, **on the basis of the best available evidence**, that no employee will suffer material impairment of health or functional capacity even if such employee has regular exposure to the hazard dealt with by such standard for the period of his working life.²⁶⁴

The court held that:

[W]e have defined **substantial evidence** as “such **relevant evidence as a reasonable mind might accept as adequate to support a conclusion**.” The reviewing court must take into account contradictory evidence in the record, but “the possibility of drawing two inconsistent conclusions from the evidence does not prevent an administrative agency’s finding from being supported by substantial evidence.”²⁶⁵

²⁶¹ 452 U.S. 490 (1981).

²⁶² See *id.* at 490.

²⁶³ *Id.* at 506-07 (citations omitted).

²⁶⁴ *Id.* at 508 (emphasis added) (quoting Occupational Safety and Health Act, 29 U.S.C. § 655(b)(5)).

²⁶⁵ *Id.* at 522-23 (emphasis added) (citations omitted) (quoting *Univ. Camera Corp. v. NLRB*, 340 U.S. 474, 477 (1951) and *Consolo v. Fed. Mar. Corp.*, 383 U.S. 607, 620 (1966)).

D. How the Mighty Have Fallen

In the hundred years or so since the *Muller* and *Coca-Cola* cases were heard by the Supreme Court, it appears that the facility in the legal community to appreciate even simple mathematical concepts such as graph reading has diminished to the point of non-existence.²⁶⁶ The flawed interpretation of the RAND study results being promulgated and relied—without a hue and cry on the misuse of statistics and the misread of simple graphs—is, itself, an illustration of the problem.²⁶⁷ If lawyers cannot determine the statistical slant given by the impartial authors of the RAND study, how can they be trusted to ferret out statistical issues in evidence preferred by an advocate? If the myriad of commentators who have argued either that *Daubert* favored one side or the other, recently culminating in an “academic war” with one side advocating that a new “sufficiency” test be utilized to address their claim of “favoritism” with others vehemently disagreeing—not on the facts—but maintaining that *Daubert* is properly doing its job (which we now see is not the case), then how can we expect judges to make appropriate decisions?²⁶⁸

E. Conclusion: BING-ing science and the Universal Translator or the Need to Re-Define Science and the Scientific Method in Law

Certainly, these data highlight the importance of continuing empirical research regarding the effect of rulings post-*Daubert* in both *Daubert* and *Frye* jurisdictions; at present the data mysteriously stops circa 2000. But other important messages surface from this work as well. Various important words surface in the *American Textile* case: “Relevant,” “material harm,” “significant health impairment,” and “best available evidence”; words, that along with traditional evidentiary concepts of “competency, materiality and significance,” should suffice to filter out pseudo-science from the courtroom—from a legal perspective.²⁶⁹

From a scientific perspective, however, judges and lawyers must have some level of basic scientific and mathematical competency before

²⁶⁶ See Jehuda Reinharz, 1985 *Science in the Service of Politics: The Case of Chaim Weizmann during the First World War*, 396 THE ENG. HIST. REV. 572, 572 (1985). Complicated patent cases were also heard by the courts, for example, the patent of Harvey (Chaim) Weizmann regarding industrial production of acetone through the fermentation process of bacteria *Clostridium aceto-butylicum* in 1914. See *id.*

²⁶⁷ See discussion *supra* Parts I.C., II.A. (discussing flaws in RAND data and providing reanalysis).

²⁶⁸ See Green & Sanders, *supra* note 12, at 1 (arguing for need of sufficiency test); Twerski, *supra* note 26, at 641 (arguing *Daubert* factors are effective).

²⁶⁹ See *American Textile Mfrs. Inst. v. Donovan*, 452 U.S. 490 *passim* (1981).

any tests can be administered. The failures in reading the RAND graphs (generally taught at the seventh grade level) are illustrative. And while conventional judicial education programs sponsor world-class scientists to rub shoulders with the judiciary, such an approach is ill-suited to judicial gatekeeping.

The objection that we do not expect judges to be scientists—as some claim is required under *Daubert*—is valid. In a day or a week, there is no way even the most scientifically adept could learn the intricacies of genetics or artificial intelligence. What judges can be taught in a day is the meaning of “science.” A weekend would suffice to teach the difference between validity and reliability—two distinct scientific concepts—mangled and comingled in footnote nine of *Daubert*.²⁷⁰ What they can be taught in a week is some understanding of the scientific method and its limitations. These are general principles that apply to all scientific evidence, even cases not relying on science per se. As far as cases involving more sophisticated or interdisciplinary material such as epidemiological and toxic tort cases, which involve statistics, biostatistics, immunology, genetics, toxicology, obstetrics, and more, at least the basics should be introduced in law school as electives for students interested in pursuing these areas of law. Better cross-examination would better help judges, and better understanding of the basics helps everyone.

²⁷⁰ *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 590 n.9 (1993) (“In a case involving scientific evidence, evidentiary reliability will be based upon scientific validity.”). It is imperative to note that scientific validity and scientific reliability are separate indices of the plausibility and reproducibility (respectively) of the work. *See supra* note 10 and accompanying text (explaining differences between concepts of validity and reliability).