

2019

Federal District Court Decisions on the Admissibility of Expert Witness Testimony: An Empirical Study and a Revision of the Conventional Narrative in the Toxic Tort Context

George Horvath
Berkeley Law, ghorvath@law.berkeley.edu

Follow this and additional works at: <https://scholarship.law.uc.edu/uclr>

Recommended Citation

George Horvath, *Federal District Court Decisions on the Admissibility of Expert Witness Testimony: An Empirical Study and a Revision of the Conventional Narrative in the Toxic Tort Context*, 88 U. Cin. L. Rev. 515 (2019)

Available at: <https://scholarship.law.uc.edu/uclr/vol88/iss2/4>

This Article is brought to you for free and open access by University of Cincinnati College of Law Scholarship and Publications. It has been accepted for inclusion in University of Cincinnati Law Review by an authorized editor of University of Cincinnati College of Law Scholarship and Publications. For more information, please contact ronald.jones@uc.edu.

FEDERAL DISTRICT COURT DECISIONS ON THE
ADMISSIBILITY OF EXPERT WITNESS TESTIMONY: AN
EMPIRICAL STUDY AND A REVISION OF THE CONVENTIONAL
NARRATIVE IN THE TOXIC TORT CONTEXT

George Horvath

ABSTRACT

There is a generally accepted narrative about the development of the rules governing the admissibility of expert witness testimony. In this narrative over most of the twentieth century, the Frye rule imposed an inflexible requirement that evidence and testimony which was generally accepted was admissible while evidence that was not generally accepted was not. But the Supreme Court's 1993 decision in Daubert installed judges as the new arbiters of admissibility of expert testimony and replaced the Frye rule with a new standard: the relevance and reliability of the scientific evidence on which the expert witnesses rely. This revolution (one strand of the generally-accepted narrative continues) has allowed judges to unfairly tilt the playing field in favor of defendants—typically large corporations—by erecting impossibly high barriers to the admissibility of plaintiffs' expert witnesses.

This Article uses an empirical analysis of fifty-seven toxic tort cases to advance a counternarrative: The scientists were never in charge; rather, judges have long had the upper hand. Before Daubert, scientific evidence presented by experts was admitted in support of claims that seemed to the judge to have merit and rejected when the case appeared to be weak, allowing the judge to dispose of actions in the pretrial phase. The Daubert "revolution" did nothing to alter that balance of power as judges continue to exclude experts and dispose of cases at the evidentiary stage with the same frequency and in the same patterns as under the Frye standard.

The analysis does not allow a determination of the factors that underlie post-Daubert judges' admissibility decisions. But, whatever those factors may be, the Daubert "revolution" has neither clearly improved the outcomes in toxic tort cases nor ensured judges' adherence to the existing set of rules governing the admissibility of expert testimony and scientific evidence.

TABLE OF CONTENTS

I. Introduction	517
II. Of Infallible Liars and Frustrated Judges.....	518
A. The Rules: From <i>Frye</i> to the <i>Daubert</i> Trilogy	521
1. The <i>Frye</i> Rule	521
a. The <i>Frye</i> Rule and Its Discontents	521
b. Pre-Federal Rules Alternatives to the <i>Frye</i> Rule	523
2. The Federal Rules of Evidence	525
a. The Federal Rules of Evidence and the Bias in Favor of Admissibility.....	525
b. The Bias in Favor of Admissibility and the Junk Science Hysteria	528
3. The Rules Under the <i>Daubert</i> Trilogy	531
B. Scholars' Views of How Judges Are Using Their New Powers.....	536
III. Toxic Tort Litigation: The Choke Point of General Causation	539
IV. Empirical Study	541
A. Prior Empirical Work	542
B. Methods	543
C. Findings	546
1. Exclusion of Plaintiffs' Experts and Grants of Summary Judgment Before and After <i>Daubert</i>	547
2. All-or-Nothing Behavior in Admissibility Rulings.....	548
3. Non-Evidentiary Factors Underlying Judges' Admissibility Rulings.....	551
a. Characteristics of Cases	551
b. Judges' Political Views and Experience.....	553
c. The Number of Plaintiffs' General Causation Experts	555
V. Revising the Narrative	556
VI. Conclusion	558

I. INTRODUCTION

There is a generally accepted narrative about the development of the rules governing the admissibility of expert witness testimony. This narrative, in rough outline, runs as follows: There was the *ancien regime*—dating back to 1923—of the *Frye* rule, under which judges were required to defer to the scientific community as the arbiter of what scientific evidence and which expert witnesses would be admitted to testify at trial. *Frye* imposed an inflexible rule: evidence and testimony that was generally accepted by the scientific community was admissible; evidence that was not generally accepted was not. But, in 1993, the Supreme Court prompted a revolution in its *Daubert* decision, which deposed the scientific community as the arbiter of admissibility. Instead, the Court installed judges as the new arbiters of admissibility, and replaced the *Frye* rule with a new standard—the relevance and reliability of the scientific evidence on which the expert witnesses rely. This revolution (one strand of the generally-accepted narrative continues) has allowed judges to unfairly tilt the playing field in favor of defendants—typically large corporations—by erecting impossibly high barriers to the admissibility of plaintiffs’ expert witnesses.

This Article uses empirical evidence derived from an analysis of fifty-seven toxic tort cases to advance a counternarrative: The reality was that the scientists were never in charge. Complain as they might about untruthful experts and their apparent infallibility, judges have long had the upper hand. In cases decided before *Daubert*, scientific experts and their scientific evidence were useful tools to be admitted in support of claims that the judge deemed to have merit, and to be rejected when the case appeared to be weak, allowing the judge to dispose of the action in the pretrial phase. The *Daubert* revolution did nothing to alter that balance of power as judges are still excluding experts and disposing of cases at the evidentiary stage with the same frequency and in the same patterns as under the *Frye* standard. But post-*Daubert* judges are no longer relying on simplistic markers such as a case’s strength to guide their admissibility decisions.

Ultimately, I conclude that judges’ decisions on the admissibility of expert witnesses and scientific evidence have never been solely about whether the rules of admissibility that applied at the time were satisfied; rather, judges have been rendering, and continue to render, admissibility decisions based on other factors. The findings of this empirical study do not allow me to determine the factors that underlie post-*Daubert* judges’ admissibility decisions. But whether judges have been acting out of a pervasive cynicism about the ability of juries to appropriately weigh expert testimony, or out of a desire to further their own political and ideological agendas, or out of some other motive, judges have not been

playing by the rules in regards to admitting scientific evidence. While I will not argue that a results-oriented approach to admissibility rulings is unacceptable in all circumstances, I will argue that as currently practiced, judges' rulings on the admissibility of expert testimony and scientific evidence have been a failure from both a results-oriented and a process-oriented perspective.

This Article proceeds in four parts. Part I presents the conventional narrative, along with a more robust history of the rules governing the admissibility of expert witness testimony and scientific evidence. This discussion will show that the conventional narrative is incomplete in many places. Part II focuses on the requirements that a plaintiff advancing a toxic tort claim must meet in order to show why an analysis of toxic tort cases may be useful in elucidating judges' actual admissibility practices. Part III presents the findings of the empirical study that was conducted by the author which shows that the conventional narrative is not only incomplete, but also incorrect in many places. Part IV proposes a counternarrative to the traditional narrative that takes into account both the more complete history presented in Part I and the results of the empirical study. Part IV also discusses areas of promise for the work of scholars who want to understand what judges are actually doing in toxic tort cases when they render admissibility rulings on scientific evidence and expert witnesses.

II. OF INFALLIBLE LIARS AND FRUSTRATED JUDGES

How do courts handle expert witnesses and scientific evidence? This question—and its next of kin: how *should* courts handle expert witnesses and scientific evidence?—has long bedeviled plaintiffs and defendants, commentators, Congress, and most importantly, courts themselves. In 1874, Judge William Lawrence Foster repeated the words of a contemporary trial attorney, who in a closing argument said that “there are three kinds of liars,—the common liar, the d—d liar, and the scientific expert.”¹ The judge's frustration was understandable. Recounting a New Hampshire case over which he presided in 1874, he noted that three experts had testified that blood cells, when deposited onto clothing, may be “restored to perfect shape” after a period of ten years, while two other experts had testified that such a transformation was possible only after two weeks.² Judicial attitudes toward experts has changed little since the nineteenth century; Judge Jack Weinstein has recently stated that “[a]n expert can be found to testify to the truth of almost any theory, no matter

1. William L. Foster, *Expert Testimony,—Prevalent Complaints and Proposed Remedies*, 11 HARV. L. REV. 169, 169 (1897). The elision is Judge Foster's.

2. *Id.* at 172.

how frivolous.”³ In anonymous surveys, federal district court judges identified the most significant problem with expert witness testimony as being that experts too often “abandon objectivity and become advocates for the side that hired them.”⁴

If the problem were limited to the availability of experts willing to testify on demand to any theory, it would be of minor importance. The availability of experts bearing opposing opinions and the power of cross-examination would minimize the experts’ impact. But, a pervasive concern of the legal system is that the veneer of scientific objectivity will lead to unquestioning acceptance of scientific evidence by juries. One court described the power of scientific evidence as being so great that “scientific proof may in some instances assume a posture of mystic infallibility in the eyes of a jury of laymen.”⁵ Similar concerns attach to the experts who would testify at trials.

These concerns lend themselves to the creation of a “judges versus experts” trope which underlies the conventional narrative about the development of the rules of admissibility of expert witness testimony and scientific evidence. The apparent conflict between judges and experts has only grown more severe with time. Science has advanced dramatically, and scientific knowledge has moved farther and farther from the understanding of laymen jurors.⁶ As a result, the balance of power is said to have shifted even further in favor of the experts, whose ability to interpret the incomprehensible arcana of science is seemingly necessary to help the jury make correct decisions.⁷

The conventional narrative regarding the admissibility of scientific experts and testimony assumes that the rules of evidence governing the admissibility of expert testimony and scientific evidence have long served the purpose of balancing the competing interests of expert witnesses (and the parties on whose behalf they testify, and the lawyers who pay their fees) with those of judges and the validity of the adversarial process that they oversee. Under the *ancien regime*,⁸ i.e., the *Frye* rule, judges had to defer to the scientific community as the arbiter of what scientific evidence

3. ROGER C. PARK & RICHARD D. FRIEDMAN, EVIDENCE: CASES AND MATERIALS 85 (12th ed. 2013) (quoting Walter Olson, *The Case Against Expert Witnesses*, FORTUNE, Sept. 25, 1989).

4. 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, 328 tbl.6 (2002).

5. *United States v. Addison*, 498 F.2d 741, 744 (D.C. Cir. 1974). *See also* *United States v. Baller*, 519 F.2d 463, 466 (4th Cir. 1975) (“Because of its apparent objectivity, an opinion that claims a scientific basis is apt to carry undue weight with the trier of fact.”).

6. *See* SHEILA JASANOFF, SCIENCE AT THE BAR: LAW, SCIENCE, AND TECHNOLOGY IN AMERICA 42-43 (1995) (outlining the history of the role of science and scientific experts in legal cases).

7. *Id.* at 43.

8. The *ancien regime* analogy was drawn by Professor David L. Faigman in a recent law review article. *See* David L. Faigman, *The Daubert Revolution and the Birth of Modernity: Managing Scientific Evidence in the Age of Science*, 46 U.C. DAVIS L. REV. 893, 907 (2013).

and which expert witnesses would be admissible to testify. The *Frye* rule was inflexible—evidence that was generally accepted had to be admitted. The judges’ hands were tied. But, beginning in 1993, the Supreme Court prompted a revolution in its decisions that form the “*Daubert* trilogy”⁹ which removed the scientific community as the arbiter of admissibility, replacing it with judges as the new arbiters. The *Daubert* trilogy also rejected the *Frye* rule for a flexible new standard: the relevance and reliability of the scientific evidence on which the expert witnesses rely. Although it is not universally accepted, the narrative of the revolutionary impact of the trilogy is constantly retold by judges,¹⁰ by commentators,¹¹ and—showing the narrative’s general acceptance—in textbooks.¹²

Scholars have extended the conventional narrative to argue that modern-day judges are using their gatekeeping powers under *Daubert* to pursue various goals that are not permitted under the current rules. Some argue that the *Daubert* revolution has allowed judges to unfairly tilt the playing field in favor of defendants by erecting impossibly high barriers to the admissibility of plaintiffs’ expert witnesses. This tilt favors large corporations and represents a retreat by courts from the project of protecting consumers. Others argue that judges make admissibility rulings and dispose of cases in conformity with their political and ideological leanings. Others further argue that judges’ decisions are motivated by a desire to avoid reversal by appellate courts.

Section A of this Part provides a background on the rules of evidence that courts and Congress have developed over the past ninety years to deal with scientific evidence and experts. This history discusses the test

9. The “*Daubert* trilogy” consists of *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993), *General Electric Co. v. Joiner*, 522 U.S. 136 (1997), and *Kumho Tire Co. v. Carmichael*, 526 U.S. 137 (1999). See Part II.A, *infra*.

10. See *e.g.*, Berger, *infra* note 133, at 290 (reporting the results of a Federal Judicial Center study in which judges assessed their admissibility decision as having changed significantly after *Daubert*).

11. See, *e.g.*, Faigman, *supra* note 8, at 907 (“*Daubert* . . . marked a sea change and gave to courts the core instrument of the enlightenment, the scientific method. *Daubert* constituted a frontal assault on the Ancien Regime, thus giving birth to modernity in the law’s use of expertise.”); Lucinda M. Finley, *Guarding the Gate to the Courthouse: How Trial Judges Are Using Their Evidentiary Screening Role to Remake Tort Causation Rules*, 49 DEPAUL L. REV. 335, 335 (2000) (“Vigorously exercising their role as evidentiary ‘gatekeepers’—a task assigned to them by the United States Supreme Court in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*—federal trial judges in products liability cases have been doing far more than screening proposed expert testimony to determine admissibility. The *Daubert* gatekeeper power has become a potent tool of tort lawmaking.”); Cf. Sophia I. Gatowski et al., *Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-Daubert World*, 25 LAW AND HUMAN BEHAVIOR 433, 435-36 (2001) (discussing only *Frye* and *Daubert* in a review of the development of the rules governing the admissibility of expert testimony and scientific evidence). But see Joseph Sanders, *Science, Law, and the Expert Witness*, 72 LAW & CONTEMP. PROBS. 63, 78 (2009) (“The so-called *Daubert* revolution has pushed courts in the United States toward a slightly more inquisitorial posture.”).

12. RONALD J. ALLEN ET AL., EVIDENCE: TEXT, PROBLEMS, AND CASES 664 (5th ed. 2011) (“It is not too much to say that *Daubert* has prompted a revolution in the manner in which expert testimony is being treated by the lower courts.”).

established in *Frye v. United States*¹³ and the alternative tests that developed in response to that decision, the 1975 adoption of the Federal Rules of Evidence,¹⁴ and concludes with a discussion of the seminal 1993 case of *Daubert v. Merrell-Dow Pharmaceuticals*¹⁵ and the two subsequent cases that form the “*Daubert* Trilogy.” Section B then discusses recent scholarly literature that has theorized what judges have actually been doing in the *Daubert* era.

A. *The Rules: From Frye to the Daubert Trilogy*

1. The Frye Rule

Surprisingly, the rule that governed the admissibility of expert testimony and scientific evidence for most of the twentieth century was established neither by Congress nor by the Supreme Court; instead, the *Frye* rule was developed in the Court of Appeals for the District of Columbia. Subsection a describes the *Frye* rule and the limitations that courts and commentators complained the rule created. Then Subsection b discusses some of the mechanisms that courts developed in an attempt to escape those limitations.

a. *The Frye Rule and Its Discontents*

The test enunciated in the Court of Appeals for the District of Columbia’s 1923 *Frye v. United States* decision has been described as the “general acceptance” test.¹⁶ In *Frye*, the court was presented with a defendant charged with murder who had confessed but subsequently recanted.¹⁷ The defendant sought to admit evidence of his innocence in the form of an expert who would testify that results from a forerunner of the modern polygraph test established the defendant’s innocence.¹⁸ In evaluating the admissibility of this “new” form of evidence,¹⁹ the court noted that:

Just when a scientific principle or discovery crosses the line between

13. 293 F. 1013 (D.C. Cir. 1923).

14. Pub. L. No. 93-595, 88 Stat. 1926 (1975).

15. 509 U.S. 579 (1993).

16. *Daubert*, 509 U.S. at 585.

17. C.T. McCormick, *Deception-Tests and the Law of Evidence*, 15 CALIF. L. REV. 484, 499 n.49 (1927).

18. *Frye*, 293 F. at 1013-14.

19. Professor McCormick’s 1927 article concludes that by the time *Frye* was decided, physiologic disturbance tests (such as measuring blood pressure changes in response to questions) had been generally accepted as scientifically valid by psychologists. *Supra* note 17, at 503.

the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.²⁰

Although the *Frye* test was never—until the *Daubert* decision—challenged at the Supreme Court, it “dominated the admissibility of scientific evidence” in both federal and state courts for seventy years.²¹ The *Frye* test conditioned the admissibility of expert testimony on the general acceptance by the scientific community of the data or technique on which an expert relied. The *Frye* court’s brief opinion did not address two key issues, however: First, it did not address how a court is to determine when a technique has been generally accepted by the relevant scientific community. Second, it did not address how—or even whether—a court should evaluate an expert’s use of the technique or data in formulating an opinion.

The *Frye* test was criticized on general grounds as being too rigid.²² Much of the specific criticism leveled at the test concerned what judges and commentators saw as the test’s tendency to exclude too much scientific evidence and too many scientific techniques for too long.²³ Some commentators argued that the time required for a new technique or theory to obtain general acceptance “deprive[d] courts of reliable evidence.”²⁴ Judges themselves struggled with the straitjacket that the general acceptance test imposed. One judge, confronted with a case in which an anesthesiologist was suspected of murdering his wife by injecting her with the anesthetic drug succinylcholine—which was not detectable by conventional tests in 1968—rejected the delay inherent in the general acceptance rule, writing that “[s]ociety need not tolerate homicide until there develops a body of medical literature about some particular lethal agent.”²⁵ He concurred that evidence of the presence of succinylcholine in the murder victim, which had been generated by a novel test developed by the medical examiner—evidence that should have been excluded under the *Frye* general acceptance rule—had been properly

20. *Frye*, 293 F. at 1014.

21. Paul C. Giannelli, *The Admissibility of Novel Scientific Evidence: Frye v. United States, a Half-Century Later*, 80 COLUM. L. REV. 1197, 1205 & n.47 (1980).

22. *Id.* at 1223.

23. *Id.*

24. *Id.*

25. *Coppolino v. State*, 223 So. 2d 68, 75 (Fla. Dist. Ct. App. 1968) (Mann, J., concurring specially).

admitted by the trial judge.²⁶ Even proponents of the rule admitted that it “retard[ed] somewhat the admission of proof based on new methods of scientific investigation.”²⁷

Other courts struggled with the rigidity of the *Frye* rule in the opposite context, where the rule would have required admitting expert testimony that was generally accepted. Courts in criminal cases found “a strong countervailing” reason to limit the state’s expert testimony: “the defendant’s right to a fair trial.”²⁸ Thus, despite its surface inflexibility, the *Frye* general acceptance rule may have been circumvented by judges to exclude scientific evidence where they feared that such evidence would “assume a posture of mystic infallibility in the eyes of a jury of laymen,” and thus result in unfairness to a criminal defendant.²⁹

Although the goal of assuring a criminal defendant’s right to a fair trial is vital, so is the goal of not allowing murderers to go free while awaiting general acceptance. Cloaking decisions based on fairness in the mantle of an objective evidentiary rule thus raises a problem: when judges conflate evidentiary standards and burdens of proof, they are making decisions that may be dispositive of cases using means that are, at best, opaque. It seems likely that even in the straitjacketed *Frye* era, judges had been crossing a line, moving into the realm of projecting their sense of the overall merits of the cases before them onto specific evidentiary rulings.

b. Pre-Federal Rules Alternatives to the Frye Rule

Courts and commentators struggled with the general acceptance test, finding it to be over- and underinclusive, difficult to apply, and focused on the wrong consideration.³⁰ Even though no alternative standard gained widespread acceptance, two other standards were developed. However, a review of the case law from before the Federal Rules of Evidence were adopted shows that neither test was widely accepted, and that courts never defined standards for these tests. Thus, courts generally avoided the strictures of the *Frye* test on an ad hoc basis.

Looking back from 1980, Professor Paul Giannelli described a “traditional” approach to admissibility that he called the “relevancy approach”³¹ as an alternative to the *Frye* test. He noted that, in 1954,

26. *Id.*

27. *United States v. Addison*, 498 F.2d 741, 743 (D.C. Cir. 1974).

28. *United States v. Brown*, 557 F.2d 541, 556 (6th Cir. 1977). *See also* Giannelli, *supra* note 21, at 1248 n.380 (describing other cases in which courts considered the prosecution’s overall burden of proof when determining the admissibility of scientific evidence).

29. *Addison*, 498 F.2d at 744.

30. *See* Giannelli, *supra* note 21, at 1223-25 (discussing problems of overinclusiveness and underinclusiveness).

31. Giannelli, *supra* note 21, at 1232-33.

Professor Charles McCormick had written that general acceptance is not a “criterion for the admissibility of scientific evidence.”³² Instead, McCormick insisted that “[a]ny relevant conclusions which are supported by a qualified expert witness should be received unless there are other reasons for exclusion.”³³ The criteria under the relevancy approach included the qualifications of the expert witness and relevancy of the expert’s conclusions, with the extent of the scientific community’s acceptance of the underlying theory or technique being merely an issue of the weight, not the sufficiency, of the evidence.³⁴

Case law contemporary with McCormick’s 1954 textbook, and indeed case law predating the adoption of the Federal Rules of Evidence, does not show that the relevancy approach was widely accepted or that the terms of the approach were well defined. In the 1954 case *State v. Olivas*,³⁵ the Arizona Supreme Court heard a challenge to the trial court’s admission of the results of a “drunkometer” test³⁶ that allegedly showed the defendant to be legally intoxicated.³⁷ In rejecting the defendant’s challenge based on the existence of “scientific disagreement” over the accuracy of the test, the court held that “scientific disagreement affects only the weight and not the admissibility of evidence.”³⁸ *Olivas* has been cited as a case applying a rule that all relevant evidence should go to the jury.³⁹ It is not clear, however, that *Olivas* actually rejected general acceptance as the sine qua non of admissibility of scientific evidence. The court, in its brief opinion, rejected only a “lack of unanimity” standard.⁴⁰ However, under the *Frye* formulation of *general* acceptance, *unanimity* is not necessarily required. At most, then, *Olivas* stands for a rejection of the most stringent possible interpretation of *Frye*.⁴¹

Coppolino v. State,⁴² decided in 1968, appears to be the last major case cited in support of the relevancy approach before the advent of the Federal Rules of Evidence. In *Coppolino*, a medical examiner devised a novel test for detecting the presence of succinylcholine, a drug that induces

32. *Id.* at 1233 (quoting McCormick, Evidence 363-64 (1954)).

33. *Id.*

34. *Id.* at 1233 and n.280.

35. *State v. Olivas*, 267 P.2d 893 (Ariz. 1954).

36. The drunkometer was an early version of the modern-day Breathalyzer. See *State v. Warren*, 252 P.2d 781, 783 (Ariz. 1953), for a discussion of the device and its use.

37. *Olivas*, 276 P.2d at 894.

38. *Id.*

39. See Giannelli, *supra* note 21, at 1233 n.276.

40. *Olivas*, 276 P.2d at 894.

41. See also *People v. Bobczyk*, 99 N.E.2d 567, 570 (Ill. App. Ct. 1951); *McKay v. State*, 235 S.W.2d 173, 174 (Tex. Crim. App. 1950) (“In all probability a scientist may be found who will disagree with practically every generally accepted scientific theory.”).

42. *Coppolino v. State*, 223 So. 2d 68 (Fla. Dist. Ct. App. 1968).

paralysis and which is not found in the human body.⁴³ The examiner determined that traces of succinylcholine were present in the body of a deceased woman, forming the basis of the examiner's testimony in the woman's murder trial.⁴⁴ *Coppolino* differed from the earlier cases in that, unlike the Drunkometer, the medical examiner's test had never been performed before; thus, there existed no body of scientific literature on the test and no degree of acceptance of the test in the forensic science community.⁴⁵ Nonetheless, the Florida appellate court held that the trial judge had not abused his discretion in admitting the evidence because he had "listened to the testimony of the expert witnesses and in an exercise of his discretion ruled that the tests in question were sufficiently reliable to justify their admission."⁴⁶

It is difficult to extract from these cases a relevancy standard; in fact, *Coppolino* seems to apply a reliability standard instead of a relevance standard. A review of these court opinions and academic commentaries does show, however, that during the *Frye* era, there were judges and commentators who advocated for the use of relevance (however determined) as the standard for admissibility, while there were others who advocated for the use of reliability (however determined) as means of escaping the rigidity of *Frye's* general acceptance test.

2. The Federal Rules of Evidence

This Section reviews the Federal Rules of Evidence, focusing on Rule 702, which governs the admissibility of expert testimony. Then, this Section discusses some of the complaints that courts and commentators raised to the perceived ease with which expert testimony was admitted, and the court's responses to these complaints.

a. The Federal Rules of Evidence and the Bias in Favor of Admissibility

The Federal Rules of Evidence, enacted in 1975, are generally "biased in favor of admissibility" of evidence.⁴⁷ Under Section 402, *any* relevant evidence is admissible, unless barred by the Constitution, federal statute, another section of the Federal Rules, or a rule promulgated by the Supreme Court.⁴⁸ Relevance is given a broad definition in Section 401:

43. *Id.* at 69 ("[The medical examiner] 'testified that some of his tests and procedures were standard ones and that some were new.'").

44. *Id.*

45. *See id.* at 70-71.

46. *Id.* at 71.

47. PAUL C. GIANNELLI, UNDERSTANDING EVIDENCE 13 (2d. ed. 2006).

48. *See* FED. R. EVID. 402.

“Evidence is relevant if: (a) it has any tendency to make a fact more or less probable than it would be without the evidence; and (b) the fact is of consequence in determining the action.”⁴⁹ Thus, there is no minimum quantum of evidence required for relevance; any tendency, no matter how small, to make a fact more or less likely, satisfies Rule 401.⁵⁰ Likewise, a piece of evidence, even if it relates far more directly to other matters, is relevant if it bears even slightly on a fact in question.⁵¹

The original version of Rule 702, which governed the admissibility of expert testimony, was likewise biased in favor of admissibility: “If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise.”⁵² The Advisory Committee Notes to the 1972 Proposed Rules reflect the intention of the drafters that Rule 702 be broadly interpreted, noting specifically that expert testimony was not limited to opinions, that expertise is not limited to scientific and technical topics, and that experts include those who have experience, but not an advanced degree, in a field.⁵³ In fact, the topic of the expert’s testimony need not have been beyond the average juror’s comprehension; testimony was admissible if it would help jurors understand material even if the material was such that was difficult but which the jury could ultimately understand.⁵⁴

Focusing solely on the language of Rule 702, there would seem to be no place left for the *Frye* general acceptance test. But the assumption that the Federal Rules had superseded *Frye* was far from generally accepted. The advisory committee notes, congressional committee reports, floor debates, and hearings—the vast bulk of the legislative history—were all silent on the impact of the Rules on *Frye*.⁵⁵ Several commentators believed that “[i]t would be odd if the Advisory Committee and the Congress intended to overrule the vast majority of cases . . . without explicitly stating so.”⁵⁶ In essence, these writers applied the maxim of

49. FED. R. EVID. 401.

50. See *id.*, Advisory Committee Notes, 1972 Proposed Rules (noting that “[a]ny more stringent requirement is unworkable and unrealistic”).

51. *Id.*

52. FED. R. EVID. 702 (1975) (amended 2000). The current version of Rule 702 incorporates the Supreme Court holdings in *Daubert*, *Joiner*, and *Kumho*. See Fed. R. Evid. 702 advisory committee’s notes to the 2000 Amendments.

53. See FED. R. EVID. 702 advisory committee’s notes to the 1972 Proposed Rules.

54. See *United States v. Downing*, 753 F.2d 1224, 1229 (3rd Cir. 1985) (quoting S. SALTZBURG & K. REDDEN, FEDERAL RULES OF EVIDENCE MANUAL 451 (3d ed. 1982)).

55. See Giannelli, *supra* note 21, at 1229.

56. *Downing*, 753 F.2d at 1234 (reviewing writings of Saltzburg & Redden, Louisell & Mueller, and Giannelli).

statutory construction that holds a statute “in derogation of the common law . . . is to be strictly considered.”⁵⁷ Some federal, and many state, courts thus continued to apply the *Frye* general acceptance test long after the adoption of the Federal Rules.⁵⁸

Others, however, saw the Federal Rules’ silence on *Frye* as being “tantamount to an abandonment of the general acceptance standard.”⁵⁹ Indeed, the overall structure of the Rules supports this contention. Rule 702 requires only that expert testimony be helpful.⁶⁰ Rule 402 requires that any relevant evidence be admitted,⁶¹ while Rule 401 defines relevance as having any tendency to make a fact more or less likely.⁶² Nowhere do the rules mention general acceptance.

The lower courts struggled with this lack of clarity. In *United States v. Downing*, the Third Circuit wrote that the “status of the *Frye* test under Rule 702 is somewhat uncertain.”⁶³ Nonetheless, the court went on to say that “a particular degree of acceptance of a scientific technique within the scientific community is neither a necessary nor a sufficient condition for admissibility.”⁶⁴ It is difficult to conceive of a more clear rejection of the *Frye* test. The court set out a three factor test of its own, under which a district court would evaluate:

- (1) the soundness and reliability of the process or technique used in generating the evidence, (2) the possibility that admitting the evidence would overwhelm, confuse, or mislead the jury, and (3) the proffered connection between the scientific research or test result to be presented, and particular disputed factual issues in the case.⁶⁵

This so called “*Downing* test” became the Third Circuit rule in criminal and civil cases.⁶⁶ Other courts attempted to harmonize the Federal Rules and the *Frye* test.⁶⁷

57. *Devers v. Scranton City*, 161 A. 540, 542 (Pa. 1932).

58. *See, e.g.*, *United States v. Kozminski*, 821 F.2d 1186, 1199 (6th Cir. 1987) (Krupansky, J. concurring); *United States v. Solomon*, 753 F.2d 1522, 1526 (9th Cir. 1985) (“Evidence based on a novel scientific technique is admissible if it is generally accepted as a reliable technique among the scientific community.”).

59. *Downing*, 753 F.2d at 1234 (quoting J. WEINSTEIN & M. BERGER, WEINSTEIN’S EVIDENCE).

60. FED. R. EVID. 702.

61. FED. R. EVID. 401.

62. FED. R. EVID. 402.

63. *Downing*, 753 F.2d at 1232.

64. *Id.* at 1237.

65. *Id.*

66. *See In re Paoli R.R. Yard PCB Litig.*, 916 F.2d 829 (3d Cir. 1990) (holding that district court erred in application of *Downing* test in toxic tort case); *United States v. Stevens*, 935 F.2d 1380 (3d Cir. 1991) (holding that district court erred under *Downing* test in excluding defendant’s proffered expert testimony).

67. *See* Amy T. Schutz, Note, *The New Gatekeepers: Judging Scientific Evidence in a Post-Frye World*, 72 N.C.L. REV. 1060, 1075-76 (1994).

In addition to establishing admissibility requirements for scientific evidence, the Federal Rules of Evidence also prescribe minimal restrictions on the qualifications for expert witnesses. Rule 702 states only that “[a] witness . . . is qualified as an expert by knowledge, skill, experience, training, or education.”⁶⁸ The 1972 Advisory Committee Notes to the proposed rules indicated that the intention of the Rules was to provide broad latitude in the definition of an expert: “[W]ithin the scope of the rule are not only experts in the strictest sense of the word, e.g., physicians, physicists, and architects, but also the large group sometimes called ‘skilled’ witnesses, such as bankers or landowners testifying to land values.”⁶⁹ Under the Federal Rules, courts thus have broad latitude to decide whether an expert’s particular knowledge, skill, experience, training, or education qualifies them to testify on the particular matter at issue as an expert.⁷⁰

Rule 702 does not specify how closely a proffered expert’s knowledge, skill, experience, training, or education must relate to the issues on which they would testify. Courts applying Rule 702 generally held a permissive attitude toward admitting experts who possessed expertise somewhat related to the point on which they were proffered to testify; they did not require an expert to possess expertise precisely on point.⁷¹ “[Expert witnesses were] not strictly confined to [their] area[s] of practice, but may testify concerning related applications; a lack of specialization affects the weight of the opinion, not its admissibility.”⁷² This “[l]iberality and flexibility in evaluating qualifications”⁷³ was based on an assumption that someone who has related expertise “can . . . make himself very much an expert in the particular [area].”⁷⁴ Thus, in terms of both experts’ qualifications and the content of their testimony, the Federal Rules establish a strong presumption in favor of admissibility.

b. The Bias in Favor of Admissibility and the Junk Science Hysteria

Concerns over the impact of the permissive approach that the Federal Rules of Evidence take to the admissibility of expert testimony formed

68. FED. R. EVID. 702.

69. FED. R. EVID. Rule 702 advisory committee’s note.

70. *See, e.g.*, *Lappe v. Am. Honda Motor Co.*, 857 F. Supp. 222, 225 (N.D.N.Y. 1994), *aff’d sub nom.* *Lappe v. Honda Motor Co. Ltd. of Japan*, 101 F.3d 682 (2d Cir. 1996).

71. *Id.* at 226. *See also id.* at 227 n.4 (citing *Gardner v. General Motors Corp.*, 507 F.2d 525 (10th Cir.1974); *Wylie v. Ford Motor Co.*, 536 F.2d 306 (10th Cir.1976); *Hammond v. International Harvester Co.*, 691 F.2d 646 (3d Cir.1982); *Dixon v. International Harvester Co.*, 754 F.2d 573 (5th Cir.1985)).

72. *Id.* at 226.

73. *Id.* at 227.

74. *Id.* at 226-27 (quoting *United States v. Viglia*, 549 F.2d 335 (5th Cir.1977), *cert. denied*, 434 U.S. 834 (1977)).

part of a crescendo of dissatisfaction with the court system during the 1980s and early 1990s. Over the thirty year period between 1960 and 1988, the caseload in the federal district courts trebled, and the ratio of pending to terminated cases rose from 0.75 to 0.97, indicating that the district courts were falling behind in their workload.⁷⁵ Some academic commentators claimed that “[t]his tremendous surge in civil litigation has taxed our court system,” resulting in “higher product prices, higher legal fees and costs, less access to the system for average Americans, and a less dynamic economy.”⁷⁶ In 1991, Vice President Dan Quayle delivered the report of the President’s Council on Competitiveness (“Quayle Report”), which claimed that the increase in litigation had resulted in direct annual costs of \$80 billion, and over \$300 billion when indirect costs were considered.⁷⁷ The Quayle Report noted that in 1989, 18 million new civil cases had been filed in state and federal courts.⁷⁸

The Quayle Report sharply criticized the use of expert witnesses on several counts. The use of contingency fees “easily turns too many expert witnesses into ‘hired guns.’”⁷⁹ And the report claimed that the ease with which expert testimony was admitted resulted in “the ability to fashion almost any opinion into expert testimony.”⁸⁰ Criticism by commentators of the use of expert witnesses extended to criminal cases as well. Professor Giannelli deplored the use of “junk science” in criminal cases, discussing, among other things, the widely-discredited testimony of a psychiatrist in *Barefoot v. Estelle* that there was a “one hundred percent and absolute” chance that the convicted murder would kill again.⁸¹

In the popular media, the “junk science” trope was widely broadcast. Newspapers retold stories of absurd jury awards to undeserving plaintiffs, such as the Philadelphia psychic who was awarded nearly one million dollars because a CT scan had deprived her of her psychic powers.⁸² The stories rarely were thorough—in the psychic’s case, the reported stories rarely noted that the plaintiff had informed the medical staff that she was allergic to intravenous contrast, which the staff had administered

75. THE FEDERAL COURTS STUDY COMMITTEE, REPORT OF THE FEDERAL COURTS STUDY COMMITTEE 5-6 (1990).

76. Gregory Brian Butler & Brian David Miller, *Fiddling While Rome Burns: A Response to Dr. Hensler*, 75 JUDICATURE 251 (1992).

77. PRESIDENT’S COUNCIL ON COMPETITIVENESS, AGENDA FOR CIVIL JUSTICE REFORM IN AMERICA 1 (1991).

78. *Id.*

79. *Id.* at 5.

80. *Id.*

81. Paul C. Giannelli, “Junk Science”: *The Criminal Cases*, 84 J. CRIM. L. & CRIMINOLOGY 105, 113 (1993) (quoting *Barefoot v. Estelle*, 463 U.S. 880 (1983)).

82. See, e.g., *Says Her Powers Vanished: ‘Psychic’ Awarded \$988,000 in Hospital CAT Scan Lawsuit*, LA TIMES (Mar. 30, 1986), http://articles.latimes.com/1986-03-30/news/mn-1672_1_allergic-reaction [https://perma.cc/6TAW-9NGT].

nonetheless, provoking a severe reaction⁸³—and judicial reductions in the awards often went unreported. Magazines recounted the bizarre “specialties” that experts claimed to possess, including “sportsology,” which had enabled one expert to opine on the millions of dollars in revenues a collegiate athlete lost when he was cut from his team for academic underperformance.⁸⁴ Perhaps no one was more influential than Peter Huber, whose 1991 book, *Galileo’s Revenge: Junk Science in the Courtroom*,⁸⁵ popularized the term “junk science,” and served as the basis for many of the Quayle Report’s assertions.

Within the academic world, however, scholars debated the extent, and even the existence, of the “litigation explosion.” Professor Deborah R. Hensler pointed out the weaknesses of the Quayle Report’s empirical bases. According to Hensler, only two out of every hundred people injured by accidents filed lawsuits to recover damages.⁸⁶ Further, the 18 million new civil cases that the Report cited included “millions of routine cases citizens are required to file under certain circumstances—for example, divorce cases and probate cases,” undercutting the argument that litigation was driving costs up and driving down American competitiveness.⁸⁷

In spite of significant pushback within the legal academy and the bar against the Quayle Report’s conclusion that there was an ongoing litigation explosion, the overall tenor of academic commentary on the admissibility of expert witness testimony and scientific evidence revealed a pervasive sense that the balance of power was tilted far too far in the plaintiffs’ experts’ favor.⁸⁸ Some judges, too, felt that the admissibility

83. *Haimes v. Temple Univ. Hosp.*, 39 Pa. D. & C. 3d 381, 384-85 (1986).

84. Walter Olson, *The Case Against Expert Witnesses*, FORTUNE, Sept. 25, 1989, https://money.cnn.com/magazines/fortune/fortune_archive/1989/09/25/72505/ [<https://perma.cc/UW68-4QUE>].

85. PETER W. HUBER, *GALILEO’S REVENGE: JUNK SCIENCE IN THE COURTROOM* (1991).

86. Deborah R. Hensler, *Taking Aim at the American Legal System: The Council on Competitiveness’s Agenda for Legal Reform*, 75 JUDICATURE 244, 245-46 (1992).

87. *Id.* at 245.

88. See Michael D. Green, *Expert Witnesses and Sufficiency of Evidence in Toxic Substances Litigation: The Legacy of Agent Orange and Bendectin Litigation*, 86 NW. U. L. REV. 643, 697 (1992) (“[T]hese observations do not augur well for the current concerns about junk science or cleaning up the expert witness business.”); John F. Baughman, *Galileo’s Revenge: Junk Science in the Courtroom*, 90 MICH. L. REV. 1614 (1992) (“[T]he evils of allowing junk science to support tort verdicts are several.”); Peter Huber, *Junk Science in the Courtroom*, 26 VAL. U. L. REV. 723, 742 (1992) (“An accretion of cranks in court follows inevitably from the great paradox of modern liability science: in attempting to control quackery outside the courtroom, we invite quacks to the witness stand.”). *But see* Jeff L. Lewin, *Calabresi’s Revenge: Junk Science in the Work of Peter Huber*, 21 HOFSTRA L. REV. 183, 189 (1992) (“Huber’s evaluation of the impact of junk science does not measure up to the standards of scientific methodology.”); Kenneth J. Chesebro, *Galileo’s Retort: Peter Huber’s Junk Scholarship*, 42 AM. U. L. REV. 1637, 1639 (1993) (describing the case against the contemporary use of expert witnesses as “a catalog of every conceivable kind of error: data dredging, wishful thinking, truculent dogmatism, and, now and again, outright fraud” (quoting HUBER, *supra* note 85, at 3)).

practices were too lenient. Judge Jack Weinstein complained that “an expert can be found to testify to the truth of almost any factual theory, no matter how frivolous [A]n expert's testimony can be used to obfuscate what would otherwise be a simple case.”⁸⁹ This frustration led one commentator to write that “[t]he liberalization of expert witness testimony inevitably created its own backlash.”⁹⁰ Some judges began applying “stricter standards against evidence from the fringes of the scientific community.”⁹¹ However, it is not clear how extensive this pre-*Daubert* judicial move away from the permissive approach to expert testimony became.⁹²

3. The Rules Under the Daubert Trilogy

In *Daubert v. Merrill-Dow Pharmaceuticals, Inc.*, the Supreme Court held that the Federal Rules of Evidence had superseded *Frye's* general acceptance test.⁹³ In place of general acceptance, the Court held that the Federal Rules now required trial court judges to “ensure that any and all scientific testimony or evidence admitted is not only relevant, but reliable.”⁹⁴

The Court located the requirement for relevance in Rule 702, which established that scientific evidence was admissible if it would “assist the trier of fact to understand the evidence or to determine a fact in issue.”⁹⁵ This is the so-called “fit” requirement—that is, the testimony had to be sufficiently “tied to” the case that it would assist the jury.⁹⁶

The Court located the reliability requirement in the clause of Rule 702 that referred to “*scientific, technical, or other specialized knowledge.*”⁹⁷ In the Court’s view, “scientific” meant that the testimony was grounded “in the methods and procedures of science”⁹⁸ and “knowledge” meant “any body of known facts or to any body of ideas inferred from such facts or accepted as truths on good grounds.”⁹⁹ The key here is that “good grounds” meant knowledge “derived by the scientific method.”¹⁰⁰ But

89. Green, *supra* note 88, at 670 (quoting Jack B. Weinstein, *Improving Expert Testimony*, 20 U. RICH. L. REV. 473, 475-76 (1986)).

90. *Id.* at 669.

91. HUBER, *supra* note 85, at 736.

92. *See id.* at 737 (“Judicial insistence on peer review is still by no means unanimous, however.”).

93. *Daubert v. Merrill-Dow Pharmaceuticals, Inc.*, 509 U.S. 579 (1993).

94. *Id.* at 589.

95. *Id.* (quoting FED. R. EVID. 702).

96. *Id.* at 591.

97. *Id.* at 589 (emphasis added).

98. *Id.* at 590.

99. *Id.*

100. *Id.*

how were judges, untrained in the sciences, to determine whether a technique or theory is scientific knowledge? The Court posited four factors for consideration: (1) whether the theory or technique can be and has been tested, (2) whether it has been subjected to peer review and publication, (3) whether it has a known error rate and standards for “controlling the technique’s operation,” and (4) whether it is generally accepted.¹⁰¹

Although *Daubert* is considered to have revolutionized the rules of admissibility for scientific evidence and expert testimony,¹⁰² there is very little that is revolutionary about the four “*Daubert* factors.” That the reliability factors include *Frye*’s general acceptance test is only part of the story; the other three factors are merely essential conditions for general acceptance. Very few techniques or theories will be widely accepted by scientists if they have not been tested; testing involves evaluating the error rate and controllability of the technique; and the medium through which scientists communicate is publication in peer reviewed journals. So, in a very real sense, *Daubert* simply broke the general acceptance test into its subatomic particles and afforded trial judges greater flexibility. On the one hand, judges could find that evidence was admissible even though it had not yet achieved general acceptance, as where a technique or theory had been tested and published through the peer review process. On the other hand, judges could exclude evidence that was generally accepted if they determined that the other factors were absent. Of course, as just noted, if a theory or technique had been generally accepted, it was a practical necessity that scientists in the field had determined that the other factors were present. To contradict this determination of scientists in the field, judges would be, as Justice Rehnquist warned in his partial dissent, playing the role of “amateur scientists.”¹⁰³

It is this last possibility, which arises from the shift in judicial attitude demanded by *Daubert*, that is generally accepted to be the truly revolutionary outcome of the decision.¹⁰⁴ The conventional narrative is that before *Daubert*, judges took a deferential approach to scientific evidence and experts; that is, the scientific community, not the trial judge, was empowered to determine what evidence was admissible.¹⁰⁵ After *Daubert*, it was the trial judge, not the scientific community, who would

101. *Id.* at 593-94.

102. See e.g., Marc S. Klein, *Daubert: Worldwide Judicial Management of Humanity’s Specialized Knowledge*, 30 U.C. DAVIS L. REV. 1229, 1230 (1997) (describing *Daubert*’s “revolutionary effect”).

103. *Daubert*, 509 U.S. at 600-01 (Rehnquist, C.J., partial dissent).

104. Margaret A. Berger, *Upsetting the Balance Between Adverse Interests: The Impact of the Supreme Court’s Trilogy on Expert Testimony in Toxic Tort Litigation*, 64 Law and Contemp. Probs. 289, 293 (2001) (noting that *Daubert* thrust judges into the position of defining scientific knowledge).

105. Ronald J. Allen & Esfand Nafisi, *Daubert and Its Discontents*, 76 BROOK. L. REV. 131, 134 (2010).

determine whether evidence was reliable and relevant, and hence, admissible. But this story does not hold true for all areas of the law; ironically, it does not hold true for the very area of the law out of which *Daubert* arose: toxic tort litigation.¹⁰⁶

The Court continued to develop its jurisprudence concerning expert witness testimony and scientific evidence in its 1997 *General Electric Co. v. Joiner* decision.¹⁰⁷ In *Joiner*, the Court established two important points. First, the Court assigned trial judges an additional responsibility of determining whether the expert's inferences, even though drawn on reliable and relevant knowledge that passes the *Daubert* test, were themselves products of a reliable scientific method. At trial, the district court judge had excluded *Joiner's* proffered expert witness testimony that polychlorinated biphenyls ("PCBs") can cause small cell lung cancer.¹⁰⁸ Finding that the expert opinion was inadmissible because it was based on animal studies, and on human epidemiological data that failed to demonstrate a correlation between exposure and lung cancer, the Supreme Court held that "a court may conclude that there is simply too great an *analytical gap* between the data and the opinion proffered."¹⁰⁹ *Joiner* thus appears to have tipped the balance struck by the rules and the Court's decisions between judges and experts farther toward the judges, empowering judges to exclude testimony when, in their view, the proposed testimony represented nothing more than the unfounded belief of the expert.¹¹⁰

Just how far the Court in *Joiner* intended to tip the balance was made clear in Justice Breyer's concurring opinion. Justice Breyer noted that the gatekeeper function assigned to judges by *Daubert* would "sometimes ask judges to make subtle and sophisticated determinations about scientific methodology and its relation to the conclusions an expert witness seeks to offer."¹¹¹ To Justice Breyer, judges were clearly empowered to make scientific, as well as legal, determinations.¹¹² Recognizing that judges are not scientists, Justice Breyer urged that in complex cases, judges should rely on court-appointed, neutral panels of experts to aid in making the admissibility decision.¹¹³

106. Klein, *supra* note 102, at 1231 ("[By the time of the *Daubert* decision] *Frye* was already dead on arrival at the Supreme Court. Thus, for the most part, *Daubert* simply represents the official death certificate."); see also *infra* Part III.C.

107. 522 U.S. 136 (1997).

108. *Id.* at 140.

109. *Id.* at 146 (citing *Turpin v. Merrell Dow Pharm., Inc.*, 959 F.2d 1349, 1360 (6th Cir. 1992)) (emphasis added).

110. *Id.*

111. *Id.* at 147 (Breyer, J., concurring).

112. *Id.*

113. *Id.* at 149-50.

The second point that *Joiner* established—on the question that was actually before the Court on appeal¹¹⁴—was that courts of appeals were to apply an abuse of discretion standard when they reviewed trial courts' decisions on the admissibility of expert witness testimony.¹¹⁵ This point is important because it tips the balance even more toward trial judges. Practically, the abuse of discretion standard means that the trial court judge's decision on admissibility will stand so long as he or she did not issue a manifestly erroneous holding, which, as previously stated, is extremely broad.¹¹⁶ After *Joiner*, evidentiary rulings that are appealed are unlikely to be reversed under the abuse of discretion standard. Plaintiffs thus get only one chance to present expert testimony.

Joiner's analytical gap analysis is problematic in two interrelated ways. First, asking trial judges to determine whether the gap between the data and the expert's opinion is too wide risks substituting the trial judge's opinion for that of the expert. Except in very rare instances, there will not be a single definitive study that directly establishes a causal relationship between exposure to a substance and a physical harm;¹¹⁷ thus, in nearly all cases, experts will need to draw an inference of causation. How far such an inference may reach is an integral part of scientific expertise as training in the sciences includes training in how much may be inferred from the data, and the peer review process often carefully vets the extent of authors' inferences based on the data they have generated.¹¹⁸ Thus, *Joiner* magnifies the concern raised in Justice Rehnquist's *Daubert* dissent that judges were now empowered to substitute their own *ipse dixit* as to which inferences may be drawn from the scientific evidence in place of those of the experts in the field.

This raises the second, related, problem created by *Joiner*: the Court provided little guidance as to *how* judges are to decide when the analytic gap is so great as to indicate that nothing but “the *ipse dixit* of the

114. Justice Stevens dissented from the part of the *Joiner* holding in which the Court created the analytical gap analysis, arguing that the Court had an insufficient record on which to review the trial court's decision. *Id.* at 150 (Stevens, J., dissenting).

115. *Id.* at 141 (majority opinion).

116. *Id.* at 141-42 (noting that under the abuse of discretion standard a trial court's ruling will not be overturned “unless the ruling is manifestly erroneous”).

117. Indeed, a skillful opponent can ensure that *no* study is sufficiently on point. For example, if a substance has been shown to cause cancer in humans, that does not prove that it can cause a specific cancer, such as leukemia. If it has been shown to cause leukemia, that does not prove it can cause AML, a specific type of leukemia. If it has been shown to cause AML, that does not prove that it can cause APL, a subtype of AML, and so on, potentially ad infinitum. This line of argument that causation had not been proved was used by the defense in *Milward v. Acuity Specialty Prods. Grp., Inc.*, 664 F. Supp. 2d 137 (D. Mass. 2009), and was critiqued by the plaintiff's counsel in Steve B. Jensen, *Sometimes Doubt Doesn't Sell: A Plaintiff's Lawyer's Perspective on Milward v. Acuity Products*, 3 WAKE FOREST J.L. & POLICY 177, 184-86 (2013).

118. *Id.* at 180.

expert”¹¹⁹ links the data and the opinion. In *Kumho Tire Co., Ltd. v. Carmichael*, the Court noted that “the law grants a district court the same broad latitude when it decides how to determine reliability as it enjoys in respect to its ultimate reliability determination.”¹²⁰ Courts may apply the “*Daubert* factors”¹²¹ to the methodology used by the expert to reach a conclusion;¹²² they “must make certain that an expert . . . employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field.”¹²³ And, courts can, under the Federal Rules of Evidence, appoint experts to serve on behalf of the court.¹²⁴ Beyond this, the Supreme Court appears to have offered very little guidance to trial judges in determining the admissibility of the inferences and conclusions drawn by an expert witness.¹²⁵

The Court completed its “revolution” with its 1999 *Kumho* decision.¹²⁶ *Kumho Tire* manufactured a tire that had experienced a blow out, resulting in a fatal traffic accident.¹²⁷ The district court had ruled that the plaintiff’s expert, whose testimony was based on technical knowledge—as opposed to scientific knowledge, about which both *Daubert* and *Joiner* were concerned—was inadmissible under a *Daubert* analysis.¹²⁸ The Supreme Court upheld the district court’s decision, ruling that the *Daubert* gatekeeper function applied to testimony based on technical or other specialized knowledge as well as on scientific knowledge.¹²⁹ Thus, the *Daubert* framework applies to all of the forms of expert witness testimony covered by Rule 702.

The Court suggested a second important point in *Kumho*: courts might

119. 522 U.S. at 146.

120. *Kumho Tire Co., Ltd. v. Carmichael*, 526 U.S. 137, 142 (1999) (emphasis omitted).

121. These factors are (1) whether the theory is falsifiable, (2) whether it has been submitted for peer review and publication, (3) whether a method has a known error rate and means of controlling the technique’s operation, and (4) whether the theory or technique is generally accepted. *Daubert v. Merrell-Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 593-94 (1993).

122. 526 U.S. at 156.

123. *Id.* at 152.

124. FED. R. EVID. 706; 522 U.S. at 149-50 (Breyer, J, concurring).

125. Two searches using WestlawNext were performed. The first filtered all Westlaw Citing References to *Joiner* so as to return only Supreme Court cases. This returned eight cases, including *Kumho*. Of the remaining seven, only two, *Weisgram v. Marley Co.*, 528 U.S. 440 (2000), and *Bragdon v. Abbott*, 524 U.S. 624 (1998), discussed *Joiner*, but not the issue of how judges should decide when an expert’s conclusions are unreliable. The second search involved searching for all mentions of “General Elec. Co. v. Joiner,” restricting the results to cases in the U.S. Supreme Court. This produced one additional case, *Sprint/United Management Co. v. Mendelsohn*, 522 U.S. 379 (2008), in which the Court held that the Tenth Circuit had erred by not applying the abuse of discretion standard in its review of the district court’s evidentiary rulings.

126. 526 U.S. 137.

127. *Id.* at 142.

128. *Id.* at 145.

129. *Id.* at 147-48.

narrow their construction of “qualified” when deciding whether an expert would be allowed to testify.¹³⁰ Under the Federal Rules, there are no specific standards establishing when a proposed expert may be considered a qualified expert. In *Kumho*, Justice Breyer’s majority opinion noted that “[t]he trial court had to decide whether this particular expert had sufficient *specialized knowledge* to assist the jurors.”¹³¹ The Court endorsed the trial court’s close examination of the expert’s qualifications.¹³² Thus, *Kumho* suggested another way in which judges might act as gatekeepers: by rejecting the testimony of a proffered expert whose expertise is not highly specialized in the area in which the expert would testify. The Court did not address this in detail and did not discuss the apparent conflict of this narrowing of the definition of a qualified expert with the permissive language of Rule 702.

B. Scholars’ Views of How Judges Are Using Their New Powers

Legal scholars have criticized judges’ handling of scientific evidence and expert witnesses in the post-*Daubert* era. This criticism falls into two general categories. In one category are claims that judges are rejecting evidence based on misunderstandings of the statistical analyses contained in scientific studies. In the other category are claims that judges are rejecting evidence in order to arrive at certain case outcomes. These claims posit that judges are knowingly rejecting valid science.

Numerous legal scholars have criticized the approaches that many trial courts, post-*Daubert*, have taken to the scientific evidence on which expert witnesses would rely.¹³³ As one line of criticism stated, judges “subject each item of expert proof proffered by plaintiffs to substantive causation law scrutiny, to see if it, standing alone, would prove both general and specific causation.”¹³⁴ In essence, this criticism holds that judges take the approach that only if an item, by itself, can prove causation may it be admitted or serve as a basis for an expert’s opinion. This “atomist” approach conflicts with the way in which scientists draw

130. *Id.* at 146-47 (noting that the Court granted certiorari “in light of uncertainty among the lower courts” over the scope of qualifying knowledge).

131. *Id.* at 156 (emphasis added).

132. *Id.* at 153 (noting that the trial court examined the expert’s education, work experience, and prior recognition as an expert witness in other cases).

133. See, e.g., Lucinda M. Finley, *Guarding the Gate to the Courthouse: How Trial Judges Are Using Their Evidentiary Screening Role To Remake Tort Causation Rules*, 49 DEPAUL L. REV. 335, 336-37 (1999); Susan Haack, *Proving Causation: The Holism of Warrant and the Atomism of Daubert*, 4 J. HEALTH & BIOMEDICAL L. 253 (2008); Margaret A. Berger, *Upsetting the Balance Between Adverse Interests: The Impact of the Supreme Court’s Trilogy on Expert Testimony in Toxic Tort Litigation*, 64-SUM LAW & CONTEMP. PROBS. 289, 290 (2010); Steve C. Gold, *A Fitting Vision of Science for the Courtroom*, 3 WAKE FOREST J. L. & POL’Y 1, 4-5 (2013).

134. Finley, *supra* note 133, at 336-37.

inferences. Scientists typically use a “weight of the evidence” approach in which they examine all of the available data, weighing each piece to determine its relative value, and ultimately arrive at a conclusion.¹³⁵ In effect, judges using an atomist approach push the admissibility decision back one step, making an admissibility decision on each piece of evidence instead of allowing the expert to determine the weight of each piece in order to draw an over-arching conclusion. Because this approach excludes vastly more pieces of scientific evidence than the weight of the evidence approach, fewer experts’ opinions have sufficient support to be admissible.

Scholars have criticized several other ways that judges look at scientific evidence. These scholars have pointed to an inappropriate conflation of the burden of proof—the preponderance of the evidence—with the admissibility standard of relevance.¹³⁶ Many courts have ruled that epidemiologic evidence, which shows how much more frequently people exposed to a substance develop a disease than those who were not exposed (the “relative risk”), is admissible only if those who were exposed are at least twice as likely to develop the disease.¹³⁷ In reality, a less-than twofold increase in relative risk still may indicate that exposure is associated with a higher likelihood of disease.¹³⁸ Whether such a finding deserves weight depends on the statistical significance, not on the magnitude of the increase. Judges who apply a “relative risk 2.0” standard do so because this seems to comport with the burden of proof in a civil action: only when the relative risk is greater than 2.0 is it more likely that any person who developed a disease did so because they were exposed to the allegedly harmful substance. But nothing in the Federal Rules or the *Daubert* trilogy indicates the preponderance of the evidence standard was to be applied to individual pieces of scientific evidence. Thus, judges who apply a relative risk 2.0 standard exclude valid scientific evidence on which experts would rely outside the courtroom.

135. *Id.* at 337. This point was clearly made by Sir Austin Bradford Hill, who is regarded as the father of modern epidemiology:

What I do not believe—and this has been suggested—is that we can usefully lay down some hard-and-fast rules of evidence that must be observed before we accept cause and effect. None of my nine viewpoints can bring indisputable evidence for or against the cause and effect hypothesis and none can be required as a *sine qua non*. What they can do, with greater or less strength, is to help us to make up our minds on the fundamental question – is there any other way of explaining the set of facts before us, is there any other answer equally, or more, likely than cause and effect?

Austin Bradford Hill, *The Environment and Disease: Association or Causation?* 58 PROC. OF THE ROYAL SOC’Y OF MED. 295, 299 (1965).

136. Finley, *supra* note 133, at 349.

137. *See, e.g.*, In re Joint Eastern & Southern Dist. Asbestos Litigation, 827 F. Supp. 1014, 1030 (S.D.N.Y. 1995) (“[E]pidemiological evidence regarding the relationship between exposure to *c* and the development of *d* may fall short of the 2.0 threshold of statistical significance.”).

138. Finley, *supra* note 133, at 348 & n.49.

Other facets that have been criticized include the presumption that animal studies are far less relevant than epidemiologic studies,¹³⁹ a requirement that the plaintiff precisely quantify their exposure to the allegedly harmful substance,¹⁴⁰ and the misuse of sets of factors to be weighed as lists of individual requirements that must be satisfied.¹⁴¹

Many legal scholars have gone beyond mere criticism of judges' handling of the scientific aspects of the scientific evidence and the expert witnesses and have looked at the effects of the admissibility decisions judges have made since *Daubert*. Professor Lucinda Finley has argued that by applying a stringent set of requirements—relative risk 2.0 and the atomist approach included—judges “have been making profoundly normative judgments about the social allocation of risk and who should bear the burden of scientific uncertainty or controversy.”¹⁴² Finley claims that the effect of placing a higher burden on plaintiffs is to have moved away from using tort litigation to implement “consumer protective social policies.”¹⁴³ She sees this as having multiple social implications, including having a disparate impact on groups whose health problems are not robustly studied, thereby reducing the ability of litigation to spur further scientific research and bringing a “premature sense of scientific closure” to issues that merit further study.¹⁴⁴ Professor Margaret Berger has been even more blunt, stating that “according to some observers . . . toxic tort law is being reformulated in the federal courts to the advantage of defendants.”¹⁴⁵

Implicit in these claims is that district court judges are acting on the basis of policy or ideological positions. A substantial body of scholarship suggests that partisanship accounts for a significant portion of judicial decision making.¹⁴⁶ However, most of this work has focused on the appellate courts, with relatively little attention devoted to the district courts. Studies of federal district court judges do show some influence of partisanship, typically seen as a judge's known party affiliation or based on the party to which the President who appointed the judge belonged.¹⁴⁷

139. See Berger, *supra* note 133 at 303.

140. *Id.* at 306-07.

141. See Haack, *supra* note 133 at 274-81.

142. Finley, *supra* note 133 at 335-36.

143. *Id.* at 336.

144. *Id.* at 338.

145. Berger, *supra* note 133 at 290.

146. See, e.g., CASS R. SUNSTEIN ET AL., ARE JUDGES POLITICAL?: AN EMPIRICAL ANALYSIS OF THE FEDERAL JUDICIARY (2006); Randall D. Lloyd, *Separating Partisanship from Party in Judicial Research: Reapportionment in the U.S. District Courts*, 89 THE AM. POL. SCI. REV. 413 (1995).

147. See Lloyd, *supra* note 146 at 417-18 (finding that district court judges hold reapportionment plans drawn by state legislatures controlled by the opposite party to a higher standard than plans drawn by nonpartisan legislatures or those controlled by the judge's party).

These studies have tended to focus on ideologically-charged issues such as legislative reapportionment. The *Daubert* trilogy was created during an ideologically-charged moment as the perception that runaway litigation and junk science were harming U.S. competitiveness was much discussed in the courts, in politics, and in the popular media. Thus, underlying claims by scholars such as Finley and Berger are an assumption that the judicial imposition of stringent admissibility standards is a manifestation of a political or ideological program of tilting the litigation playing field in favor of corporate defendants. However, no scholarly work has attempted to link individual judges' political and ideological leanings to their treatment of scientific evidence and expert testimony, nor to changes in judicial behavior after the "*Daubert* revolution."

III. TOXIC TORT LITIGATION: THE CHOKE POINT OF GENERAL CAUSATION

Toxic tort litigation offers an avenue by which to understand how much of a change the *Daubert* trilogy has had on judicial evidentiary practices. Toxic tort litigation involves a claim of injury resulting from exposure to an allegedly harmful substance.¹⁴⁸ The harms may be acute, as in the case of illness arising immediately upon one's inhalation of a noxious gas,¹⁴⁹ but most commonly these actions concern harms that manifest years after the initial exposure such as various cancers. Tort actions have deep, historical roots, but toxic tort litigation is a relatively recent phenomenon. Although isolated toxic tort cases have been brought by plaintiffs who became acutely ill upon exposure for over 150 years, the development of the scientific tools and techniques that allow inferences of causation to be drawn has been the engine that has driven modern toxic tort litigation.

During the nineteenth and early twentieth centuries, a multidisciplinary group of scientists developed epidemiology, the application of statistical analysis to the occurrence of "disease or injury in human populations."¹⁵⁰ Epidemiologic studies can demonstrate a correlation showing that people exposed to substance *X* are more likely to develop disease *Y* than are

148. L. NEAL ELLIS JR., INTRODUCTION, TOXIC TORT LITIGATION 3 (Arthur F. Foerster & Christine Gregorski Rolph eds., 2d ed. 2013). "Substance" is typically defined capaciously, and so toxic tort cases can involve X-rays, nuclear radiation, and other forms of energy. *See, e.g.*, *Cano v. Everest Minerals Corp.*, 362 F. Supp. 2d 814 (W.D. Tex. 2005) (toxic tort case involving claim that cancers resulted from exposure to ionizing radiation).

149. *See* *Hunt v. Lowell Gas Light Co.*, 90 Mass. 169 (1864) (alleging illness arising from defendant gas company's negligent release of gas).

150. FEDERAL JUDICIAL CENTER, NATIONAL RESEARCH COUNCIL, REFERENCE MANUAL ON SCIENTIFIC EVIDENCE 286 (3d ed. 2011).

people not exposed to X .¹⁵¹ The first major triumph of modern epidemiologists was the correlation of cigarette smoking with lung cancer, which was demonstrated by Richard Doll and Austin Bradford Hill in a study published in 1950.¹⁵²

But while epidemiologic studies can demonstrate a correlation, they cannot prove that X actually causes Y , because there may be other, unidentified factors that come along with X that are the true cause of Y . Because of this, scientists have found it difficult to agree on how to infer causation. This is especially important in the commonly litigated situation where a substance allegedly causes a disease, but it takes years for the disease to develop. Some scientists, including Bradford Hill, argued in favor of using a “weight of the evidence” approach in this situation, examining all of the available evidence but holding no single part to be a necessary condition to establish causation.¹⁵³ Others consider the nine factors that Bradford Hill examined to be a checklist, with certain elements—typically those looking at epidemiologic data to assess the strength of the correlation between X and Y —to be necessary conditions, without which one should not infer causation.¹⁵⁴

In the 1950s and 1960s, scientists developed additional tools that have become important for modern toxic tort litigation. Animal studies, chemistry, biochemistry, and toxicology all developed into techniques that permitted scientists to construct models that provide mechanistic explanations of how some substances could harm humans.¹⁵⁵

Plaintiffs in a toxic tort case must prove that they were exposed to the allegedly harmful substance, that they suffered harm as the result of the exposure, and, most importantly, that there is a causal relationship between the exposure and the harm.¹⁵⁶ Courts generally divide causation into two distinct elements: general causation (i.e., the capacity of a substance to cause the alleged harm in humans) and specific causation (i.e., the fact that the substance caused the harm in the plaintiff).¹⁵⁷ In essence, the general causation inquiry asks “can it?” while the specific

151. *Id.* at 551-52.

152. Richard Doll & A. Bradford Hill, *Smoking and Carcinoma of the Lung*, 2(4682) BR. MED. J. 739 (1950).

153. *See* Hill, *supra* note 135. In his Royal Society address, Bradford Hill listed nine types of data that scientists should consider when deciding whether X causes Y ; he insisted that no one type was dispositive.

154. Carl V. Phillips & Karen J. Goodman, *The Missed Lessons of Sir Austin Bradford Hill*, 1 EPIDEMIOLOGIC PERSPECTIVES AND INNOVATION *1, *2, <http://www.epi-perspectives.com/content/1/1/3> (noting that in epidemiology classes, Bradford Hill’s considerations are “[t]ypically presented as a checklist approach to assessing causation”).

155. *See* FEDERAL JUDICIAL CENTER, *supra* note 150, at 604-05.

156. ELLIS, *supra* note 148, at 4.

157. FEDERAL JUDICIAL CENTER, *supra* note 150, at 552; Alan Rudlin, Alexandra Cunningham & Thomas R. Wascom, TOXIC TORT LITIGATION *supra* note 148, at 140-43.

causation inquiry asks “did it?”

The necessity of proving causation creates an early opportunity for a defendant to end a plaintiff’s case through an evidentiary challenge and a motion for summary judgment.¹⁵⁸ If the defendant can show that there is no genuine dispute of material fact concerning its own expert witness’s assertion that the defendant’s substance could not or did not cause the plaintiff’s harm, it is entitled to summary dismissal of the case against it.¹⁵⁹

To prove general and specific causation in toxic tort cases, almost all courts apply a rule that “expert . . . testimony is unquestionably required to assist the jury.”¹⁶⁰ Without a plaintiff’s expert to assert that substance *X* may cause or has caused harm *Y*, a defendant’s motion for summary judgment will almost certainly succeed.¹⁶¹

Under the generally accepted narrative, trial judges, who were once mandated by the *Frye* rule to defer to the scientific community as to whether scientific evidence and expert testimony based on that evidence were admissible, now have the power to terminate cases at the evidentiary challenge stage. Having been liberated by *Daubert*, judges may now determine for themselves when evidence is reliable and relevant science, and when an expert’s inferences attempt to cross too broad an analytical gap in order to exclude a plaintiff’s expert witness testimony. Judges may use the choke point of admissibility of plaintiffs’ causation experts to dispose of cases based on their early perception of the merits of those cases by excluding all of plaintiff’s causation experts. Further, judges may use the choke point of admissibility of plaintiffs’ causation experts to change the law in ways that reflect their own political or ideological preferences.

IV. EMPIRICAL STUDY

Unfortunately, little empirical work has examined the factors that actually drive judges’ admissibility rulings. This Part presents the results of a pilot empirical study of federal district court rulings on challenges to

158. See FED. R. CIV. P. 56.

159. *Id.*

160. *Junk v. Terminix Int’l Co.*, 628 F.3d 439, 450 (8th Cir. 2010 (quoting *Ranes v. Adams Labs., Inc.* 778 N.W.2d 677, 688 (Iowa 2010)); see also *Korte v. Exxonmobil Coal USA, Inc.*, 164 F. App’x 553, 556 (7th Cir. 2006) (“Expert testimony is needed to establish causation in cases alleging an adverse health effect when the ‘medical effects [of exposure to the toxin] are not within the ken of the ordinary person.’”) (quoting *Goffman v. Gross*, 374 F.3d 549, 552 (7th Cir. 1995); *Porter v. Whitehall Labs., Inc.*, 791 F. Supp 1335 (1992)). The Sixth Circuit appears to be the only outlier. See *Gass v. Marriott Hotel Svcs., Inc.*, 558 F.3d 419 (6th Cir. 2009) (“Plaintiffs are not required to produce expert testimony on causation where Defendants have failed to offer scientific evidence regarding the effects of Demand CS or Suspend SC.”).

161. See *Berger*, *supra* note 133, at 290.

plaintiffs' general causation experts, both before and after the Supreme Court decisions that compose the *Daubert* trilogy. The first goal of the study was to determine whether there actually was a “*Daubert* revolution,” by asking whether judges' dispositions of toxic tort cases at the evidentiary challenge stage changed after the trilogy. The second goal of the study was to identify the factors that sway judges' rulings, in particular those rulings that exclude all of a plaintiff's general causation experts, which effectively terminates plaintiffs' cases.

A. *Prior Empirical Work*

A number of recent studies have empirically examined the impact of the *Daubert* trilogy on admissibility decisions. Professors Andrew Jurs and Scott DeVito used a database of several million cases that originated in state courts to examine the rates at which civil defendants removed cases from state to federal courts.¹⁶² By comparing removal rates from state and federal courts before and after each adopted the *Daubert* standard (or retained the *Frye* rule), Jurs and DeVito conclude that civil defendants view *Daubert* as a stricter standard.¹⁶³ Professors David M. Flores, James T. Richardson, and Mara L. Merlino examined 191 cases in the District of South Carolina; however, only 25 of these cases involved challenges to the admissibility of plaintiff's expert witnesses.¹⁶⁴ They found that post-*Daubert*, it was 4.69 times more likely that a defendant would raise a challenge to a plaintiff's expert witness.¹⁶⁵ Of the 25 cases involving a challenge to a plaintiff's expert, all but six settled, leaving too small a data set to compare outcomes before and after *Daubert*.¹⁶⁶ Neither study specifically examined the actual practices of federal judges. Jurs and DeVito focused mainly on defendants' perceptions of the relative stringency of the *Frye* and *Daubert* standards.¹⁶⁷ Their findings would tend to support the conventional narrative that *Daubert* allowed judges to exclude more expert testimony and scientific evidence.¹⁶⁸ Flores, Richardson, and Merlino's study also suggests that defendants view *Daubert* as having raised the bar to admissibility, since defendants post-

162. Andrew Jurs & Scott DeVito, *The Stricter Standard: An Empirical Assessment of Daubert's Effect on Civil Defendants*, 62 CATH. U. L. REV. 675 (2013).

163. *Id.* at 680.

164. David M. Flores, James T. Richardson & Mara L. Merlino, *Examining the Effects of the Daubert Trilogy on Expert Evidence Practices in Federal Civil Court: An Empirical Analysis*, 34 S. ILL. U. L.J. 533 (2009).

165. *Id.* at 555.

166. *See id.* at 560.

167. Jurs & DeVito, *supra* note 162, at 678-79.

168. *See id.* at 680 (finding that defendants preferred the *Daubert* rule to the *Frye* rule because it led to more frequent exclusion of plaintiffs' experts).

Daubert have acted as though judges are more likely to exclude plaintiff's expert witnesses.¹⁶⁹

Unlike these prior studies, the current study discussed in this Article directly examines judges' evidentiary rulings. By limiting the examination to toxic tort cases, this study seeks to explore more directly not only the impact that the *Daubert* trilogy has had on the number of cases disposed of at the evidentiary challenge stage, but also the factors that underlie evidentiary decisions both before and after *Daubert*.

B. Methods

Using the West database accessed through WestlawNext and employing a variety of search terms, I assembled a set of fifty-seven federal court decisions¹⁷⁰ on challenges to the admissibility of plaintiffs' general causation experts in toxic tort cases. Twenty-six of these decisions were rendered before *Daubert* and thirty-one after *Daubert*.

The primary endpoint assessed in this study is the outcome of a defendant's challenge to the admissibility of a plaintiff's general causation experts. Court opinions were read carefully to identify the challenged experts who would give testimony on general causation. The number of proffered general causation experts under challenge and the number whose testimony was admitted were recorded. For analytical purposes, a dummy variable was assigned, with "1" assigned to decisions admitting any or all of a plaintiff's general causation experts and "0" assigned to decisions admitting none of a plaintiff's general causation experts. A secondary endpoint was the judge's ruling on a defendant's motion for summary judgment if a motion was made at the same time. Cases summarily disposed of at this stage were coded as "1," while cases that survived were coded as "0."

To determine which, if any, factors were correlated with judges' admissibility rulings, I identified three potential areas of investigation. First was whether judges may have been ruling on the basis of fairly obvious weaknesses in the plaintiffs' cases. If judges are truly as concerned about the impact of expert witness testimony on juries as they have indicated,¹⁷¹ they might use the choke point of general causation as

169. See Flores, Richardson & Merlino, *supra* note 164, at 554 (finding a large increase in defendants' challenges to the admissibility of plaintiffs' experts following the *Daubert* decision).

170. The focus of this Article is on federal district court practices, so the majority of the cases in this database are from the district court level. However, five appellate court decisions are included where no published opinion of the district court ruling under review was available. The criteria for including appellate court opinions were that they contain sufficient information about the underlying case and the district judge's stated reasons for admitting or excluding the proffered expert witness testimony that the factors under evaluation in this study could be coded.

171. See *supra* notes 1-5 and accompanying text.

a means of preventing cases they perceive as meritless from getting to a jury, as previously discussed *supra*. Second was whether judges' own ideology or judicial experience informed their decision making. And third was whether judges might use a plaintiff's ability to present more than a single general causation expert as a proxy for the merits of the case.

To evaluate the potential impact of overt factual weaknesses in plaintiffs' cases on judges' admissibility rulings, I used a set of five factors that I call "adverse case characteristics." These factors were identified based on analysis of the first fifteen cases added to the database (most of which were decided in the post-*Daubert* era). The factors were then applied to the entire database. The five factors are:

(1) Failure of the plaintiff to identify a specific harm that they suffered. Plaintiffs in many toxic tort cases allege specific harms, such as leukemia¹⁷² or asthma.¹⁷³ The allegation of a specific harm permits the finder of fact to evaluate the likelihood that the defendant's substance caused that harm. There are two circumstances where such an evaluation may be so difficult as to make it impossible to meet the legal standard of causation. In the first, a plaintiff may allege multiple vague or nonspecific complaints. One example is the combination of sinusitis, toxic encephalopathy, small airways disease, and dysosmia;¹⁷⁴ another is a combination of fatigue and difficulty concentrating. In the second circumstance, multiple plaintiffs may each allege a single, distinct harm, but each may allege a *different* harm.¹⁷⁵ Not only do these two kinds of cases make legal determinations difficult, but the broad nature of the alleged harms begs the question of whether any single specific cause can ever be identified, and why the particular defendant should be held liable.

(2) Failure of a plaintiff to specify the substance responsible for the alleged injury. Many toxic tort plaintiffs specify that they were harmed by a single substance such as benzene¹⁷⁶ or the insecticide chlorpyrifos.¹⁷⁷ Others, however, allege that they were exposed to unspecified chemicals contained in a toxic soup on the grounds of a rail yard¹⁷⁸ or to multiple chemicals in hundreds of fragrances.¹⁷⁹ Again, the difficulties in establishing causation where the causative

172. *Castellow v. Chevron USA*, 97 F. Supp. 2d 780 (S.D. Tex. 2000).

173. *Cartwright v. Home Depot U.S.A., Inc.*, 936 F. Supp. 900 (M.D. Fla. 1996).

174. *Sanderson v. Int'l Flavors and Fragrances, Inc.*, 950 F. Supp. 981 (C.D. Cal. 1996).

175. *Knight v. Kirby Inland Marine, Inc.*, 363 F. Supp.2d 859 (N.D. Miss. 2005) (one plaintiff alleging Hodgkin's lymphoma and the other bladder cancer).

176. *Baker v. Chevron USA, Inc.*, 680 F. Supp.2d 865 (S.D. Ohio 2010).

177. *Pritchard v. Dow Agro Sciences*, 705 F. Supp.2d 471 (W.D. Pa. 2010).

178. *Wicker v. Consol. Rail Corp.*, 371 F. Supp.2d 702 (W.D. Pa. 2005).

179. *Sanderson*, 950 F. Supp. at 981.

agent is not identified weaken the plaintiff's claims against a specific defendant.

(3) Failure of the plaintiff to state a believable timeline linking the exposure to the defendant's substance to the alleged harm. A plaintiff who alleges that a worsening of his asthma nearly ten years after a single exposure to a chemical,¹⁸⁰ and one who alleges the development of widely-metastatic cancer of a slowly-growing type within a year of exposure,¹⁸¹ both present facially questionable claims.

(4) Failure of a plaintiff to state a believable level of exposure to the allegedly harmful substance. A plaintiff who alleges that her child's congenital defects were due to airborne chemicals released from an explosion at a manufacturing plant several miles away, to which she was exposed for less than ten minutes,¹⁸² presents an immediately questionable claim.

(5) Cases in which other courts have already closely examined either the proposed causal mechanisms to which the plaintiff's experts would testify or the qualifications of the plaintiff's experts and found them to be wanting.

This set of factors was designed to examine whether a "commonsense" look at a plaintiff's case correlates with admissibility rulings. The presence of one or more of these adverse case characteristics was coded as a "1," and the absence of all five was coded as "0."

To evaluate the potential impact of judges' characteristics, the political party to which the president who appointed the judge rendering the opinion, and the number of years that the judge had been on the bench at time he or she rendered the opinion, were recorded. For the analysis of party effect, judges appointed by a Democratic president were coded as "0" and those appointed by a Republican president were coded as "1." For analysis of the impact of duration of federal judicial service, judges with less than ten years on the federal bench were coded as "0" and with ten or more years as "1." This resulted in nearly equal sized cohorts.

To evaluate the potential impact of plaintiffs' experts, the study used a simple quantification: plaintiffs who proffered only a single general causation expert were coded as "0," while those who proffered more than one general causation expert were coded as "1."

Statistical analyses were performed using two freely available online

180. *O'Conner v. Commonwealth Edison Co.*, 807 F. Supp 1376 (C.D. Ill. 1991).

181. *Peterson v. Sealed Air Corp.*, 1990 WL 186779 (N.D. Ill. 1990).

182. *Boyles ex rel. Boyles v. Am. Cyanamid Co.*, 796 F. Supp. 704 (E.D.N.Y. 1992).

calculators.¹⁸³ Simple comparisons were examined using Chi-squared tests, while a multiple logistic regression analysis was used to determine the independence of the effects of the examined factors. A p-value of 0.05 was considered to be significant.

C. Findings

The fifty-seven rulings in the data set were issued between April of 1963, and January of 2014. Other than *Roberts v. Union Carbide Co.*, the 1963 case, the pre-*Daubert* cases were clustered between 1985 and 1992. Thus, this group of cases predominantly represent the era of the Federal Rules of Evidence. The post-*Daubert* cases were weighted slightly toward the past fifteen years. Thus, these cases reflect both the period during which the “*Daubert* trilogy” was being developed and the more recent period during which no major legislative or Supreme Court changes to the admissibility doctrine were made.

The analysis of this data yields three significant findings, two of which challenge the conventional narrative.

First, judges before and after *Daubert* excluded plaintiffs’ expert witness testimony and granted motions for summary judgment at nearly the same rate. Thus, the narrative’s tale of a progression from judges having minimal power over the admissibility of experts to judges having nearly unfettered power does not describe the reality of toxic tort litigation in the federal courts.

Second, when judges are presented with multiple general causation experts by a plaintiff, they very frequently admit or exclude all of the experts. Likewise, when experts seek to rely on multiple scientific studies, judges tend to admit or exclude the studies in an all-or-nothing fashion. A close examination of some of the cases in which judges excluded all studies relied upon by a plaintiff’s experts suggests that the admissibility rulings were not based on the evidentiary rules that applied at that time. Thus, the conventional narrative is correct that judges decide on the admissibility of expert testimony and scientific studies on the basis of factors other than general acceptance (under *Frye*), or reliability and relevance (under the Federal Rules of Evidence and the *Daubert* trilogy).

Lastly, before *Daubert*, judges were excluding plaintiffs’ experts and thus disposing of cases based on perceived superficial weaknesses in the plaintiffs’ overall cases. However, after *Daubert*, neither superficial weaknesses, the ability of the plaintiff to obtain expert testimony, nor judges’ political orientation appear to correlate with their decisions.

183. John C. Pezzulo, Logistic Regression, <http://statpages.org/logistic.html> (last visited Mar. 28, 2014); P. Wessa, Free Statistics Software, Office for Research Development and Education, version 1.1.23-r7, http://www.wessa.net/rwasp_multipleregression.wasp/ (last visited Mar. 28, 2014).

All of these findings are discussed in greater detail below.

1. Exclusion of Plaintiffs' Experts and Grants of Summary Judgment
Before and After *Daubert*

The conventional narrative maintains that before *Daubert*, judges admitted nearly all proffered expert testimony, and that after *Daubert*, judges excluded far more expert testimony. This study, however, refutes the story of overly-permissive admissibility before *Daubert*. Before *Daubert*, judges excluded all of a given plaintiff's general causation experts in 72 percent of cases. The analysis also refutes the story of drastic change following *Daubert*. Post-*Daubert*, judges excluded all of a plaintiff's general causation experts in 81 percent of cases. This difference is not statistically significant.¹⁸⁴

Other comparisons also suggest that little has changed in the pattern of judicial rulings on the admissibility of plaintiffs' general causation experts. Because proving general causation is a choke point in toxic tort litigation,¹⁸⁵ when a judge excludes all of the plaintiff's general causation experts, the case may be terminated by a grant of summary judgment. Before and after *Daubert*, judges granted motions for summary judgment made along with challenges to the admissibility of plaintiffs' general causation experts in 89 and 85 percent of cases, respectively. Since not all rulings on defendants' challenges to the admissibility of plaintiffs' general causation experts also ruled on motions for summary judgment, grants of defendants' motions of summary judgment resulted in termination of 68 percent of all examined cases before *Daubert* and 55 percent after *Daubert*.¹⁸⁶ Again, the difference is not statistically significant.

At this level of analysis, very little has changed since *Daubert*. Establishing general causation continues to be a major choke point in toxic tort litigation, with judges' exclusion of plaintiffs' general causation experts still effectively terminating over half of all cases. The high rate of exclusion of expert witness testimony before *Daubert* highlights the unreliability of self-reported data and the assumptions of commentators. And the absence of a significant change in the rate of exclusion before and after *Daubert* shows that claims of *Daubert*'s revolutionary impact are vastly overstated.

184. $p = 0.47$.

185. See *supra* Part II.

186. $p = 0.31$.

2. All-or-Nothing Behavior in Admissibility Rulings

Under *Frye*, the judge's process was supposed to be fairly mechanical and limited to determining whether a theory or a scientific study was generally accepted. Under the Federal Rules of Evidence, judges were to determine whether theories or studies were relevant and reliable, although it was unclear whether general acceptance was still required. Under *Daubert*, it was clear that general acceptance was only one among many factors a court could use to determine relevance and reliability. Under all of these approaches, each theory, methodology, or study that an expert cited should have been evaluated independently. A reasonable expectation is that where an expert cites multiple studies as forming the basis of their opinion, some of those studies would be excluded and some admitted. Under the Federal Rules and the *Joiner* decision, the reliability of the expert's inferences drawn from the admitted materials would be assessed. The likely outcome of such a process is that some but not all studies would be admitted and some but not all experts would be admitted. Not that this would happen in every case. Some cases might be so weak that experts could do little more than present largely irrelevant scientific studies and draw unsupported inferences from them. In these cases, judges should exclude all of the cited studies and all of the proffered experts. In some particularly strong cases, experts could rely on any number of relevant and reliable studies to draw well-supported inferences. In these cases, judges should admit all of the cited studies and all of the proffered experts. But particularly where experts seek to use a large number of studies or where plaintiffs proffer multiple general causation experts, judges who apply the rules would likely demonstrate selectivity, admitting some but not all studies and experts in many cases.

To determine whether judges were being selective, and whether, as the conventional narrative would suggest, they became more selective once they had the tools *Daubert* supposedly gave them, this study examined a smaller subset of cases in which judges ruled on challenges to more than one proffered general causation expert or on more than one scientific study upon which an expert sought to base an opinion.

Before *Daubert*, judges admitted some and rejected other plaintiffs' general causation experts in only 14 percent of cases. After *Daubert*, the number remained low—20 percent. In all of the other cases—86 percent and 80 percent before and after *Daubert*, respectively—judges either admitted all or excluded all of the proffered general causation experts. Thus, it does not appear that judges became more selective in their admissibility decisions on plaintiffs' experts in the aftermath of *Daubert*.

Because this frequent all-or-nothing response to multiple proffered experts and scientific studies is inconsistent with the selective approach called for under the Federal Rules of Evidence and the *Daubert*

framework, it seems likely that judges' decision-making process in post-*Daubert* cases is based on factors other than reliability and relevance. It is, however, also possible that a large majority of toxic tort cases are meritless, and that the plaintiffs' experts in these cases proffer only irrelevant scientific studies and draw only unsupported inferences based on those studies. Judges would thus have no choice but to engage in an all-or-nothing exclusion pattern based on the analysis mandated by the Federal Rules and *Daubert*. However, the cases in this data set provide evidence that refutes this contention. Using the adverse case characteristics described above as a marker for overtly weak cases, it is possible to examine whether plaintiffs presenting superficially weaker cases have difficulty engaging general causation experts. Plaintiffs presenting superficially strong cases—those with no adverse case characteristics—proffered more than one general causation expert in 52 percent of cases, while those presenting superficially weaker cases—those with at least one adverse case characteristic—proffered more than one general causation expert in an identical 52 percent of cases. Thus, the apparent strength of a plaintiff's case does not appear to impact the number of experts they can engage. While this only indirectly addresses the contention that weak cases force plaintiffs' experts to rely on irrelevant studies and to draw unsupported inferences, the fact that plaintiffs presenting relatively stronger or weaker cases can proffer the same number of experts¹⁸⁷ suggests that, to the relevant community of experts, there is similar scientific merit in the claims of both groups of plaintiffs.

Further, a close reading of cases in which a judge excluded all of the studies on which a plaintiff's general causation expert sought to rely shows judges using mechanisms that suggest their admissibility rulings are results oriented and based on considerations other than the legal rules. For example, in *Smith v. Ortho Pharmaceutical Corp.*, decided two years before *Daubert*, the district court barred the plaintiff's two experts on causation from testifying, holding that although the experts were both qualified under Rule 702, their proposed testimony was not based on reliable evidence.¹⁸⁸ The experts had formulated their opinions—that nonoxynol-9, the active ingredient in Ortho's spermicide, was capable of causing trisomy 18, a genetic defect—based in large part on four published scientific studies.¹⁸⁹

In that case, the court treated scientific evidence in several ways that

187. Plaintiffs in cases with no adverse case characteristics proffered an average of 2.2 general causation experts, while those in cases with at least one adverse case characteristic proffered an average of 2.0 general causation experts ($p = 0.32$).

188. *Smith v. Ortho Pharm. Corp.*, 770 F. Supp. 1561 (N.D. Ga. 1991).

189. *Id.* at 1576-80.

legal scholars have criticized.¹⁹⁰ First, the court overstated the import of empirical studies that failed to show a correlation between exposure and disease when it stated that “[i]f the statistical likelihood is negligible, it establishes a reasonable degree of medical certainty that there is no cause-and-effect relationship absent some other evidence.”¹⁹¹ By ignoring the crucial issue of whether negative studies had included large enough groups to have statistical significance, the court incorrectly interpreted a study that fails to provide evidence of causation as proving the absence of causation. Second, the judge rejected one study that showed a greater-than-threefold increase in risk of another genetic duplication, trisomy 21 (which causes Down’s Syndrome), because the study did not address trisomy 18 (the plaintiff’s condition).¹⁹² In essence, the court took an “atomistic” approach, requiring plaintiff’s experts to base their opinion on a single study that directly proves causation rather than multiple studies that collectively supported the expert’s opinion. This is the opposite of what scientists do in practice, which is to synthesize all of the available data using a “weight of the evidence” approach. Thus, the judge imposed a judicially determined analytical process on scientific experts. Finally, the court rejected one study because the authors stated that “the results suggest no more than a ‘tentative confirmation of a link.’”¹⁹³ Again, this represents an atomistic approach to the available evidence and, if widely applied, would lead courts to reject almost any study since scientists’ reluctance to pronounce any study definitive leads the insertion of such boilerplate disclaimers into most published studies. Most importantly, this case shows a court applying the all-or-nothing approach.

Ultimately, it appears that the court placed dispositive weight on the fact that the FDA had examined the safety of nonoxynol-9 on two occasions and on both had determined that “the weight of the present evidence does not support an association between the use of spermicides and birth defects.”¹⁹⁴ But the FDA’s findings should affect the weight accorded to the plaintiff’s experts’ testimony, not the admissibility of that testimony. The weight afforded to the evidence should be a question for a jury. Judge Ward did not comment on the irony of his endorsement of the FDA’s use of a weight-of-the-evidence analysis while applying an atomistic approach to the plaintiff’s weight-of-the-evidence analysis. This disparate treatment of the analysis on which the court based its decision and the analysis of the expert whose testimony it rejected

190. *See supra* notes 121-132 and accompanying text.

191. *Smith*, 770 F. Supp. at 1576.

192. *Id.* at 1581.

193. *Id.* at 1578.

194. *Id.* at 1563-64. The court did not comment on the irony of its endorsement of the FDA’s use of a weight-of-the-evidence approach, while applying an atomistic approach of the plaintiff’s evidence.

provides further support for the argument that the judge's decision on admissibility was actually a results-oriented ruling made by means of an evidentiary ruling.

Based on the findings that judges frequently admit or exclude scientific evidence and plaintiffs' general causation experts in an all-or-nothing fashion and that judges appear to consider factors other than reliability and relevance in their admissibility decisions, part of the conventional narrative appears to be possibly correct: judges do appear to be using admissibility rulings to reach certain goals.

3. Non-Evidentiary Factors Underlying Judges' Admissibility Rulings

If the explanation for judges' all-or-nothing admissibility decisions lies beyond the application of the rules of evidence, this raises an obvious question: what factors other than the relevance and reliability of scientific theories and studies and experts' inferences are judges considering? This study examined three possible factors. The only significant correlation was between certain, easily determined, characteristics of weak cases and the exclusion of all of a plaintiff's general causation experts; this correlation was present only in the pre-*Daubert* cases. Two other potential factors—judges' ideology and experience and judges' use of the number of plaintiffs' experts as a proxy for the merits of the cases—did not appear to be significantly correlated with judges' evidentiary rulings.

a. Characteristics of Cases

Certain factual weaknesses of plaintiff's toxic tort case may be readily apparent, even before either side's experts have testified. As described above, this study considered any of the following features to be a weakness that would be readily apparent to a judge at the time of an evidentiary challenge: failure to identify a specific harm, failure to identify a specific substance that caused the alleged harm, failure to describe a timeline relating the exposure to the harm that is consistent with commonly known facts, failure to allege a level of exposure that could believably result in the harm alleged, and prior courts' adverse evaluations of the causal mechanisms espoused by and the qualifications possessed by the plaintiff's general causation experts. The study considered the presence of any one of these five features to mark a case as possessing "adverse case characteristics."

Adverse case characteristics were present in 79 percent of cases before *Daubert* and 42 percent after *Daubert*.¹⁹⁵ This statistically significant

195. $p = 0.005$.

decrease resulted mainly from a clustering of adverse case characteristics in a small number of post-*Daubert* claims, such that these claims contained two or more adverse characteristics. This clustering was unusual in the pre-*Daubert* cases. Thus, cases brought in the post-*Daubert* period tended to be either strong (having no adverse characteristics) or very weak (having two or more adverse characteristics).

Judges' admissibility rulings in the pre-*Daubert* era exhibited a very strong correlation with the presence of any adverse case characteristics. In cases with any adverse characteristic, judges were highly likely to exclude all of a given plaintiff's general causation experts, doing so in 89 percent of cases. In contrast, in cases with no adverse characteristics, judges excluded all of plaintiff's general causation experts in only 29 percent.¹⁹⁶ Thus, cases that exhibited any adverse characteristic were far less likely to survive beyond a challenge to the admissibility of the plaintiff's general causation experts.

Analysis of the post-*Daubert* cases reveals a strikingly different pattern. In this group, judges ruling on cases with or without any adverse characteristics excluded all of the plaintiff's general causation experts in 84 and 79 percent of cases, respectively.¹⁹⁷ Thus, in the post-*Daubert* era, judges no longer appear to be deciding to exclude all of a plaintiff's general causation experts based on their perceptions of the overall merits of the cases. This change is particularly striking in light of the finding that cases in the post-*Daubert* era tended to be more distinctly strong or weak, even at the stage at which challenges to expert witnesses arise.

The conventional narrative is partially consistent with the finding that pre-*Daubert* judges were rendering admissibility rulings in an apparent attempt to prevent weak cases from going to the jury. Judges and commentators have long voiced concerns about the excessive influence that apparently infallible scientific experts may have on lay juries.¹⁹⁸ Pre-*Daubert* judges' admissibility decisions based on superficial case weaknesses may reflect such a distrust of juries' abilities to properly weigh expert witness testimony; judges were likely preventing weak cases from ever reaching the jury where convincing scientific testimony could confuse the jury into ruling for the weaker side.

But other parts of the conventional narrative about judges' worries about the effect of apparently infallible experts appear to have been wrong. It appears that all along judges had the upper hand. The pre-*Daubert* cases examined in this study suggest that in the *Frye* and FRCP eras, judges could control case outcomes by excluding all of a plaintiff's

196. $p = 0.006$.

197. $p = 0.64$.

198. See *supra* note 5 and accompanying text.

scientific evidence and expert testimony on general causation.

The conventional narrative appears to be largely incorrect in the post-*Daubert* era as well. Although high, the rate at which judges excluded scientific evidence and expert testimony was unchanged. Hence, *Daubert* does not appear to have effected a revolutionary expansion in judges' ability to exclude evidence and experts. But something has indeed changed. The adverse case characteristics, which reflect a commonsense understanding of cases' weaknesses, no longer correlate with the exclusion of all of plaintiffs' expert testimony post-*Daubert*. Rather, *Daubert* appears to have freed federal judges to dispose of cases based on other considerations. The remainder of this study examined some possible considerations that are used by judges.

b. Judges' Political Views and Judicial Experience

A substantial body of scholarship suggests that partisanship accounts for a significant portion of judicial decision making.¹⁹⁹ Most of this work has focused on the appellate courts, with relatively little attention devoted to the district courts. Studies of federal district court judges do support some influence of partisanship, typically defined as a judge's known party affiliation or based on the party to which the President who appointed the judge belonged.²⁰⁰ These studies have examined decisions on politically-charged topics such as legislative reapportionment.²⁰¹

As discussed above, although admissibility rulings may seem far removed from these kinds of politically charged issues, toxic tort cases do contain a large political dimension. Many cases have the potential to impose large costs on corporate defendants, and all toxic tort cases raise the question of the appropriate—or, depending one's ideology, the efficient—allocation of risk and cost. Thus, judges' political and ideological preferences are reasonable factors to investigate in trying to ascertain what is driving their admissibility decisions.

In this study, party affiliation had no significant effect on judges' admissibility decisions. Judges' party affiliation was determined by the party to which the president who appointed the judge belonged. Judges appointed by Democratic and Republican presidents excluded all of a given plaintiff's general causation experts in 82 and 74 percent of cases,

199. See, e.g., CASS R. SUNSTEIN ET AL., ARE JUDGES POLITICAL?: AN EMPIRICAL ANALYSIS OF THE FEDERAL JUDICIARY (2006); Randall D. Lloyd, *Separating Partisanship from Party in Judicial Research: Reapportionment in the U.S. District Courts*, 89 THE AM. POL. SCI. REV. 413 (1995)

200. See Lloyd, *supra* note 199, at 417-18 (finding that district court judges hold reapportionment plans drawn by state legislatures controlled by the opposite party to a higher standard than plans drawn by nonpartisan legislatures or those controlled by the same party).

201. See *id.*

respectively.²⁰² When decisions were stratified by whether they were rendered before and after *Daubert*, there were still no statistically significant differences related to party affiliation, although, before *Daubert*, judges appointed by Democratic or Republican Presidents excluded all of a given plaintiff's general causation experts in 83 and 68 percent of cases, respectively.²⁰³ Neither of these differences was statistically significant, indicating that party affiliation did not influence judges admissibility decisions either before or after *Daubert*.

This finding argues against the claims of some legal scholars that judges' admissibility decisions have become based on politics and ideology, and specifically counters the argument that judges have used their *Daubert*-conferred gatekeeper powers to tilt the playing field in favor of corporate defendants. Although using the party affiliation of the President who appointed a judge as a proxy for the judge's ideology is imperfect, the analysis does not support the arguments made by scholars such as Lucinda Finley and Margaret Berger that there has been a strong post-*Daubert* shift in toxic tort litigation motivated by a desire to move away from "consumer protective social policies" toward more corporation-protecting policies.²⁰⁴ Rather, the study can be interpreted as showing that such attitudes had been present—and were being implemented—well before *Daubert*.

The study also postulated that the length of time a judge had spent on the federal district court bench might influence the judge's admissibility practices. Specifically, the study asked whether less experienced judges might be more or less inclined to admit testimony and might feel more or less pressure to shrink their dockets. The study hypothesized that judges with fewer years on the bench are more likely to admit plaintiffs' general causation experts.

However, the study concluded that the length of time a judge had spent on the federal district court bench also had no impact on their admissibility decisions. Judges with less than ten years of experience—a threshold which was chosen to divide the judges into two nearly-even sized groups—excluded all plaintiffs' general causation experts in 78 percent of cases, while judges with ten or more years of experience excluded all plaintiffs' general causation experts in 77 percent of cases. Setting the dividing line at five years of experience also failed to yield a significant difference.²⁰⁵

202. $p = 0.41$.

203. $p = 0.46$.

204. See *supra* notes 129-132 and accompanying text.

205. Judges with five or fewer years of experience excluded all plaintiffs' general causation experts in 64 percent of cases, while judges with more than five years of experience excluded all plaintiffs' general causation experts in 80 percent of cases ($p = 0.23$ using a two-tailed analysis).

Thus, neither political affiliation nor federal judicial experience appears to correlate with judges' admissibility rulings.

c. The Number of Plaintiffs' General Causation Experts

Judges evaluating the merits of a case might also use the number of proffered general causation experts as a proxy for the strength or weakness of a plaintiff's case. Plaintiffs whose claims were weak because there was little reason to believe the substance to which they were exposed could cause the harm might be able to find fewer general causation experts to testify in their behalf. Courts and commentators have claimed that under *Frye*, judges used the number of experts as a proxy for the merits of a case, and criticized this practice "because of the importance it places on the judge's subjective ability to 'count heads' among experts in the scientific community."²⁰⁶

In the toxic tort cases included in this study's database, "counting heads" did not seem to be a common occurrence. Looking at the number of general causation experts proffered by the plaintiffs, that number did not correlate with judges' decisions to exclude all of the plaintiffs' general causation experts. Plaintiffs who had at least one general causation expert admitted to testify had proffered an average of 2.4 experts, while those whose general causation experts were all excluded had proffered an average of 2.0 experts, a difference that is not statistically significant.²⁰⁷ Dividing opinions into those issued before and after *Daubert* shows the same nonsignificant difference in each group.²⁰⁸

Admittedly, this quantitative analysis does not fully address the "counting heads" argument because that argument assumes that judges count the heads on both the plaintiff's and the defendant's sides. Unfortunately, evidentiary rulings often address only the plaintiffs' experts since most often it is the defendant who raises the challenge. Thus, the number of defendants' general causation experts often cannot be determined by reading the judges' rulings.

This analysis does, however, tend to refute the theory that judges use the number of a plaintiff's general causation experts as a proxy for the strength of the case and then base their admissibility rulings off the number of experts presented by the plaintiff.

206. *Ellis v. Int'l Playtex, Inc.*, 745 F.2d 292 (11th Cir. 1984).

207. $p = 0.26$ (using a one-tailed test, assuming a unidirectional response).

208. Prior to *Daubert*, plaintiffs in this study who had at least one expert admitted to testify had proffered a mean of 2.6 experts, and plaintiffs who had no experts admitted had proffered a mean of 2.2 experts. The p value was 0.7. After *Daubert*, the mean number of experts proffered was 2.2 and 1.9, respectively. The p value was .66.

V. REVISING THE NARRATIVE

As the analysis in Part I shows, the conventional narrative of the rules governing the admissibility of expert witness testimony and scientific evidence is largely incomplete. The data from the empirical study presented in this Article suggests that the conventional narrative is also incorrect in several respects. This Part proposes a new narrative that is both more consistent with the data and more complete in the hope of explaining what federal judges are actually doing when they make rulings at the evidentiary challenge stage of toxic tort cases.

The new narrative begins before *Daubert*, when judges excluded plaintiffs' general causation experts in the vast majority of cases. Whether deciding under the *Frye* general acceptance test or under the relevance and reliability requirements of Rule 702, judges were actually using their rulings on challenges to the admissibility of plaintiffs' experts to dispose of superficially weak cases. Presumably, these judges were deeply cynical about juries' abilities to weigh the testimony of apparently infallible scientific experts. But, even if cynicism about juries' abilities was not the reason, judges clearly held the upper hand and firmly controlled which experts were admitted to testify in toxic tort cases.

The impact of the *Daubert* trilogy has been negligible in at least one way: judges have excluded plaintiffs' expert witnesses and disposed of plaintiffs' cases at nearly the same rate after the trilogy as they did before *Daubert* and its trilogy. Thus, from a plaintiff's perspective, little has changed. However, the trilogy has been revolutionary in other ways. Judges are no longer excluding experts and disposing of cases on the basis of superficial weaknesses. Further, judges do not seem to be acting in pursuit of political or ideological ends. At this time, although the data suggests that some of the motivations that other scholars have proposed are not correct, the judges' motivations remain unknown.

This new narrative suggests some promising areas for future study by scholars interested in understanding judges' actual practices. The ascendance of the law and economics movement has occurred over the same time period that this study covers. Although analysis of judges' party affiliation did not detect an effect on their admissibility rulings, it may be fruitful to examine whether judges' embrace of economic analysis correlates with those rulings. Also, it may be beneficial to examine whether judges' dispositions of toxic tort cases reflected a concern with the efficient allocation of costs; this might be studied by correlating the potential economic impact of an adverse outcome on the defendant (would it affect only the defendant, or an entire industry) and on the plaintiff (would it affect an individual plaintiff, or an entire class) with the judge's evidentiary ruling.

Another potentially promising area of future research is whether judges

post-*Daubert* are using characteristics of experts as proxies for the merits of plaintiffs' cases. In *Kumho*, Justice Breyer seemed to suggest that judges should make their criteria for qualifying experts more stringent.²⁰⁹ Although this study did not detect an impact of the number of general causation experts a plaintiff proposed, judges may be looking at the relative number of general causation experts put forth by each side, or they may be looking more closely at the qualifications of each expert. Thus, judges may be using factors such as an expert's publication record or the "fit" between the expert's field of expertise and the subject of the testimony the expert proposes to give in order to determine the strength of a plaintiff's case.

How we understand contemporary judges' behavior—and hence the impact of the *Daubert* trilogy—might depend on whether we focus on process or on outcome. If what we focus on is outcome, what we really want is to know whether or not toxic tort suits are disposed of correctly (assuming that a "correct" disposition can objectively be determined); this study can provide only a partial answer. Judges are now disposing of some cases that lack obvious weaknesses for reasons that are not clear. It is possible that an examination of other factors, such as the relative degree of specialized expertise possessed by each sides' experts or the economic impact of a judgment against one side or the other will suggest judges' underlying motivations. But whatever these motivations are—and even if judges are simply "following the rules"—the current outcomes are undesirable because cases asserting colorable claims are being disposed of based on the pretense of a failure to establish general causation. Conversely, judges are now allowing a similar percentage of facially weak cases to go to the jury, even though there is a long history of judges stating concerns that experts' apparent infallibility may corrupt the decision making of lay jurors. Thus, from a results-oriented perspective, the answer to the question of how we judge the current behavior is with at least partial disapproval. Ultimately, though, if we are concerned with outcomes, we need empirical evidence on the societal impact of judges' admissibility decisions in toxic tort cases.

If, however, we are more concerned with process—whether judges are following the rules—this study shows that the answer is that judges did not appear to be following the rules before *Daubert* and they do not appear to have been following the (new) rules after *Daubert*. Before *Daubert*, judges were disposing of cases at the general causation admissibility stage based on overt weaknesses not tied to general causation. But what about contemporary judges' behavior? Again, the specific answer to the question about the bases for judges' admissibility rulings is unclear based

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

on this study, but it is likely that judges are not simply applying the rules as set out in Federal Rule of Evidence 702 and the *Daubert* trilogy.

VI. CONCLUSION

This Article shows that the accepted narrative about the admissibility of expert witness testimony is nuanced and, in many particulars, incorrect. The *Daubert* revolution was in part a non-event. At least in the world of toxic tort litigation, *Frye* was already dead when *Daubert* was decided, and *Daubert* simply gave judges a little more flexibility to do what they were already doing—excluding most experts proffered by plaintiffs. In terms of overall case dispositions at the evidentiary challenge stage, very little has changed: similar numbers of cases are terminated by a defendant's challenge to the admissibility of the plaintiffs' general causation experts. But there has been a revolutionary change as well: federal judges appear to be deciding cases in the post-*Daubert* era on the basis of a very different set of considerations. This study suggests that while, before *Daubert*, judges used the evidentiary challenge stage as means of disposing of superficially weak cases, after *Daubert*, this is no longer the case. Further empirical research is needed help to understand the practices of contemporary judges, and thus to fully evaluate the impact of the Federal Rules of Evidence and the *Daubert* trilogy, but the study presented in this Article provides a strong foundation for future empirical studies.