First *Annual Judicial Symposium*

JUSTICE AND SCIENCE

Course Materials

July 2005

All views, opinions and conclusions expressed are those of the authors and/or speakers, and do not necessarily reflect the opinion and/or policy of NFJE and its leadership.

150 North Michigan Avenue, Suite 310
Chicago, Illinois 60601

Produced in the United States of America

Course Materials Table of Contents

The *Daubert* Trilogy in the States

David E. Bernstein

Professor
George Mason University School of Law

Jeffrey D. Jackson

Visiting Assistant Professor of Law
Washburn Law School

DAVID E. BERNSTEIN is a Professor at the George Mason University School of Law in Arlington, Virginia, where he has been teaching since 1995. Professor Bernstein is a graduate of the Yale Law School, where he was senior editor of the *Yale Law Journal* and a John M. Olin Fellow in Law, Economics, and Public Policy. Professor Bernstein's recent law review articles have appeared in the *Yale Law Journal*, *Michigan Law Review*, *Texas Law Review*, *Georgetown Law Journal*, *Vanderbilt Law Review*, and *California Law Review*, among others. He is the author of *You Can't Say That! The Growing Threat to Civil Liberties from Antidiscrimination Laws* (Cato Institute 2003). Professor Bernstein is also the co-author of *The New Wigmore: Expert Evidence* (Aspen Law and Business 2003), author of *Only One Place of Redress: African-Americans, Labor Regulations, and the Courts from Reconstruction to the New Deal* (Duke 2001) and co-editor of *Phantom Risk: Scientific Inference and the Law* (MIT 1993). Professor Bernstein teaches Torts, Products Liability, Evidence, Constitutional Law, and Scientific and Expert Evidence.

JEFFREY D. JACKSON comes to Washburn Law from the Kansas Supreme Court where he was staff attorney for Death Penalty and Constitutional issues. Prior to that, he was a law clerk for the Honorable Mary Beck Briscoe in the U.S. Court of Appeals for the Tenth Circuit, law clerk to The Honorable Justice Robert E. Davis at the Kansas Supreme Court, an associate at Bennett & Dillon L.L.P., in Topeka and staff attorney for the Kansas Court of Appeals. Jackson received his B.B.A. in economics from Washburn University in 1989, his J.D. from Washburn Law in 1992 and his LL.M. in Constitutional Law at Georgetown University Law Center in 2003. At Washburn Law, Jackson was assistant editor for the Washburn Law Journal. Jackson is admitted to practice in Kansas, Missouri, U.S. District Court for the District of Kansas, U.S. District Court for the Western District of Missouri and the U.S. Court of Appeals for the Tenth Circuit. He is a member of the Kansas Judicial Council Death Penalty Advisory Committee.

The *Daubert* Trilogy in the States

Table of Contents

The Daubert Trilogy in the States

Abstract: The *Daubert* trilogy of Supreme Court cases—*Daubert*, *Joiner*, and *Kumho Tire,* codified in Federal Rule of Evidence 702—has established new rules the admissibility such evidence in federal court.

The situation in state courts is far more unsettled. First, a significant number of courts have continued to adhere to the tests they used before *Daubert,* either *Frye* general acceptance test or some other test.

Even among those states which have adopted *Daubert*, its application has been decidedly nonuniform. Only a few states have adopted the *Daubert* trilogy in its entirety. Some states have adopted *Daubert*, but not yet adopted *Kumho Tire* or *Joiner*. Others have adopted *Daubert* and *Kumho* Tire, but not *Joiner*, or have adopted only part of *Joiner*. Still other states view the *Daubert* trilogy as only instructive or consistent with their own traditional state tests.

This article analyzes the degree to which the holdings of the *Daubert* trilogy have been adopted by state courts. This analysis shows that there is a rich diversity of tests within the states, so much so that, contrary to the prevailing impression, the *Daubert* trilogy is not yet the majority standard even among the states that have rejected *Frye*.

Citation: David E. Bernstein and Jeffrey D. Jackson, The *Daubert* Trilogy in the States, 44 Jurimetrics J. ___–___ (2004).

> In 1993, the Supreme Court decided what became the first in the "*Daubert* trilogy" of cases concerning the admissibility of expert testimony, *Daubert v. Merrell Dow Pharmaceuticals, Inc.*[1] In *Daubert*, the Court determined that Federal Rule of Evidence 702 mandated that scientific evidence be subject to a reliability test, rather than the common law "general acceptance test" set forth in *Frye v. United States*.[2] In place of *Frye*, the Court imposed upon judges the "gatekeeping" responsibility of assessing whether "the reasoning or methodology underlying the testimony is scientifically valid and…whether that reasoning or methodology properly can be applied to the facts at issue."[3] The Court set forth several general factors that might be considered in reaching a decision on whether to admit scientific expert evidence, including: (1) whether the theory or technique can be or has been tested; (2) whether the theory or technique has been subjected to peer review and publication, as such review "increases the likelihood that substantive flaws in the methodology will be detected"; (3) in the case of the particular technique, the known or potential rate of error; and (4) whether the theory or technique enjoys general acceptance within the relevant scientific community.[4]

Daubert changed the nature of the admissibility determination by judges from *Frye's* deference to the views of scientists in a relevant field to an independent evaluation of the proffered evidence.[5] Courts and commentators disagreed, however, regarding whether this "revolution"[6] in how judges were to go about deciding whether to admit scientific evidence would lead to more permissive or more restrictive admissibility rulings.

Both sides of the debate could point to evidence favoring their position. Those who believed that *Daubert* was relatively permissive noted that the *Frye* general acceptance test had

been considered by many an inflexible and unduly conservative test that excluded innovative scientific techniques.[7] *Daubert*, by promising flexibility and case-by-case analysis of admissibility, seemed to liberalize admissibility rules by allowing courts to reconsider the validity of polygraph tests and other forensic techniques that had been excluded on general acceptance grounds but which some scientists believed were reliable. On the other hand, *Daubert* also gave courts the opportunity to reconsider the admissibility of common forensic techniques, such as handwriting analysis, that were generally accepted among forensic scientists, but had never been proven reliable.[8]

Similarly, before *Daubert* a string of recent federal circuit court opinions had applied *Frye* to toxic tort evidence and had interpreted *Frye* rather stringently.[9] Because plaintiffs' proffered evidence in toxic tort cases is often sufficiently novel that it will not yet have achieved general acceptance, the *Frye* test, especially if applied to an expert's conclusions, threatened to severely impede the ability of toxic tort plaintiffs to present admissible evidence. *Daubert*, by emphasizing the flexibility of the admissibility determination and by admonishing courts that they must focus on an expert's methodology, not his conclusions, appeared to some to be a more permissive alternative to *Frye*.[10] The death of *Frye* seemed an especially serious blow to advocates of strict scrutiny of scientific evidence in toxic tort cases because their most well-known champion, Peter Huber, was a strong advocate of *Frye*.[11]

However, Huber and other advocates of strict scrutiny expressed satisfaction with *Daubert*.[12] Despite the recent precedents noted above, *Frye* had rarely been applied to toxic tort evidence. Moreover, courts had often applied *Frye* in a cursory manner, examining only the general acceptance of an expert's overarching methodology and not also whether that methodology was used in the particular case in a generally accepted way. *Daubert*, by contrast, required an inquiry into the reliability of proffered scientific evidence, an inquiry that, it was argued, would inevitably require courts to review experts' reasoning process,[13] something they rarely did under *Frye*.

Debate along these lines raged for several years, until the Court decided the second case in the *Daubert* trilogy, *General Electric Co. v. Joiner*,[14] in 1997. *Joiner* clarified the *Daubert* test in two important respects. First, *Joiner* made it clear that courts could scrutinize the reliability of an expert's reasoning process as well as the expert's general methodology and that "nothing in either *Daubert* or the Federal Rules of Evidence requires a district court to admit opinion evidence that is connected to existing data only by the *ipse dixit* of the expert."[15] Instead, courts were free to conclude that "there is simply too great an analytical gap between the data and the opinion proffered."[16] The *Joiner* Court, moreover, did not simply remand the case but sent a clear "strict scrutiny message" by reversing the Eleventh Circuit and upholding the district court's exclusion of the marginal causation evidence that plaintiffs often rely on in toxic tort cases. Those who had hoped that *Daubert* had liberalized the admissibility standards for evidence in toxic tort cases conceded defeat.[17]

Further, *Joiner* made it clear that the decision of the trial court judge as to whether to admit particular scientific evidence was to be reviewed only for an abuse of discretion.[18] The Court rejected the notion propounded by several circuits that they should engage in especially stringent review of decisions excluding scientific evidence proffered by plaintiffs in toxic tort and products liability cases.[19]

The Daubert trilogy was later codified in Federal Rule of Evidence 702, as amended in December 2000:

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

Meanwhile, many states changed their evidentiary rules in reaction to Daubert. Although binding only on federal courts, a number of state courts quickly adopted the Daubert test for the admissibility of scientific evidence.[20]

By 1998 one commentator claimed that "thirty-three states have adopted *Daubert* in essence."[21] This statement greatly overstated *Daubert's* influence. It appears that the author's count included all jurisdictions that adopted a test other than *Frye*, which remains the rule in a significant minority of states.[22] As discussed below, however, by mid-2003 only twenty-seven states had adopted a test consistent with *Daubert*. Also, "consistent with" is a relatively broad category, encompassing states that have not directly adopted *Daubert* but have stated that *Daubert* is consistent with their state tests or is otherwise instructive. Also, courts that have adopted or stated their approval of *Daubert* have not necessarily adopted *Joiner* or *Kumho Tire's* expansion of *Daubert*.

Some of the states that adopted *Daubert* had been *Frye* jurisdictions, and their courts endorsed *Daubert* because they believed that it was a more permissive test than *Frye*.[23] Other state courts thought *Daubert* was consistent with liberal admissibility rules they had previously adopted. However, as some commentators predicted,[24] *Daubert*, particularly as extended by *Joiner* and *Kumho Tire*, became far broader than *Frye* ever was. Unlike *Daubert*, *Frye* was limited to "novel" scientific evidence; unlike *Joiner*, *Frye* typically was applied only to general methodology, not to reasoning; unlike *Kumho Tire*, *Frye* did not apply to nonscientific evidence.

Not surprisingly, then, some of the states that had jumped quickly on the *Daubert* bandwagon have been more reluctant to adopt *Joiner* and *Kumho Tire*. Only nine of the *Daubert* states have either explicitly or implicitly adopted all of the holdings of the trilogy.[25] Several states, meanwhile, have rejected all or part of *Joiner* or *Kumho Tire*.

This article analyzes the degree to which the holdings of the *Daubert* trilogy have been adopted by various state courts. Part I discusses states that have adopted the *Daubert* trilogy in its entirety, either explicitly or implicitly. Part II discusses states that have adopted the holdings of *Daubert* and *Kumho Tire* but have not fully adopted *Joiner*. Part III discusses states that have adopted *Daubert* but not *Kumho Tire*. Part IV discusses states that have refused to explicitly adopt *Daubert* but view some or all of the trilogy as instructive or consistent with their own state tests. Finally, Part V discusses states that have rejected *Frye* but apply tests for the admissibility of expert testimony that bear little resemblance to *Daubert*. This analysis shows that there is a rich diversity of tests within the states, so much so that the *Daubert* trilogy is not the majority test even among the states that have rejected *Frye*.

I. States Adopting the *Daubert* Trilogy in its Entirety

Only nine states—Arkansas,[26] Delaware,[27] Louisiana,[28] Massachusetts,[29] Mississippi,[30] Nebraska,[31] Oklahoma,[32] Texas,[33] and Wyoming[34]—have either explicitly or implicitly adopted the

full holdings of the *Daubert* trilogy. Even this may be an overstatement, however. The two judicial opinions in Louisiana adopting *Joiner*'s scrutiny of the expert's reasoning process and abuse of discretion standards are Court of Appeals cases.[35] The Louisiana Supreme Court has not yet ruled on the viability of *Joiner* in Louisiana.

II. States Adopting *Daubert* and *Kumho* Tire But Not Joiner

Six states have adopted the reasoning and holdings of both *Daubert* and *Kumho Tire* but have not adopted *Joiner*. In Kentucky,[36] Ohio,[37] North Carolina,[38] Rhode Island,[39] and South Dakota,[40] the issues involved in *Joiner* have not yet arisen. New Hampshire, by contrast, has adopted both *Daubert* and *Kumho Tire*[41] but has repudiated *Joiner's* abuse of discretion standard, at least insofar as it applies to situations where "the reliability or general acceptance of novel scientific evidence is not likely to vary according to the circumstances of a particular case."[42] The New Hampshire Supreme Court held that the level of scrutiny that it applies to lower courts' admissibility decisions will vary depending on "the complexity of the evidence involved and the impact the evidence likely will have on the trial itself," and where the reliability of the theory underlying the technique would not vary with each case, it would review the lower court's decision with regard to reliability *de novo*.[43]

III. States Adopting *Daubert* But Not *Kumho Tire* or *Joiner*

Seven states have adopted *Daubert's* reasoning, but have not adopted *Kumho Tire's* holding that the reliability test applies to all expert evidence or *Joiner's* holding that the reliability test may be applied to an expert's reasoning process. As discussed below, in most of these states these issues have not yet explicitly arisen, but a few states have implicitly or explicitly repudiated *Kumho Tire*.

Of the states that have not yet explicitly ruled on *Kumho Tire*, Vermont is the most likely to adopt *Kumho Tire's* holding. The Vermont Supreme Court adopted *Daubert* in 1995,[44] and more recently referred approvingly to *Kumho Tire* in discussing its *Daubert* test.[45] It appears, therefore, that Vermont is likely to adopt *Kumho Tire's* holding if squarely faced with the question.

Alaska has adopted *Daubert* and has also adopted *Joiner's* abuse of discretion standard.[46] However, the Alaska Supreme Court has not expressed a view on *Kumho Tire* or *Joiner's* expansion of the *Daubert* test. The Alaska Court of Appeals favorably addressed *Kumho Tire* in a pair of unpublished opinions, *Bourdon v. State*[47] and *Vent v. State*.[48] In *Bourdon*, the court, citing *Kumho Tire*, applied Daubert to child abuse expert testimony.[49] In *Vent*, the court noted scholarly criticism of extending *Daubert* beyond scientific testimony but ultimately followed *Kumho Tire* in applying *Daubert* to expert testimony regarding the psychology of confessions.[50]

New Mexico, one of the first state courts to adopt *Daubert*,[51] has not yet explicitly adopted *Kumho Tire*, although the Court of Appeals suggested *in dicta* that it would adopt *Kumho Tire*.[52] It is not clear whether the New Mexico Supreme Court would adopt *Kumho Tire* if given the opportunity,[53] nor has the court expressed an opinion on the substantive aspect of *Joiner*.

Similarly, although Connecticut has adopted *Daubert*,[54] the Connecticut Court of Appeals has noted that "[o]ur courts…have neither adopted nor rejected the *Kumho Tire Co., Ltd.*, rule."[55] The commentary to Connecticut Code of Evidence §7–2, which incorporates the *Daubert* standard, states that it "should not be read either as including or precluding the *Kumho*

Tire rule."[56] Connecticut has adopted *Joiner's* abuse of discretion standard[57] but has not addressed whether to adopt *Joiner's* scrutiny-of-the-reasoning-process holding.[58]

In *State v. O'Key*,[59] the Oregon Supreme Court, although expressly stating that *Daubert* was not binding, adopted a similar "gatekeeping" rule for scientific evidence.[60] However, Oregon has not made a decision whether to adopt *Kumho Tire*. The Oregon Court of Appeals has noted that "[i]t is by no means clear" whether courts are to assume a gatekeeping role for evidence other than scientific evidence.[61] The Oregon Supreme Court has, however, expressly repudiated *Joiner's* abuse of discretion standard of review, stating that "[n]otwithstanding the usual deference to trial court discretion, we as an appellate court retain our role to determine the admissibility of scientific evidence under the Oregon Evidence Code."[62] The Oregon Supreme Court has also implicitly rejected *Joiner*. The court held that courts may only scrutinize an expert's general methodology and not his reasoning.[63]

West Virginia, another early adopter of *Daubert*,[64] has also failed to adopt *Kumho Tire*. In *West Virginia Div. of Highways v. Butler*,[65] the West Virginia Supreme Court explicitly declined to adopt *Kumho Tire*, applying instead a more permissive standard for nonscientific expert evidence.[66] Two years later in *Watson v. Inco Alloys Int'l, Inc.*,[67] the court again rejected *Kumho Tire*.[68] However, Justice Davis, the author of the majority opinion, urged the court to adopt *Kumho Tire* in the future.[69] In *State v. Leep*,[70] the West Virginia Supreme Court noted the controversy but decided not to settle it, stating that it would "reserve further consideration thereof for a more factually appropriate case."[71]

Montana, also an early adopter of *Daubert*,[72] has implicitly rejected *Kumho*.[73] In *State v. Cline*, the Montana Supreme Court confined consideration of the *Daubert* factors to "novel scientific evidence."[74] The Montana Supreme Court has not explicitly taken a position on *Joiner*, but it appears to follow *Joiner's* abuse of discretion standard.[75]

Alabama applies *Daubert* only in very limited circumstances. Alabama Code §36-18-30 mandates that the *Daubert* test be applied to "[e]xpert testimony or evidence relating to the use of genetic markers contained in or derived from DNA for identification purposes…"[76] The Alabama Supreme Court has refused to apply *Daubert* outside of this specific circumstance.[77] The Alabama Supreme Court has not addressed the applicability of either factor of *Joiner*.

IV. States Holding That *Daubert* is Instructive

There are five states that have not adopted *Daubert* but utilize the *Daubert* factors in interpreting their own tests. As might be expected, these courts vary widely on the application of the rest of the *Daubert* trilogy. Tennessee, for example, has adopted *Daubert's* test in all but name.[78] Tennessee has also adopted *Kumho Tire's* holding extending the *Daubert* factors to all expert evidence, as well as *Joiner's* scrutiny of the reasoning process and standard of review.[79]

Hawaii has expressly refrained from adopting *Daubert*.[80] However, the Hawaii Supreme Court has held that, because the Hawaii Rules of Evidence are patterned after the Federal Rules, *Daubert* is "instructive" in interpreting the state rule.[81] The Hawaii Supreme Court has stated that, to be admissible, evidence must be both "relevant" and "reliable" and that the *Daubert* factors are helpful in assessing reliability.[82] The court applies this analysis to all expert evidence, based on *Kumho Tire*.[83] Further, the court has partially adopted *Joiner's* abuse of discretion test

but reviews ultimate conclusions *de novo*.[84] Hawaii courts have not addressed whether courts may scrutinize an expert's reasoning under the Hawaii Rules of Evidence.

Indiana has held that *Daubert* is "helpful, but not controlling" in interpreting its rule of evidence.[85] Indiana has also implicitly rejected *Kumho Tire's* extension of courts' *Daubert* gatekeeping role to nonscientific expert testimony. The Indiana Supreme Court recently held that Indiana's version of Rule 702 applies only to scientific evidence and not to expert testimony based on specialized knowledge.[86] Although the Indiana Supreme Court has not passed on *Joiner's* holding regarding scrutiny of the reasoning process, the Indiana Court of Appeals has followed *Joiner* on this issue.[87] While not explicitly addressing *Joiner's* abuse of discretion standard, the Indiana Supreme Court has held that decisions on reliability should be reviewed for abuse of discretion.[88]

Iowa has held that district courts are not required to consider the *Daubert* factors, but they may do so if they find that the factors are "helpful," especially in complex cases.[89] Similarly, judges have the discretion to apply the *Daubert* factors to all expert evidence in such cases, not just to scientific evidence.[90] The focus, according to the state supreme court, should be entirely on the principles and methodology utilized by the expert, not on the conclusions generated.[91] This language was taken from *Daubert*, but without the clarifying language from *Joiner* that conclusions and methodology are not completely distinct and that reasoning may therefore be scrutinized. Implicitly, then, Iowa has rejected that aspect of *Joiner*. The virtually unlimited discretion allowed Iowa courts in the application of the factors is consistent with *Joiner's* standard of review.

Colorado has adopted a "totality of the circumstances" test for scientific expert evidence that focuses on whether the evidence involved is both relevant and reliable.[92] In assessing the reliability, courts are directed to conduct a broad inquiry and may consider the factors set forth in *Daubert*.[93] The use of the *Daubert* factors apparently applies only to scientific evidence, rather than to all expert evidence, so Colorado seems to have rejected *Kumho Tire*.[94] Colorado has not discussed *Joiner's* "reasoning process" holding, but it appears that the expansive discretion granted to trial courts in determining relevance and reliability would not foreclose it.[95] This expansive discretion is also consistent with *Joiner's* abuse of discretion standard of review.

Maine has not adopted *Daubert*. However, its analysis of its own version of pre-2000 Fed. R. Evid. 702 parallels *Daubert*, and the Maine Supreme Court cites *Daubert* along with its own cases.[96] It appears that Maine applies this analysis to all expert evidence.[97] The review is for "abuse of discretion or clear error."[98] No cases have addressed the question of whether *Joiner*'s reasoning process analysis applies.

V. Non-*Frye* States That Reject *Daubert*

Idaho's test for admissibility of expert evidence is simply to follow the language of Idaho Rule of Evidence 702, which is identical to pre-2000 Fed. R. Evid. 702.[99] Because this standard is vague, the Idaho Court of Appeals has looked to *Daubert* for guidance on several occasions.[100] However, the Idaho Supreme Court recently reaffirmed its rejection of the application of *Daubert* to expert testimony and cited with approval a "bare analysis" of expert testimony conducted by a trial court.[101]

New Jersey has refused to adopt *Daubert* but follows an analysis that is somewhat similar for scientific evidence in civil cases.[102] Under this analysis, "an expert must be able to identify the factual basis for his conclusion, explain his methodology, and demonstrate that both the factual

basis and underlying methodology are scientifically reliable."[103] New Jersey has not applied this analysis beyond scientific evidence, and, despite the test's superficial similarity to *Daubert* and *Joiner*, the New Jersey courts are known to be quite liberal about admitting expert scientific testimony in civil cases. Further, the standard of review in New Jersey of trial court admissibility rulings is unclear.[104]

Like *Daubert*, North Dakota's test for the admissibility of scientific evidence requires trial courts to inquire into the relevance and reliability of the expert testimony.[105] However, the North Dakota rule "envisions generous allowance of the use of expert testimony if the witnesses are shown to have some degree of expertise in the field in which they are to testify."[106] This standard is far more liberal than *Daubert* and *Joiner*.[107]

South Carolina's test, like *Daubert*, imposes a "gate-keeping" requirement for scientific evidence.[108] However, South Carolina's test uses different factors to assess reliability: "(1) the publications and peer review of the technique; (2) prior application of the method to the type of evidence involved in the case; (3) the quality control procedures used to ensure reliability; and (4) the consistency of the method with recognized scientific laws and procedure."[109] The standard of review is abuse of discretion, but, unlike the *Daubert* factors, which are merely illustrative, South Carolina courts must apply their states' four factors.[110] In practice, South Carolina's test appears to be far more forgiving than *Daubert-Joiner*.[111]

Utah applies an apparently more rigorous standard than *Daubert* to the admission of scientific evidence.[112] In *State v. Crosby*, the Utah Supreme Court determined that its long-held standard was similar to, although less flexible than, *Daubert*.[113] Under Utah's standard, the court must determine whether scientific evidence is "inherently reliable" before it can be admitted. Such a determination involves exploration of "the correctness of the scientific principles underlying the testimony, the accuracy and reliability of the techniques utilized in applying the principles to the subject matter before the court and in reaching the conclusion expressed in the opinion, and the qualifications of those actually gathering the data and analyzing it."[114] The trial court should "carefully explore each logical link in the chain that leads to the expert testimony given in court and determine its reliability."[115] However, the Utah test applies only to expert testimony "based on newly discovered principles"[116] or testimony based on "novel scientific methods and techniques."[117] Other types of expert testimony, such as medical diagnoses, are subject to a much more liberal test.

Virginia imposes a duty on the trial court, as in *Daubert*, to make a threshold finding of reliability whenever "unfamiliar" scientific evidence is offered.[118] However, Virginia has expressly left open the question of whether the *Daubert* factors could be applied to this determination.[119]

Wisconsin, while not ruling out the possibility that it will some day adopt *Daubert*,[120] for now eschews any substantive test for the admissibility of expert testimony—mere qualifications are sufficient:

The fundamental determination of admissibility comes at the time the witness is "qualified" as an expert. In a state such as Wisconsin, where substantially unlimited cross-examination is permitted, the underlying theory or principle on which admissibility is based can be attacked by cross-examination or by other types of impeachment.[121]

Nevada has a similarly liberal test for the admission of expert testimony.[122] The Georgia Supreme Court has not addressed *Daubert*, but the state Court of Appeals has consistently held

that *Daubert* is not the rule in Georgia, leaving Georgia with a very liberal test for the admissibility of scientific evidence.[123]

As the analysis in this article shows, only a minority of state courts have wholeheartedly adopted the *Daubert* trilogy. Many states adopted *Daubert* when they thought that it clearly liberalized admissibility standards for scientific evidence relative to the old *Frye* general acceptance test. Such states are not inclined to adopt *Joiner* or *Kumho Tire*, which made it clear that *Daubert* will often lead to the exclusion of evidence that would either be admitted under *Frye*, or, even more significant, evidence that *Frye* would not have applied to in the first instance. Even states that are generally sympathetic to *Daubert* and its relatively stringent reliability test have not rushed to endorse *Joiner* and *Kumho Tire*. Rather, they have adopted those parts of the *Daubert* trilogy that were consistent with their prior case law and have taken a wait-and-see approach to novel issues. Many states, meanwhile, either see *Daubert* as merely instructive or have rejected it entirely. While more state courts likely will adopt the *Daubert* trilogy as the relevant issues arise, if the trilogy is to become the standard for the admissibility of expert evidence around the country, it likely will be a result of state adoption of rules equivalent to Federal of Evidence 702, as amended in 2000.

Footnotes

[1] 509 U.S. 579 (1993).

[2] 293 F. 1013 (D.C. Cir. 1923). Under the *Frye* test, courts focused on whether the scientific principle at issue had "gained general acceptance in the particular field in which it belongs." *Id.* at 1014.

[3] *Daubert*, 509 U.S. at 593.

[4] *Id.* at 593–94.

[5] *See* Faigman *et al.*, Modern Scientific Evidence: The Law and Science of Expert Testimony §1–30, at 13 (2d ed. 2002).

[6] *Id*.

[7] For the relevant arguments, see Paul Giannelli, *The Admissibility of Novel Scientific Evidence:* Frye v. United States, *a Half Century Later*, 80 Colum. L. Rev. 1197 (1980).

[8] As Michael Saks notes, "the *Frye* test suffers from a special paradox: because less rigorous fields will reach a state of 'general acceptance' more readily than more rigorous fields, courts employing *Frye* will more readily admit the offerings of less dependable fields and less readily admit the offerings of more dependable fields." Michael J. Saks, *The Aftermath of* Daubert: *An Evolving Jurisprudence of Expert Evidence*, 40 Jurimetrics J. 229, 230 (2000); *see also* Faigman *et al.*, *supra* note 5, §1–24, at 10.

[9] Daubert v. Merrell Dow Pharms., Inc., 951 F.2d 1128 (9th Cir. 1991); Christophersen v. Allied-Signal Corp., 939 F.2d 1106 (5th Cir. 1991) (*per curiam*) (en banc); Sterling v. Velsicol Chem. Corp., 855 F.2d 1188 (6th Cir. 1988).

[10] *See*, *e.g.*, Kenneth J. Chesebro, *Taking* Daubert's *"Focus" Seriously: The Methodology/Conclusion Distinction*, 15 Cardozo L. Rev. 1745 (1994); Michael H. Gottesman, *Admissibility of Expert Testimony After* Daubert: *The "Prestige" Factor*, 43 Emory L.J. 867, 869–72 (1994).

[11] *See* Peter W. Huber, Galileo's Revenge: Junk Science in the Courtroom (1991).

[12] David E. Bernstein & Peter W. Huber, *Defense Perspective*, 1 Shepard's Expert & Sci. Evid. Q. 59, 60 (1993) ("The trend towards stricter scrutiny of scientific evidence began in the late-1980s; in the aftermath of *Daubert* it will accelerate."); David Bernstein, *Hauling Junk Science Out of the Courtroom*, Wall St. J., July 13, 1993, at A16 ("[A]s standards are established, [*Daubert*]…means that junk science will have a far harder time making it to court."); Marc S. Klein, *The Revolution in Practice and Procedure: "*Daubert *Hearings*," 1 Shepard's Expert & Sci. Evid. Q. 655, 656 (1994).

> Often overlooked is that Huber did not advocate application of the traditional *Frye* test but instead advocated "a sophisticated, modern application of *Frye* [that] looks to the methods behind a scientific report." Huber, *supra* note 11, at 200. Huber's primary concern was that courts assess the underlying reliability of scientific evidence, using scientific standards; he used *Frye* as a shorthand for strict scrutiny that would incorporate appropriate scientific standards and was not especially attached to the general acceptance test, as such. An author of this Article had worked for several years with Huber on scientific evidence issues before *Daubert* and recalls Huber's delight when *Daubert* was decided. His endorsement of *Daubert* was not "spin"—he really believed that the *Daubert* Court established a sound framework for the admissibility of scientific evidence.

[13] These commentators found support in the Court's statement that Rule 702 "requires a valid scientific connection to the pertinent inquiry as a precondition to admissibility." *Daubert*, 509 U.S. 579, 592 (1993).

[14] 522 U.S. 136 (1997).

[15] *Id.* at 146.

[16] *Id.* (citing Turpin v. Merrell Dow Pharms., Inc., 959 F.2d 1349, 1360 (6th Cir. 1992), *cert denied*, 506 U.S. 826 (1992).

[17] *See* Michael H. Gottesman, *From* Barefoot *to* Daubert *to* Joiner*: Triple Play or Double Error*?, 40 Ariz. L. Rev. 753, 755 (1998). Professor Gottesman represented the plaintiffs in *Joiner* and *Daubert*.

> By contrast, the admission of forensic evidence in criminal cases remains relatively routine. Legal commentators agree that the *Daubert* trilogy has had far less of a constricting effect on forensic science evidence compared with its effect on evidence in torts cases, most likely because defense attorneys in routine criminal cases lack the resources and expertise to challenge the admission of scientific evidence. Moreover, because all three cases in the *Daubert* trilogy arose in the civil context, lower courts seem more inclined in to overcome their traditional inertia about admitting scientific evidence in that context. In any event, the relative ease of admission of forensic science evidence does not suggest that *Daubert* is more liberal regarding such matters than is *Frye*, because forensic science evidence is also routinely admitted in *Frye* jurisdictions.

[18] *Joiner*, 522 U.S. at 142–43.

[19] *Id.* at 150.

[20] Early adopters of *Daubert* included State v. Foret, 628 So. 2d 1116 (La. 1993); Com. v. Lanigan, 641 N.E.2d 1342 (Mass. 1994); State v. Alberico, 861 P.2d 192 (N.M. 1993); State v. Hofer, 512 N.W.2d 482 (S.D. 1994); State v. Brooks, 643 A.2d 226 (Vt. 1993); Wilt v. Buracker, 443 S.E.2d 196 (W. Va. 1993); *see also* Joseph R. Meaney, *From* Frye *to* Daubert: *Is a Pattern Unfolding?*, 35 Jurimetrics J. 191, 192–93 (1995).

> These courts typically reasoned that because their own expert evidence statutes were modeled after Fed. R. Evid. 702, they should be interpreted in the same manner as the Supreme Court interpreted that rule. *See, e.g.*, *Foret*, 628 So. 2d at 1123; *Lanigan*, 641 N.E.2d at 1348–49; *Alberico*, 861 P.2d at 203 (all noting that state rules are identical to Fed. R. Evid. 702 and applying *Daubert*).

[21] Heather G. Hamilton, *The Movement from* Frye *to* Daubert*: Where Do the States Stand?*, 38 Jurimetrics J. 201, 209 (1998).

[22] *See* David E. Bernstein, Frye, Frye, *Again: The Past, Present and Future of the General Acceptance Test*, 41 Jurimetrics J. 385, 401 (2001). The *Frye* states are Alabama, Arizona, California, Florida, Illinois, Kansas, Maryland, Michigan, Minnesota, Mississippi, Missouri, New Jersey, New York, North Dakota, Pennsylvania, and Washington. *See* Courtaulds Fibers, Inc. v. Long, 779 So. 2d 198 (Ala. 2000); Turner v. State, 746 So. 2d 355 (Ala. 1998); Logerquist v. McVey, 1 P.3d 113 (Ariz. 2000); People v. Leahy, 34 Cal. Rptr. 2d 663 (Cal. Ct. App. 1994); Murray v. State, 692 So. 2d 157 (Fla. 1997); Donaldson v. Cent. Ill. Pub. Serv. Co., 767 N.E.2d 314 (Ill. 2002); Kuhn v. Sandoz Pharms. Corp., 14 P.3d 1170 (Kan. 2000); Burral v. State, 724 A.2d 65 (Md. 1999); Anton v. State Farm Mut. Auto. Ins. Co., 607 N.W.2d 123 (Mich. Ct. App. 1999); Goeb v. Tharaldson, 615 N.W.2d 800 (Minn. 2000); Kansas City S. Ry. Co. v. Johnson, 798 So. 2d 374 (Miss. 2001); Long v. Mo. Delta Med. Ctr., 33 S.W.3d 629 (Mo. Ct. App. 2000); State v. Harvey, 699 A.2d 596 (N.J. 1997); People v. Wesley, 83 N.Y.2d 417 (N.Y. 1994); City of Fargo v. McLaughlin, 512 N.W.2d 700 (N.D. 1994); Blum *ex rel.* Blum v. Merrell Dow Pharms., Inc., 764 A.2d 1 (Pa. 2000); State v. Copeland, 922 P.2d 1304 (Wash. 1996).

[23] *E.g.*, Mayhorn v. Logan Med. Found., 454 S.E.2d 87 (W. Va. 1994). Post-*Daubert* but pre-*Joiner*, many federal court decisions stated that *Daubert* was a more liberal test than *Frye*. *See, e.g.*, United States v. Bonds, 12 F.3d 540, 568 (6th Cir. 1993) ("We find that the DNA testimony easily meets the more liberal test set out by the Supreme Court in *Daubert*."). On the other hand, some *Frye* states rejected *Daubert* because they wanted to retain a purportedly stricter rule. *See*, *e.g.*, People v. Leahy, 882 P.2d 321 (Cal. 1994) (emphasizing the purported relative liberality of the *Daubert* test); Brim v. State, 695 So. 2d 268, 271 (Fla. 1995) ("Despite the federal adoption of a more lenient standard in [*Daubert*], we have maintained the higher standard of reliability as dictated by *Frye*."); State v. Carter, 524 N.W.2d 763, 778 (Neb. 1994) (referring to "the more lenient relevancy standard of *Daubert*"); Blum *ex rel* Blum v. Merrell Dow Pharms., Inc., 764 A.2d 1, 2 (Pa. 2000) (observing that *Daubert* relaxes to some extent the impediments to the admission of novel scientific evidence).

[24] *See supra* note 7.

[25] *See infra* Part I.

[26] Farm Bureau Mut. Ins. Co. of Ark. v. Foote, 14 S.W.3d 512 (Ark. 2000) (adopting *Daubert*); Coca-Cola Bottling Co. of Memphis, Tenn. v. Gill, 100 S.W.3d. 715 (Ark. 2003) (adopting *Kumho Tire* and *Joiner*).

[27] M.G. Bancorporation, Inc. v. LeBeau, 737 A.2d 513 (Del. 1999) (adopting *Daubert*, *Kumho Tire* and *Joiner*'s abuse of discretion standard); Minner v. Am. Mortgage & Guar. Co., 791 A.2d 826 (Del. Ch. 2000) (adopting *Joiner*'s scrutiny of the reasoning process).

[28] *Foret*, 628 So. 2d at 1116 (adopting *Daubert*); Darbonne v. Wal-Mart Stores, Inc., 774 So. 2d 1022 (La. Ct. App. 2000) (adopting *Kumho Tire*); Lanasa v. Harrison, 828 So. 2d 602 (La. Ct. App. 2002) (adopting *Joiner*'s abuse of discretion); Lemaire v. CIBA-GEIGY Corp., 793 So. 2d 336 (La. Ct. App. 2001) (adopting *Joiner*'s scrutiny of the reasoning process).

[29] Commonwealth v. Lanigan, 641 N.E.2d 1342 (Mass. 1994) (adopting *Daubert*); Canavan's Case, 733 N.E.2d 1042 (Mass. 2000) (adopting *Joiner* and *Kumho Tire*).

[30] Miss. R. Evid. 702.

[31] Schafersman v. Agland Coop., 631 N.W.2d 862 (Neb. 2001) (expressly adopting *Daubert*, *Joiner*, and *Kumho*).

[32] Christian v. Gray, 65 P.3d 591 (Okla. 2003) (adopting *Daubert* for civil cases); Harris v. State, 13 P.3d 489 (Okla. Crim. App. 2000) (applying *Daubert*).

[33] E.I. du Pont de Nemours & Co., Inc. v. Robinson, 923 S.W.2d 549 (Tex. 1995) (*Daubert*); Gammill v. Jack Williams Chevrolet, Inc., 972 S.W.2d 713 (Tex. 1998) (announcing test consistent with *Kumho Tire* and *Joiner*'s scrutiny of the reasoning process); Exxon Pipeline Co. v. Zwahr, 88 S.W.3d 623 (Tex. 2002) (applying standard of review consistent with *Joiner*'s abuse of discretion standard).

[34] Bunting v. Jamieson, 984 P.2d 467 (Wyo. 1999) (*Daubert* and *Kumho Tire*); Williams v. State, 60 P.3d 151 (Wyo. 2002) (*Joiner*).

[35] *See supra* note 31.

[36] Kentucky has adopted *Daubert*, *Kumho Tire*, and *Joiner*'s abuse of discretion standard. *See* Mitchell v. Commonwealth, 908 S.W.2d 100 (Ky. 1995) (adopting *Daubert*), *overruled on other grounds* by Fugate v. Commonwealth, 993 S.W.2d 931 (Ky. 1999); Goodyear Tire & Rubber Co. v. Thompson, 11 S.W.3d 575 (Ky. 2000) (adopting *Kumho Tire* and *Joiner*'s abuse of discretion standard). It has not considered whether to adopt *Joiner*'s scrutiny of the reasoning process.

[37] Ohio has adopted *Daubert*, *Kumho Tire*, and, in a court of appeals decision, *Joiner*'s abuse of discretion standard. *See* State v. Nemeth, 694 N.E.2d 1332 (Ohio 1998); Lewis v. Alfa Laval Separation, Inc., 714 N.E.2d 426 (Ohio Ct. App. 1998). It has not considered whether to adopt *Joiner*'s scrutiny of the reasoning process.

[38] North Carolina has adopted both *Daubert* and, in a court of appeals opinion, *Kumho Tire*. *See* State v. Goode, 461 S.E.2d 631 (N.C. 1995); Taylor v. Abernathy, 560 S.E.2d 233 (N.C. Ct. App. 2002). Its court of appeals has also adopted an abuse of discretion standard compatible with *Joiner*. *See* State v. Spencer, 459 S.E.2d 812 (N.C. Ct. App. 1995). It has not considered whether to adopt *Joiner*'s scrutiny of the reasoning process.

[39] Rhode Island has adopted *Daubert* and *Kumho Tire*. *See* Raimbeault v. Takeuchi Mfg. (U.S.), Ltd., 772 A.2d 1056 (R.I. 2001). It has also adopted *Joiner*'s abuse of discretion standard. *See* DiPetrillo v. Dow Chem. Co., 729 A.2d 677 (R.I. 1999). It has not considered whether to adopt *Joiner*'s scrutiny of the reasoning process.

[40] South Dakota has adopted *Daubert* and *Kumho Tire*, as well as *Joiner*'s abuse of discretion standard. *See* State v. Hofer, 512 N.W.2d 482 (S.D. 1994) (*Daubert*); Rogen v. Monson, 609 N.W.2d 456 (S.D. 2000) (*Kumho Tire* and *Joiner*'s abuse of discretion standard).

[41] *See* Baker Valley Lumber, Inc. v. Ingersoll-Rand Co., 813 A.2d 409 (N.H. 2002).

[42] *See* State v. Dahood, 814 A.2d 159, 161–62 (N.H. 2002) (concerning the admissibility of the horizontal gaze nystagmus test).

[43] *Id.*

[44] *See* State v. Streich, 658 A.2d 38 (Vt. 1995).

[45] *See* State v. Kinney, 762 A.2d 833, 841 (Vt. 2000) ("We recognize that *Daubert*, and the more recent decision in *Kumho Tire Co. v. Carmichael*…emphasized the gatekeeper function of the trial court to determine that novel scientific or technical evidence is sufficiently reliable and relevant before it is admissible.") (citation omitted).

[46] *See* State v. Coon, 974 P.2d 386 (Alaska 1999).

[47] 2002 WL 31761482 (Alaska Ct. App. Dec. 11, 2002).

[48] 2003 WL 294364 (Alaska Ct. App. Feb. 12, 2003).

[49] *Bourdon*, 2002 WL 31761482, at *7–8.

[50] *Vent*, 2003 WL 294364, at *7–8. In noting criticism of extending *Daubert* to nonscientific evidence, the court quoted Edward J. Imwinkelreid's comment that "although the *Daubert* Court may have selected the optimal test for the admissibility of scientific evidence, that test is useless as a criterion for the admissibility of other types of expert testimony." *Id.* at *7 (quoting Edward J. Imwinkelreid, *The Next Step After* Daubert*: Developing a Similarly Epistemological Approach to Ensuring the Reliability of Non-scientific Expert Testimony*, 15 Cardozo L. Rev. 2271, 2285 (1984)).

[51] *See* State v. Alberico, 861 P.2d 192 (N.M. Ct. App. 1993).

[52] *See* Banks v. IMC Kalium Carlsbad Potash Co., 62 P.3d 290 (N.M. Ct. App. 2001). In determining whether New Mexico's case adopting *Daubert* should apply to scientific evidence in a proceeding before the Workers Compensation Administration, the Court of Appeals noted that "[t]he *Alberico/Daubert* analysis is not limited to novel scientific theories" and cited *Kumho Tire* for this proposition. *Id.* at 293.

[53] *See* State v. Torres, 976 P.2d 20 (N.M. 1999) (explicitly rejecting *Joiner*'s abuse of discretion standard in favor of a *de novo* review of the decision whether to apply *Daubert* and a discretionary review of the findings).

[54] *See* State v. Porter, 698 A.2d 739, 746 (Conn. 1997).

[55] Doe v. Thames Valley Council for Com. Action, Inc., 797 A.2d 1146, 1161–62 n.21 (Conn. App. Ct. 2002).

[56] Conn. Ct. R. §7–2 (West 2003).

[57] *See* State v. Pappas, 776 A.2d 1091 (Conn. 2001).

[58] The only mention of *Joiner*'s scrutiny of the reasoning process is *in dicta* in an unpublished Connecticut Superior Court opinion, *K&V Scientific Co. v. Ensign-Bickford Co.*, 2002 WL 31662326, at *2 (Conn. Super. Ct. Nov. 6, 2002).

[59] 899 P.2d 663 (Or. 1995).

[60] *Id.* at 680 n.7.

[61] State v. Sanchez-Cruz, 33 P.3d 1037 (Or. Ct. App. 2001).

[62] Jennings v. Baxter Healthcare Corp., 14 P.3d 596, 604 (Or. 2000) (quoting State v. Brown, 687 P.2d 751, 775 (Or. 1984)).

[63] *Id.* at 606.

[64] *See* Wilt v. Buracker, 443 S.E.2d 196 (W. Va. 1993).

[65] 516 S.E.2d 769 (W. Va. 1999).

[66] *Id.* at 774 n.4 (stating "[w]e decline to adopt the *Kumho* analysis in this case").

[67] 545 S.E.2d 294 (W. Va. 2001).

[68] *Id.* at 301.

[69] Davis stated that:

> At this time, the majority declines to expressly address whether we will adopt the new federal procedure regarding expert testimony. However, the author of this opinion, separate from the majority, does not believe that *Kumho* would be a death knell to the admission of non-scientific expert testi-

mony. Indeed, *Kumho* has been approved by a majority of state courts who have taken it under consideration.

The author of this opinion believes that it is the restrictive interpretation of *Kumho* anticipated by some commentators that is causing confusion. However, there are two specific reasons that *Kumho does not* realistically present any new barrier to the admissibility of expert testimony that is based on technical or other specialized knowledge. First, the *Kumho* test is a flexible one that does not require application of the specific factors suggested in *Daubert*, which were also intended to be applied flexibly… . Second, *Kumho*, as an extension of *Daubert*, applies only to expert testimony that is not subject to judicial notice.

Id. at 301 n.11 (citations omitted).

[70] 569 S.E.2d 133 (W. Va. 2002).

[71] *Id.* at 143 n.21.

[72] *See* State v. Moore, 885 P.2d 457 (Mont. 1994).

[73] *See* State v. Cline, 909 P.2d 1171 (Mont. 1996).

[74] *Id.* at 1177.

[75] *See id.* at 1178 (examining the district court's admission of novel scientific testimony for abuse of discretion).

[76] Ala. Code §36-18-30 (West 2002).

[77] *See* AAA Cooper Transp. v. Philyaw, 842 So. 2d 689, 690 n.1 (Ala. 2002).

[78] *See* McDaniel v. CSX Transp., Inc., 955 S.W.2d 257, 265 (Tenn. 1997) (stating that "[a]lthough we do not expressly adopt *Daubert*, the non-exclusive list of factors to determine reliability are useful in applying our Rules 702 and 703").

[79] *See* State v. Stevens, 78 S.W.3d 817 (Tenn. 2002) (adopting *Kumho Tire*'s analysis and *Joiner*'s scrutiny of the reasoning process, as well as providing for abuse of discretion review consistent with *Joiner*).

[80] *See* Acoba v. General Tire, Inc., 986 P.2d 288, 300 n.6 (Haw. 1999) ("To date this court has neither expressly approved nor rejected the *Daubert* criteria. We decline to do so at this time."); State v. Vliet, 19 P.3d 42, 53 (2001) (stating "[t]he prosecution is correct in contending that this court has not adopted the *Daubert* test…and we expressly refrain from doing so").

[81] *Vliet*, 19 P.3d at 53.

[82] *Id.* at 54–55.

[83] *See id.* at 56–57 (stating that "[w]ithin this framework, we do not consider it essential or necessary that a trial court embark upon a preliminary determination of whether the proffered expert testimony should be characterized as scientific, technical, or otherwise specialized knowledge").

[84] *See id.* at 55–56 n.24 (explaining its partial adoption of a standard consistent with *Joiner*, but noting that it differs from *Joiner* as to the standard to be applied to ultimate conclusion of admissibility).

[85] McGrew v. State, 682 N.E.2d 1289, 1290 (Ind. 1997). Indiana Rule of Evidence 702(b), unlike the pre-2000 Federal Rules of Evidence, requires that expert scientific evidence may not be admitted unless "the court is satisfied that the scientific principles upon which the expert testimony rests are reliable." Ind. R. Evid. 702(b) (West 2003).

[86] *See* Malinski v. State, 794 N.E.2d 1071, 1085 (Ind. 2003).

[87] *See* Lytle v. Ford Motor Co., 696 N.E.2d 465, 472 (Ind. Ct. App. 1998) (applying, without further comment, *Joiner*'s scrutiny of the reasoning process).

[88] *See McGrew,* 682 N.E.2d at 1292.

[89] Leaf v. Goodyear Tire & Rubber Co., 590 N.W.2d 525, 533 (Iowa 1999).

[90] *Id.* at 532–33.

[91] *See id.* at 533.

[92] *See* People v. Shreck, 22 P.3d 68, 77 (Colo. 2001).

[93] *Id.* at 77–78.

[94] *See id.* at 73–79. The rule stated by the Court in *Shreck* is stated as the rule for scientific evidence in Colorado. *Id.* at 73. The court in *Shreck* did refuse to limit the rule to "novel" scientific evidence. *Id.* at 78 n.12.

[95] *See id.* at 78–79 (discussing the liberal standards and discretion of the court in determining reliability and relevance).

[96] *See* State v. MacDonald, 718 A.2d 195, 198 (Me. 1998).

[97] *See* Green v. Cessna Aircraft Co., 673 A.2d 216, 218–19 (Me. 1996) (citing *Daubert* regarding the admissibility of expert testimony about failure of parts on an aircraft); *see also* State v. Tomah, 736 A.2d 1047, 1055 (Me. 1999) (stating that expert evidence must meet Rule 702).

[98] *MacDonald*, 718 A.2d at 198.

[99] *See* State v. Merwin, 962 P.2d 1026, 1029–30 (Idaho 1995).

[100] *See*, *e.g.,* State v. Siegal, 50 P.3d 1033, 1042–43 (Idaho Ct. App. 2002); State v. Parkinson, 909 P.2d 647, 652 (Idaho Ct. App. 1996).

[101] Carnell v. Barker Mgmt., Inc., 48 P.3d 651, 656–57 (Idaho 2002).

[102] *See* Kemp *ex rel.* Wright v. State, 809 A.2d 77 (N.J. 2002). New Jersey continues to follow the *Frye* test in criminal cases. *E.g.*, State v. Harvey, 699 A.2d 596, 621 (N.J. 1997).

[103] *Kemp*, 809 A.2d at 86.

[104] *See Harvey*, 699 A.2d at 620 (noting the Eleventh Circuit's decision in *Joiner* but coming to no conclusion).

[105] *See* Hamilton v. Oppen, 653 N.W.2d 678, 683 (N.D. 2002).

[106] *Id.* (quoting Anderson v. A.P.I. Co. of Minn., 559 N.W.2d 204, 206 (N.D. 1997)).

[107] *See id.* (stating that qualification of testimony is discretionary).

[108] *See* State v. Council, 515 S.E.2d 508 (S.C. 1999).

[109] *Id.* at 517.

[110] *Id.* at 517–18.

[111] *Id*.

[112] *See* State v. Crosby, 927 P.2d 638 (Utah 1996).

[113] *Id.* at 642.

[114] State v. Rimmasch, 775 P.2d 388, 403 (Utah 1989).

[115] *Id.*

[116] Alder v. Bayer Corp., AGFA Div., 61 P.3d 1068, 1083 (Utah 2002) (quoting State V. Rimmasch, 775 P.2d 388 (Utah 1989)).

[117] *Id.* at 1084.

[118] See Satcher v. Commonwealth, 421 S.E.2d 821, 835 (Va. 1992) (quoting Spencer v. Commonwealth, 393 S.E.2d 609, 621 (Va. 1990)).

[119] *See* John v. Im, 559 S.E.2d 694, 698 (Va. 2002) ("We note, however, that we have not previously considered the question whether the *Daubert* analysis employed by the federal courts should be employed in our trial courts to determine the scientific reliability of expert testimony. [Footnote omitted.] Therefore, we leave this question open for future consideration.").

[120] Conley Publ'g Group Ltd. v. Journal Communications Inc., 665 N.W.2d 879, 892 (Wis. 2003).

[121] *See* State v. Walstead, 351 N.W.2d 469, 487 (Wis. 1984).

[122] Dow Chem. Co. v. Mahlum, 970 P.2d 98, 108 n.3 (Nev. 1998).

[123] Orkin Exterminating Co. v. Carder, 575 S.E.2d 664, 669 (Ga. Ct. App. 2002); Norfolk S. Ry. Co. v. Baker, 514 S.E.2d 448, 451 (Ga. Ct. App. 1999).

Expert Information and Expert Evidence

A Preliminary Taxonomy

Jennifer L. Mnookin

Professor of Law
University of Virginia School of Law

Visiting Professor
Harvard Law School

Samuel R. Gross

Thomas and Mabel Long Professor of Law
University of Michigan Law School

The authors would like to thank D. Michael Risinger and the participants at the symposium on Expert Admissibility: Keeping Gates, Goals and Promises, Seton Hall University School of Law, February 21–22, 2003. The author would like to thank Faith Shippen for helpful research assistance.

JENNIFER L. MNOOKIN, PH.D. recently joined the law faculty of the UCLA School of Law. Previously she was Professor and Barron F. Black Research Professor at the University of Virginia School of Law, and Visiting Professor of Law at Harvard Law School during the 2003–2004 school year. An expert on evidence law, Mnookin focuses on scientific, forensic, and visual evidence in her scholarship. She has written articles on fingerprinting and its origins, the history of handwriting identification evidence, the effects of photography on the 19th-century criminal justice system, and the early use of film as legal evidence. She is a co-author of *The New Wigmore: Expert Evidence* (Aspen Law and Business 2003). Mnookin received her A.B. from Harvard and her J.D. from the Yale Law School, where she was senior editor of the *Yale Law Journal*. She also earned a Ph.D. in the History and Social Study of Science and Technology at MIT. Professor Mnookin has taught a variety of courses, including evidence, scientific evidence, torts, and law and literature. In summer 2003, she taught scientific evidence to the judges participating in UVA's LL.M. program in Judicial Process.

SAMUEL R. GROSS, the Thomas and Mabel Long Professor of Law, graduated from Columbia College in 1968 and earned a J.D. from the University of California at Berkeley in 1973. He was a criminal defense attorney in San Francisco for several years, and worked as an attorney with the United Farm Workers Union in California and the Wounded Knee Legal Defense Committee in Nebraska and South Dakota. As a cooperating attorney for the NAACP Legal Defense and Educational Fund Inc. in New York and the National Jury Project in Oakland, California, he litigated a series of test cases on jury selection in capital trials and worked on the issue of racial discrimination in the use of the death penalty. He was a visiting lecturer at Yale Law School and came to the University of Michigan from the Stanford Law School faculty. Professor Gross teaches evidence, criminal procedure, and courses on the use of the social sciences in law. His published work focuses on the death penalty, false convictions, racial profiling, eyewitness identification, the use of expert witnesses, and the relationship between pretrial bargaining and trial verdicts.

Expert Information and Expert Evidence

A Preliminary Taxonomy

Table of Contents

Expert Information and Expert Evidence

A Preliminary Taxonomy

I. Introduction

Federal Rule of Evidence 702 speaks in very general terms. It governs every situation in which "scientific, technical or other specialized knowledge will *assist* the trier of fact," and provides that, in that situation, "a witness *qualified* as an expert by knowledge, skill, experience, or education, may testify thereto in the form of an opinion or otherwise… ."[1] In 2000, following a trio of Supreme Court cases interpreting Rule 702, the Rule was amended to include a third requirement, in addition to the helpfulness of the testimony and the qualifications of the witness: *reliability*. Under Rule 702 as amended, a qualified witness may only provide expert testimony "if (1) the testimony is based upon sufficient facts or data, (2) the testimony is the product of *reliable* principles and methods, and (3) the witness has applied the principles and methods *reliably* to the facts of the case."[2]

In the wake of that trio—*Daubert v. Merrell Dow Pharmaceuticals, Inc.*,[3] *General Electric Co. v. Joiner*,[4] and *Kumho Tire Co. v. Carmichael*[5]—we have all tended to focus our attention on the third and most recent of these requirements for expert evidence: reliability. Literally thousands of pages have been written about both the proper criteria for evaluating the reliability of expert evidence and the institutional competence of judges to evaluate scientific reliability.[6] Moreover, since Rule 702 is a rule of admissibility, commentators write and talk about reliability primarily as a threshold question: is the evidence reliable enough to be admissible? That question, naturally, is only interesting in cases in which the expert evidence seems to be comparatively unreliable—as that term is understood and applied—and therefore potentially inadmissible. In each of this trilogy of leading cases, for example, the expert evidence at issue was excluded at trial as unreliable, and in each case the exclusion was affirmed.[7]

Many have suggested that the problem of assessing reliability is especially acute with respect to non-scientific expert evidence.[8] At least compared to alternative forms of knowledge-production, research science involves formalized methodological norms, articulated standards, and conscious research design. By contrast, many forms of potential expert knowledge—from the clinical doctor's diagnosis to the historian's description to the tire safety expert's analysis—are based on experience, tacit knowledge, even hunch.[9] Evaluating the reliability of knowledge not produced through formal methods thus raises especially difficult questions. As important as an examination of method, however, and much less noted, is another dimension: the degree of certainty that the expert posits in what she offers. One of the central problems with much expert testimony introduced in court—both scientific and non-scientific alike—is that experts claim as matters of fact or probability opinions that should be couched in more cautious terms, as possibilities or hypotheses.

We believe that the monocular focus on issues of scientific method and reliability has obscured some broader points; therefore, in this Article we try to step back (gingerly) and take a broader view. Instead of beginning with the problems of reliability, we start by briefly detailing the array of informational issues facing a consumer of expert evidence, thereby putting the

attention-getting problems of reliability into a broader context. We then attempt to review and classify the full range of expert testimony, much of which is unproblematic, AND some of which *is* problematic but routinely admissible nonetheless. We offer first a brief taxonomy, an outline with examples. We then develop our classification scheme by exploring each kind of expert statement in more detail. In these discussions, we purposefully do not distinguish between scientific knowledge and other forms of knowledge, instead framing our taxonomy around specific kinds of statements and their functions.

Our taxonomy has four purposes. First, we want to emphasize the extraordinary range of information that is presented in court by expert witnesses. Some of this information is 'scientific' and some of it is 'non-scientific'; but often, whichever categories it falls in, it is wholly unproblematic. Many if not most expert witnesses testify without objection and present information that may be critical for the factfinder but is not in dispute. Second, for many categories of expert evidence, even when there may be some degree of controversy or disagreement, there really is no *Daubert* or *Kumho Tire* problem. For certain categories of expert information, a focus on the credentials of the expert is generally sufficient, and courts need not (and typically do not) make any additional elaborate inquiry into validity. Third, our taxonomy reminds us of the continuing importance within the evaluation of expert testimony of more mundane credibility issues than 'reliability' in the *Daubert* sense: specifically, bias, competence, and lack of clarity. Finally, our discussion reveals the sometimes overlooked importance of paying attention not only to what the expert says, but to how she says it. Often, whether testimony is based on scientific study or more casual forms of observation, what makes an expert's conclusion unreliable is that it is expressed with a confidence not warranted by the evidence. The clinical observations of a physician or an engineer or a mechanic or a fingerprint examiner may be quite appropriate as the basis for testimony, but the degree of certainty expressed by the witness should reflect both knowledge and its limits, both what is known and what is not.

II. The General Nature of Expert Evidence

A. The Task of Providing Expert Information: Doing and Telling

"Expert evidence" is a species of the genus "expert information." In everyday life, we rely on experts constantly—doctors, lawyers, carpenters, mechanics. Most of the time we want them to do things—set bones, write wills, build walls, fix brakes—not talk about them. Most of us have learned the hard way that mechanics and contractors (if not doctors) may have pleasing personalities and give clear, plausible explanations, but do lousy work. Still, we usually want these experts to tell us what they are doing, even if that is not their main task, and sometimes that information is crucial since we, the non-expert consumers, must make critical decisions. It is that informative function that interests us here, since expert *witnesses* only provide information, evidence; what they *do* in court is *tell*.

Sometimes an expert might be able to answer a question immediately, from existing knowledge: "Anthrax is a life-threatening bacterial disease that cannot be transmitted directly from one infected person to another." For other questions, she may need to collect additional information, in one or both of two categories: (1) *data about the particular case*, ranging from the minimal ("let me take a quick look") to the extensive (collecting archival data on hiring and compensation in a large company and then running a set of statistical analyses); (2) *general*

knowledge on the topic, which can mean anything from checking a standard reference to an original scientific study that adds to the body of general knowledge in the field and may have value beyond the lawsuit. Whatever the scope of the work, the structural significance is the same: The *telling* part of an expert's task—for an expert *witness*, this is the only job—often requires some preliminary *doing*: study.[10]

Testimony based on deliberate study is not the exclusive preserve of expert witnesses. Lay investigators who set out to discover what happened on a particular occasion may also testify to some of what they find out. This category includes not only the prototypical police and private investigators, but also any other witnesses—most often, parties or their employees—who deliberately collect information about significant past events. What makes studies by expert witnesses different is the permissible scope of testimony describing those studies. The lay investigator who "solves" a murder may testify about the tracks he saw in the dust and the shell casings he found in the defendant's car, but he may not relate most of the hearsay statements he heard when conducting interviews, and he may not give his own interpretation of the evidence. By contrast, the pathologist who testifies to the cause of death may freely rely on a wide range of second-hand information in making up her mind, may testify in detail to her opinions, and may be permitted to describe a wide range of otherwise inadmissible evidence along the way.[11] Unlike a lay investigator, an expert may generate and testify to new information that did not exist before by conducting tests (or occasionally even studies) herself.

B. Issues for Consumers of Expert Information: Validity, Competence, Clarity, and Bias

1. Validity

The most basic question about expert information is whether it is within a valid category of expertise: Is there a field of knowledge that has credible tools to produce valid answers to questions such as this? Can astrologers determine personality characteristics from the time and place of a person's birth? Can forensic scientists determine a person's identity from bite marks? Can biologists determine identity from analysis of DNA? In each case, the issue is not the competence of the particular analysis but the legitimacy of the discipline's claim to be able to provide this sort of information. This aspect of validity—*field validity*—is fundamental but usually uncontroversial. It is important in court because it defines the boundaries of permissible expert testimony—biologists will be allowed to testify but astrologers will not—but most evidence that is offered is well within the recognized borders. The debates preceding and following *Daubert* have highlighted the importance of field validity, and several accepted forms of expert testimony—from psychiatric evaluations to handwriting identification—have been challenged on this basis.[12]

Moving down a level of abstraction takes us to a second aspect of validity: *method validity*. Given that a discipline is accepted as a legitimate category of expertise on a particular topic, are the methods that were used in this instance capable of producing valid answers? The present version of Rule 702, codifying *Daubert* and *Kumho Tire*, makes method validity a prerequisite for expert evidence: expert testimony is only admissible if it is "*the product of reliable principles and methods.*"[13] In *Daubert* itself, for example, no one doubted the *field* validity of epidemiology, but the lower courts went on to hold nonetheless that the particular *method* used by the plaintiff's experts—reanalysis of data from published studies—was not a valid basis for testimony by quali-

fied epidemiologists.[14] Method validity addresses the specific techniques deployed by the expert: Is PCR a valid procedure for comparing samples of DNA? Is differential diagnosis a valid method to determine the cause of a disease? *Daubert, Kumho Tire*, and *Joiner* all concern method validity.[15] In their wake, objections on this basis have become more common and more successful, particularly so in civil cases.[16] In run-of-the-mill cases, however, expert evidence continues to follow well-worn and well-accepted paths.[17]

Method validity and field validity are obviously linked—a field is invalid to the extent that its methods fail to do what they claim—but the distinction is important. Field invalidity implies method invalidity, but the converse is not the case. Few question the value of medicine as a field, and few doubt that some medical doctors make foolish medical claims, based on unsound, if not specious, methods. One of the most difficult tasks for lay judges in evaluating expert evidence is to distinguish between acceptable and unacceptable practices by qualified members of a legitimate discipline. One solution is to pass the buck back to the expert field itself, and accept the standards it imposes on itself; this is the logic behind the *Frye* standard of "general acceptance" for novel scientific techniques.[18] However (ignoring for the moment other difficulties with *Frye*), this dodge does not work when the validity of the field itself is in question.[19]

2. Competence

Assuming that expert information reflects sufficiently valid methods within a legitimate field, we need to know whether a particular statement is a competent example of its genre. This issue has two components: (1) *Qualifications*: Does the expert at hand have the knowledge and skill that are necessary to produce the information? Rule 702 makes this a requirement: to testify as an expert, a witness must be "*qualified*…by knowledge, skill, experience, training, or education…";[20] and (2) *Execution*: Did the expert (assuming she can ever do so) perform competently on this occasion? Even a qualified expert working in a recognized and valid methodology can do an incompetent job in a particular case. Rule 702, following *Kumho Tire*,[21] requires that expert testimony be "based upon sufficient facts or data," and that the expert have "*applied* the principles and methods [she used] *reliably* to the facts of the case."[22] Until recently, American courts were notoriously forgiving on qualifications, and probably almost equally so on execution, as preconditions for admissibility. The operating theory was that in most cases these are issues that go to the weight of the evidence. There are some indications that courts have become more exacting in recent years, though the cases are by no means consistent on this point.

3. Clarity

You can only make use of information to the extent that you understand it. When the informant is an expert—someone whose career is devoted to arcane information—the problem can be acute, as most of us know from common experiences with builders, accountants and computer technicians. Clarity is not generally an important factor in determining the admissibility of expert testimony in court. In theory, lack of clarity could be a basis for an objection to expert evidence under Rule 403 because its "tendency to confuse" substantially outweighs its probative value,[23]or under Rule 702 on the ground that the evidence is too confusing to "assist" the trier of fact.[24] In practice, these objections are unlikely to succeed—or even to be made. Because clarity seems to have great influence on the evaluation of experts by juries and judges, we are generally willing to rely on the self-interest of the parties to produce clear expert evidence in court. In fact,

a special danger associated with expert testimony is that the parties and their experts will sacrifice accuracy for the sake of appealingly clear but erroneous or over-simplified presentations.[25]

4. Bias

If we understand what an expert tells us, and know that she has correctly applied a valid method to the pertinent facts, our problems are at an end. In most situations, however, we cannot know these things directly for the simple and obvious reason that we ourselves are not experts in the relevant field. Instead, we use various indirect measures of the quality of expert information, of which the most pervasive, in court and out, is bias. What is the first question that comes to mind when a mechanic at a highway road stop tells you that you need two new tires, immediately? As with clarity, bias is not a common basis for excluding expert testimony. In our system, the evaluation of bias is a core function and special prerogative of juries as triers of fact; therefore we do not exclude witnesses who may be biased, but allow juries to weigh that factor in judging their evidence. Our practice for experts is, in this respect, basically the same as for lay witnesses. We systematically neglect well-considered plans for the use of unbiased (or at least, non-partisan) expert testimony as a supplement to potentially biased party-sponsored expert evidence.[26]

Conceptually, bias has three components: (1) Is the *field* biased? Do chiropractors, *as a group*, always say that the problem is a misalignment of the spinal column? This is not a common issue. In an extreme case, evidence of this sort could conceivably be the basis for an objection that evidence from such an expert is invalid—in the terminology of this Article, that the area of knowledge lacks field validity—or that experts trained in that discipline are unqualified to testify. (2) Is this *person* biased across a range of cases? Is she an expert who always says that the plaintiff had a pre-existing condition, or that toxic exposure caused the disease? In extreme cases, this could be an argument against admissibility on the theory that the expert is unqualified; in practice this argument almost always "goes to the weight" of the expert's evidence. (3) Is the expert's *performance* biased? Does the expert have a grudge against this defendant, or, more likely, was she paid a handsome fee to say what she said in this case? That too is a type of bias that is considered appropriate as a basis for discounting an expert's evidence, but not for excluding it.

…

At this point, a note about vocabulary is in order. *Daubert* and Rule 702 talk extensively about "reliability."[27] In the sciences, *reliability* is often used to mean "consistency" or "repeatability."[28] A test that produces the same result on successive applications is said to be *reliable*, while a test that produces accurate results is said to be *valid.*[29] The Court in *Daubert* uses "reliability" instead to mean something closer to "validity" or "trustworthiness."[30] What the *Daubert* Court calls "reliability" blends our categories of *validity* and *competence*, focusing especially on method validity and competence of execution. To be consistent with *Daubert* and Rule 702, in this Article we too use *reliability* to mean "trustworthiness": If evidence is *reliable*, we can trust it, we can rely on it.

III. A Taxonomy of Expert Information

When an expert's job is to inform, what sort of information do we want? The scheme that follows is a first cut. The categories we list are not entirely analytically distinct, and a single expert will typically do more than one of these tasks in a given project. We do not divide expert

testimony by field, or even by methodological approach. Our point, rather, is to note that there are several different *types* of information that we commonly seek from experts—in court and out—and to try to describe that range. In the scheme that follows, categories of expert information are arranged, more or less, along a rough continuum from specific "factual" bricks that the consumer may use to build a wall, to comprehensive general "conclusions" that the lay listener must accept or reject as a whole. We start with a capsule form of the taxonomy—an outline with examples—and proceed to a more detailed discussion in which we try apply the general issues we have touched on—validity, competence, clarity, and bias—and consider the factors that bear on the admissibility of such expert information in court.

A. Capsule Form	
1. Description:	
(a) Observation:	"The brake shoes are worn to within 1 mm of the bracket." "This x-ray shows multiple fractures of the tibia."
(b) Translation:	"In English, that means 'I agree.'" "As used in this stock exchange, '5p@36' means 'I offer to sell 5000 shares of preferred stock at $36 per share.'" "The term 'nick' refers to a five dollar bag of cocaine or crack."
(c) Calculation:	"The total payments over the 10 year term of the loan, with compound interest at 7 percent per annum, will come to $789,566." "A racial disparity this large or larger would occur by chance alone less often than 1 time in 100,000." "Total indebtedness, as of 9/10/2001, was $2,896,755."
2. Instruction:	
(a) Facts:	"Blood has a pH value of 7.4." "Mitochondrial DNA is transmitted entirely from a mother to her offspring." "Most of the loss of eyewitness memory occurs in the first few hours after an observation." "Bendectin is not a teratogen."
(b) Norms:	"The customary treatment for premature labor is bed rest and a tocolytic agent." "Epidemiological study is the most persuasive way to establish a correlation between ingestion of a substance and a disease."
3. Assessment:	
(a) Condition: (diagnosis)	"Your chimney is unsafe." "The patient is suffering from rheumatoid arthritis."
(b) Causation:	"The skull fracture may have been caused by a blow from a heavy object." "Tire failure was probably caused by a defect in the fabrication of the tire." "The tumor was caused by exposure to dioxin."

(c) Consequences: (prognosis)	"The patient will be permanently paralyzed from the fifth lumbar vertebra down." "If current levels of phosphates are maintained, all frog populations in the basin will be extinct within ten years." "You need a new chimney."
(d) Identity: (common source)	"The latent fingerprints on the knife came from the defendant's right hand." "It's a forgery."
(e) Value: (or quality)	"The present value of the plaintiff's expected lifetime earnings before the accident was approximately $1.6 million." "This house has market value between $375,000 and $405,000." "This play has literary merit."

B. Fuller Discussion

1. Description

Much information from experts is primarily descriptive rather than evaluative. Expert testimony focuses heavily on opinions, sometimes elaborate opinions, but, as Rule 702 recognizes, experts may testify "otherwise."[31] The boundary between description and assessment is fuzzy, and expert opinions are frequently, perhaps invariably, built into the foundations of their descriptive statements. All the same, the distinction is meaningful and useful. In this Article we discuss three types of descriptive statements: observation, translation, and calculation.

Field validity is not a major concern for evaluating expert information that is primarily descriptive, although it could be. A specialty that claims to be able to observe people's thoughts would run into deep trouble. Questions of method validity frequently figure in the background—in decisions on what to look for, for example (X-ray the neck, or the entire spine?), or what to calculate (Is chi-square an appropriate statistic test of the independence of the distributions of two variables?), or how to calculate it (Can one use the SAS statistical package to compute chi-squares?). And competence, of course, is crucial.

a. Observation

Much of what experts learn to do is see and hear things that the rest of us miss. The degree and nature of the skill involved varies greatly from one type of expertise to another, but the advantages that many experts have in perception are well known, from umpires (good ones) and coaches, to orchestra conductors, to optometrists. An expert's observation may be uncontroversial—the expertise involved may consist entirely of identifying the right object (brake shoe) and measuring it with the right tool (a caliper)—or it might depend on an exercise of judgment that other experts might dispute ("Strike!"). One way or another, the question that this sort of expertise answers is: "What did you perceive—see, hear, feel or smell?" In court, expert observations—describing pathologies or injuries to people for example, or damage to structures—are often crucial, either as the crux of expert testimony or as part of the foundation for evaluative opinions. And these observations are sometimes controversial since experts, like everybody else, may disagree about what they see.

The distinction between observation and assessment is one of degree. The nurse who says that a patient's blood pressure was 135/85 is *interpreting* sensory information—what she saw on the pressure gauge and heard on her stethoscope—not reporting it. Moreover, her interpretation incorporates a host of unstated assumptions about the tools she used and the procedure she followed. A doctor who says that a patient is suffering from rheumatoid arthritis is also interpreting information, but the range of information is likely to be greater (multiple tests, reported and observed symptoms, family history) and much of it probably consists of second- and third-hand accounts. We classify the first statement (blood pressure) as an expert *observation*, the second (arthritis) as an *assessment*. Some cases may be harder to call, but the dividing line is less important than the underlying issue: To what extent does the expert integrate information about the particular case from a variety of sources, rely on reports from others, and make complex analytic judgments?

In general, the first thing we want to know about an expert observation is the competence of the observer: How good is she at this task? Since we do not usually have direct measures of skill, we tend to rely on proxies: How good was her training? Does she use accepted techniques in an apparently competent manner? Out of court, we try to use experts who have, in our own experience, provided accurate information in the past; failing that, we look for those who are said to be accurate by friends and acquaintances, or who are reputed to be accurate. In court, we count on testimony on qualifications and on cross-examination to provide evidence on the expert's training and experience.[32]

Bias, of course, may be as important as competence. The chimney sweeper who tells you (as they all do) that there are ten-inch cracks in your chimney, is probably also a contractor who rebuilds chimneys. The key to minimizing bias, out of court, is to remove the incentives to distort—for example, by consulting a diagnostician who knows he will have no role in the treatment. In court, we neglect the most effective method of minimizing bias—using non-partisan experts—and rely instead on cross-examination and rebuttal to expose it.[33]

The primary foundational requirement for the admission of observational expertise under existing rules is evidence that the expert is "qualified," which bears on her competence. On this issue, the treatment of different sorts of experts is likely to diverge. If the field has elaborate formal qualifications—in particular, if it requires graduate or professional education, and/or certification—courts typically accept these formal qualifications as sufficient evidence of competence.[34] Most, perhaps all scientists fall into this category, along with many non-scientists. On the other hand, if the skill is based primarily or exclusively on "experience"—the harbor pilot or the chicken sexer—courts may demand concrete evidence from past practice that the expert can perform this task accurately.[35]

By far the most effective way to minimize the danger of error in observational evidence of any sort is to reproduce the observation. For lay witnesses, we rarely have that luxury; for experts, we can often ask another expert to do it again. If it matters enough, we have a second doctor examine the X-rays, or a second lab repeat the test. In litigation, this task may be performed by an expert for the side opposed to the party that called the first expert. When this is possible, the adversarial system probably performs quite well, on the whole, in producing high quality evidence—provided that the advocates for both sides have sufficient resources and motivation to do a good job. This is a significant qualification.[36]

The worst problems with expert descriptions occur in criminal prosecutions. Most experts in criminal cases are state employees who are called by prosecutors: police officers, medical examiners, technicians at state crime laboratories, and so forth. Many of them are inadequately trained, and some are unscrupulous. Their testimony is rarely subjected to review by defense experts; in many cases, they are hardly even cross-examined by the inexperienced, overworked, and underpaid court-appointed defense attorneys. Any witness can make avoidable mistakes or lie, but an expert can make a career of it. In recent years we have seen examples of this repeatedly in criminal prosecutions across the United States: a pathologist who testified to conducting hundreds of autopsies on bodies he never touched,[37] forensic scientists who made up findings in hundreds of cases to suit the police and prosecutors,[38] and crime labs that were riddled with incompetence.[39] We have no way of knowing how frequently similar problems go undetected, but by any accounting they are by far the worst misuse of expertise in American courts.[40] This is an avoidable scandal. If we are going to rely on the adversary system to guarantee competence and honesty in expert evidence we must actually have an adversary process. When one side is absent, the result may be disastrous.

b. Translation

Almost all expert information involves some form of "interpretation." A radiologist who describes an X-ray as showing a bone fracture is "interpreting" the X-ray; for that matter, a layperson who says that the defendant "agreed" with the policeman is "interpreting" the words the defendant spoke, or the nod of his head. We use "translation" in a narrower sense, to refer to the restatement in one system of symbolic communication (in American courts, everyday English) of a message that was conveyed in a different system of communication.[41] In this case, the question for the expert is some version of "What did she say?"

The archetypal "translation" is from one common language into another. This is no doubt the common type of expert translation in American courts, but other sorts of translation are also used with some regularity. Perhaps the most controversial is expert evidence, generally from a police officer, about the meaning of particular words and phrases in some esoteric underworld community, typically the world of drug dealers.[42] Similar interpretive tasks, however, occur in any context in which a comparatively small group of insiders in a profession or trade develops a specialized jargon for the issues they deal with on a regular, repetitive basis: You may need an expert in the diamond trade to tell you what a sentence from a diamond trader means,[43] or a pharmacist to interpret what your internist wrote on a prescription form.[44] As with expert observation, translation (by the expert herself, or by another expert) is frequently a foundational element of more evaluative expert testimony even when translation is not a central or explicit task.

The main question for a translator is the same as for an expert observer, competence: How good is she at her job? Other than common intelligence, the main requirements for competence are proficiency in the two systems of communication that need to be made mutually intelligible; a Russian language interpreter in an American court must be able to understand and communicate effectively in both Russian and English. In addition, of course, a biased translator can distort the message he transmits. As with expert observers, the best check on mistranslation is replication by a second translator who is not likely to share the weaknesses and biases of the first.

The Federal Rules of Evidence recognize that translation is a form of expert testimony, but nonetheless have a special provision for interpreters who perform this function, Rule 604:

"An interpreter is subject to the provisions of these rules relating to qualification as an expert and the administration of an oath or affirmation to make a true translation."[45] In practice, this means that interpreters who provide evidence (as opposed to those who translate proceedings for the benefit of non-English speaking participants) are treated schizophrenically, depending on the context. Live testimony cannot proceed if there is a dispute about the meaning of the spoken words. Therefore, an interpreter who translates testimony from the witness stand—unlike almost every other expert witness in an American court—is selected by the judge, preferably from an official list of pre-qualified experts, takes the special oath provided by Rule 604 (if they are in federal court) and provides (at least *prima facie*) authoritative evidence on the meaning of the spoken words.[46] The use of official interpreters no doubt greatly reduces the danger of biased translation. If the official certification program works as intended, it may also help guarantee competence.

A different set of rules applies to an interpreter who translates words that were spoken or written at some earlier time, as memorialized in a document or an electronic recording. In that situation, the translator is treated as an ordinary "language expert [under Rule 702] who [takes] the stand under oath, and subject[s] himself to cross-examination. Rule 604 is inapplicable."[47] Since immediate, authoritative translation is not a functional necessity when the words are recorded, the usual rules for expert witnesses apply, and competing experts may offer different interpretations of statements in a foreign language.[48] It is not clear to us how disputes over the accuracy of translations of *live* testimony are handled, except that the paucity of cases and commentary suggest that this is not a problem that is frequently litigated.

The distinctive feature of the use of this type of expertise is the emphasis on clarity. Because testimonial interpreters are imperative to the ongoing functioning of the system of oral testimony in the face of a language barrier, it is essential that they be comprehensible to the participants as they perform their function. As a result, the Court Interpreters Act includes a specific provision for the replacement of an interpreter who does not "communicate effectively"[49]—a problem that in other expert contexts we let the adversaries sort out.

Translation from less well-organized systems of specialized communication presents a different type of problem. When the question is "What signs and words do drug dealers use to convey their meaning without being understood by outsiders?" it is much harder to find well-informed disinterested experts than when the question is "What does this German sentence mean in English?" This can make it harder to find unbiased interpreters, and harder to check on the accuracy of a translation. The classic troublesome case is the police officer who qualifies as an expert in some variety of underworld jargon.[50] In that situation, the problem is frequently magnified by the fact that the officer-expert, who has an obvious interest in the outcome of the case, also testifies as a lay witness to critical acts by the criminal defendant or by others. When that happens, the process of qualifying the officer as an expert—both the imprimatur given by the court and the character evidence on which it is based—may improperly bolster the officer's *lay* testimony.[51] And, of course, this takes place in the presentation of the prosecution's case in a criminal trial, where our concern for accuracy ought to be at its highest.

c. Calculation

For this type of expertise, the question to the expert is some variant of "What does it add up to?" Totaling up the grocery bill, even if you have to add sales tax, hardly requires an expert, but some tasks that involve nothing more than a complex set of arithmetic calculations (pre-

paring a tax return) may benefit from an expert's help. If the job is big enough and complicated enough (calculating the indebtedness of a substantial company), or involves more sophisticated and difficult mathematical operations (calculating inferential statistics, or the force generated by a falling object) expert help is essential for most of us in any context, and always in court where the trier of fact is a lay jury or judge. In that situation, an expert who testifies to calculations will usually also explain what this type of calculation *means*, a separate task that we discuss under the heading "instruction."

Calculation, like translation, is a task of making sense of information that is already available. It can be done by any qualified expert once the data exist. As with observation and translation, the central issue is the competence of the calculator—a question that usually can be routinely answered by reference to her training and experience. In court, the qualifications of the expert are the only requirement for admissibility of this sort of expertise; mathematics itself is correctly considered to be universally reliable.

Strictly speaking, calculation itself is absolutely predictable; mathematical operations leave no room for ambiguity. Experts do disagree, however, on antecedent questions: *what* to count and *how* to do so. As the assumptions that are built into the process become more complex and debatable, the task may slip from "calculation" to "assessment." As we have drawn the line in this Article, determining the indebtedness of a company is near the borderline. We classify it, perhaps arbitrarily, as a "calculation," although, as we all learned from the Enron and Arthur Andersen scandals, there are more ways than one to make the many underlying judgments on how to categorize various possible liabilities and assets. On the other hand, we describe the task of estimating the present value of a person's lifetime earnings if he had survived an accident as an "assessment." The difference is not so much that this task is more complex or controversial than calculating indebtedness, but that it requires integrating information on, and making assumptions about, more disparate issues—productive life span, possible career paths, future employment trends and wage rates, productivity and inflation in the economy, and so forth. As with observation, the location of the boundary line between the more descriptive task of "calculation" and the more analytic task of "assessment" is not crucial, as long as we realize that some calculations are based in part on controversial assumptions; that many assessments have calculations embedded in them; and that there are close cases that are hard to call.

As calculation shades into assessment, questions of method and bias become increasingly salient. The question may be not so much what the indebtedness of Enron adds up to, but what the expert *means* by "indebtedness." If both sides have access to competent experts, legal disputes that involve calculations are likely to focus on these choices and on the expert assessments on which they are based, or on the accuracy of the underlying data, rather than on the mathematical calculations themselves. Since the results of the mathematical operations are determinate, there should be no more room to dispute the *calculation* that produces an estimate of the expected lifetime earning of a deceased plaintiff than to dispute the arithmetic (as opposed to the items) on a tab in a restaurant. If there is a dispute, replication (in a restaurant or in court) is a perfect check on mathematical accuracy. In theory, calculations and their underlying assumptions can always be teased apart, the assumptions spelled out and separated from the calculations that follow. In practice this may make the information too dense to absorb—we do not really want to know how and why our accountant calculated the tax as she did—and in any case it often is not done.

2. Instruction

Experts, to state the obvious, educate—they provide lay people with useful information. In court, the only function of an expert witness (or for that matter, any witness) is to educate, in this very general sense: to supply information that helps the trier of fact make decisions.[52] We focus here on a particular kind of education—"instruction"—by which we mean general information about some common issue or phenomenon ("Inadequate drainage can cause ground water to undermine a foundation") rather than specific information about a particular problem or case ("The wall collapsed because of inadequate drainage at its base"). Instruction answers such questions as "What do experts know about that topic?" and "How do things like that work?"

Instruction provides background knowledge but not the case-specific answers.[53] An expert witness who gives purely instructional testimony in a trial can literally repeat the same performance in a different courtroom with a different cast of characters, in another case that raises a parallel issue with different specific facts. Instructional testimony may lead to an inference that suggests or even requires a specific decision in a given case, but it is not itself information about any particular case. The statement "Bendectin is not a teratogen"—once controversial, now no longer so—is an example of instruction: it is a general empirical claim. By contrast, the statement, "Ingestion of Bendectin by the plaintiff mother did not cause this limb defect in her offspring," is no longer instruction. It follows logically from the general instructional statement, if true; but it also may be true even if that general statement is not. In any case, the second assertion is a claim about the specific events involved in the litigation—a type of statement that we classify and discuss as an *assessment* of causation.

As usual, effort can make the boundaries fuzzy. Consider, for example, the archetypal early twentieth-century hypothetical question in which the lawyer would ask an expert for his opinion about the cause of death of a "hypothetical person." In form, the lawyer would ask for an opinion on the cause of death of *anybody* in like circumstances, but the actual question (which sometimes went on for pages) would incorporate a whole series of *particular* facts that happened to be in evidence in the case at hand, and the expert would provide an opinion assuming these many ostensibly hypothetical facts to be true.[54] Strictly speaking, the testimony is framed in general non-case-specific terms and, therefore, fits our definition of "instruction." But the effect, of course, if we look past the formal structure of the testimony to its substance, is an extremely fact-specific opinion that is best described as an "assessment."

Whatever the dividing line between instruction and assessment, most experts—and certainly most expert witnesses—go back and forth across it and provide both types of information. This need not be the case. Eyewitness identification experts, for example, if permitted at all, are typically allowed to provide instruction only. They may testify, for example, that in general cross-racial identification is more likely to be erroneous than intra-racial identifications, or that eyewitness confidence is not generally a good predictor of eyewitness accuracy. But case-specific testimony evaluating the merits of a particular eyewitness's identification, if offered, would typically be excluded as an invasion of the province of the jury.[55] Most experts, however, do provide case-specific assessments, and usually incorporate instruction into that testimony, to make their conclusions both more comprehensible and more persuasive.

In this Article we discuss two different kinds of instructional statements: claims about matters of fact in the physical world, and statements about social norms—customs, common practices, ethical and professional standards, research methodologies. In both contexts the

expert gives her views on the state of the world—instruction on norms, in this sense, is factual rather than normative—but the purposes to which each kind of instruction are put in court are quite different. Much (though by no means all) factual instruction is uncontested, and typically, even when controversial, it is not a central issue in a case but rather serves as the background for an assessment. Occasionally, however, a disputed general factual proposition—in the classic case, "Bendectin is a human teratogen"—is an essential element of a claim. By contrast, in malpractice claims against professionals—lawyers, accountants, psychotherapists, and especially doctors—instructional opinions about norms—the professional standard of care for the conduct at issue—are often at the heart of the case. As we shall see, these differences in nature and purpose have implications for the admissibility of these two types of instructional testimony.

a. Facts

If you want to know an esoteric fact, you ask an expert, or—much the same thing—check a book or an article written by one. Most of the time we can rely on the information we get in this manner because the answer, although news to us, is well-known and uncontroversial among those who specialize in the area. This can be true for information at any level of generality, from the universal ("the speed of light in vacuum is 299,792,458 meters per second") to the highly particular ("73 percent of the respondents on this survey did not know the name of either of their United States senators").

In court, much expert information of this sort could be presented without testimony by any witness. Frequently, the statements at issue—that human blood has a pH of 7.4, for example, or that every person (except an identical twin or a clone) has a unique DNA sequence—are suitable subjects for judicial notice because they are not open to reasonable dispute and are "capable of accurate and ready determination by resort to sources whose accuracy cannot reasonably be questioned." [56] In practice, such information is generally introduced in the course of expert testimony, in part because the statements accompany other information provided by the expert that could not properly be judicially noticed, and in part because a well-prepared expert witness is more likely than the judge to present the information in a manner that is helpful to the side that called her.

A layperson cannot know in advance which propositions are not in dispute in medicine or physics or genetics. Out of court, we can guard against being misled (as usual) by asking more than one expert, or consulting more than one reference, or relying on a source—the United States Census Bureau, for example—whose authority we do not doubt. In court, the adversarial process defines what is in dispute, and undeniable expert facts are likely to form part of the uncontested background for the assessments that are contested. The parties may fight about whether a substance found in the defendant's possession was cocaine, but they are not likely to fight about the chemical make-up of cocaine itself. They might disagree about whether a DNA-typing test was done correctly in the particular case, or about how to interpret the evidentiary significance of finding the defendant's DNA at the crime scene, but they are unlikely to do battle over the general premise that apart from identical twins, a person's DNA genome is unique. On background issues of that sort, there is no debate about factual accuracy, and hence no question of reliability or admissibility. There is still room for bias, of course, in the selection and organization of the material. A testifying expert (like any teacher) chooses what to present from a large array of possible pieces of information; she may select the background facts most useful to her party's case and omit other information that is less helpful to her client. This form of bias is

almost invariably addressed by cross-examination and rebuttal testimony rather than exclusion. Assuming the opposing party is motivated and prepared, these may well be adequate mechanisms for checking partisan impulses to provide partial information.

As we move away from well known facts about which there is little or no disagreement, matters become more complicated. In theory, this is a problem that can be addressed: after more tests, more study, and more examination, we might someday reach answers that, if not wholly conclusive, may at least achieve widespread consensus. For a straightforward issue—the population of Detroit in 2000—a single, well-executed study (the decennial census), despite its imperfections, may provide as good an answer as we will ever have. But for difficult and complex questions (the origin of HIV; the structure of quarks) the process is likely to be complicated and uncertain, and even many comparatively easy issues are never addressed, or the studies are not complete when we need them. Whether because of the complexity of the issue or the paucity of the data (or both), at any point in time, a general factual question may have no clear answer. Knowledgeable people may disagree on what the answer is, or whether it is known, or even whether the question is answerable. And even if there is a reasonably clear answer out there somewhere, it may be no mean feat to find it. None of us can absorb even a tiny fraction of the general knowledge that exists in our extremely complex culture. We have no choice but to rely on experts, those who know the issue best—which merely pushes the problem back a step, to the identification of the most knowledgeable experts to rely on.

How do we determine matters of instructional fact outside of the courtroom? On routine issues, even debatable ones—and often on big ones as well—we do not invest a lot of energy in screening expert instruction. We look in well-known sources or ones we are familiar with (the Internet, the New York Times), we ask experts we have used before who are referred to us by friends and acquaintances, or we see what highly-credentialed experts have to say. If it *really* matters, however—if a lot of money, or the future of an institution, or someone's life or health is at stake—we increase our search costs. We spend time and energy to identify leading experts and to learn what they think about the matter.

What exactly are we searching for? Since we cannot identify accurate expert instruction directly, we look for the best available proxy: the consensus of well-informed experts in the field. Of course there may not be a consensus, and if there is, it may be out-of-date or just plain wrong. But figuring out the consensus viewpoint, if there is one, is the starting point of any well-conducted inquiry into general facts outside of litigation; more often than not, it is the ending point as well.

In the context of litigation, identifying a consensus among experts in court can be very hard. The expert's credentials are not much help. An imperfect proxy for knowledge in the best of circumstances, credentials lose nearly all of whatever value they may have when expert witnesses are chosen by the parties, from among those who will say what the parties want to hear, in part precisely because of their seemingly impressive credentials.[57] Worse, the parties who call expert witnesses have no common interest in identifying a consensus among experts. On the contrary, if there is any room at all for a dispute among experts, one side or both may have an interest in obscuring the extent of general agreement within the field and creating the appearance of an active dispute when there is none. The experts the parties call may take positions that are at best outliers within their own specialties but sound perfectly plausible to outsiders. If the factfinder is a jury, it may not even have the power to put questions to the party experts,

much less seek another opinion or investigate another source.[58] If the plaintiff calls an expert who testifies that silicone gel breast implants are known to cause auto-immune diseases, and the defendant calls an expert who testifies that there is no evidence that these devices produce that effect, the factfinder is in no position to judge which statement reflects a consensus in the field, if one exists, and which statement is considered unlikely or simply false. In any other contexts, the answer would be to check with experts who are not chosen by and identified with those parties whose interests are directly at stake—a panel assembled by the NIH rather than researchers hired by the pharmaceutical company whose product is being studied. That procedure is available in court under Federal Rule of Evidence 706, but, as we have mentioned, it is rarely used, and almost never in routine cases.

If credentials do not help much, and unbiased expertise is not sought, what legal tools are left to address conflicting expert claims about general factual propositions? We can exclude the evidence, or we can do our best to make use of what we get. In practice, we rely on the second sort of remedy more often than the first.

Daubert and its progeny—including the post-*Daubert* version of Rule 702—require courts to assess the validity of this sort of expert evidence. But how? By what standard? If we are not concerned about the validity of the field (if we are dealing with medicine, for example, and not astrology), a consensus of informed experts about the matter of general fact should certainly be sufficient to justify admission, but is probably too much to require. On one hand, if the cause of a phenomenon is unknown—and generally agreed to be unknown—this requirement would prevent the jury from getting any meaningful expert assistance on causation, since the only thing an expert witness could testify to is the consensus that experts do not know. On the other hand, for the large array of issues on which experts disagree on whether there is a consensus, or on what the consensus is, this rule would be impossible to administer. Could judges possibly determine the existence of an expert consensus, item for item, as a precondition to permitting any expert witness to make general factual claims? The difficulties of applying *Frye* pale by comparison.

A judge could demand a lesser degree of agreement as condition for expert instruction. A mild version of this inquiry into the consensus of the field seems to us to be a reasonable requirement for expert instruction. Because the judge's role is merely to decide whether the information may be considered at all, the test should be whether other experts in the field generally agree that the factual claim is plausible. This is a different kind of agreement than that demanded by *Frye* ("general acceptance" of the method or technique), and it does not require any level of agreement on the ultimate truth of any asserted fact. The focus, rather, is on the legitimacy of the factual claim: is it (at a minimum) considered a debatable point among those who know best?

Rule 702 speaks to the issue of instructional testimony only obliquely.[59] This category of testimony does not fit neatly into the Rule's structure, which assumes that expert testimony will apply "to the facts of the case."[60] Nonetheless, Rule 702 can be interpreted to impose a reliability requirement on general factual instruction, as for all expert testimony. To the extent that courts do scrutinize scientific testimony about general facts, we suspect that their inquiry focuses on the kind of consensus we have described: the degree of plausibility of the factual proposition within the witness's field. In general, however, the courts make no such inquiry. They simply require that an expert be "qualified," and then—except in a small group of unusual but important cases—allow them to make a wide range of general factual claims within their fields of expertise. Once a witness has been permitted to testify as an expert under Rule 702, judges usually leave

the task of correcting and explaining their instructional statements to the opposing parties and the expert witnesses they call.

Occasionally, a party may call an expert witness who says things that are simply bizarre. This seems to happen most often in criminal cases, where resources and standards of practice—especially on the defense side—are sometimes abysmally low. In the 1980s, for example, an anthropologist named Louise Robbins was allowed to testify for the prosecution in several cases that it is possible to identify a person from nothing more than a shoe print—the *wearer*, not the shoe.[61] The only explanation is incompetence by the judges, the defense attorneys, or both. (Her testimony was also excluded several times. Even so, the fact that she did testify repeatedly—indeed, that prosecutors even considered using her—is some measure of what's *not* considered a problem.) Usually, however, it is not the statement itself that is implausible but the confidence with which it is uttered. The problem is not the assertion that exposure to asbestos *might* cause colon cancer, but the claim that it *does*.

On reflection, it appears that the level of certainty that a witness attaches to a statement is likely to be the dominant problem for this type of expert testimony. If an expert testifies, for example, that a scientific study published in a peer-reviewed journal found that ephedra had a teratogenic effect on mice when ingested at a certain dose, this is as unproblematic as instructional testimony can be; the expert is reporting a fact that can easily be checked. The expert might then go on to testify, however, on the basis of that study (perhaps in conjunction with other information, perhaps not), that in her opinion ephedra *is* a human teratogen. Although phrased as an "opinion" this second claim is still "instruction" rather than "assessment," since the conclusion remains general rather than case-specific. But there is much more reason to be concerned about reliability at this higher level of generality. The study might be biased or poorly-conceived, and even if it is flawless, it might not be a sufficient basis for the expert's general conclusion. Even so, the issue is not so much the possible causal link the expert draws between ephedra and birth defects, but her confidence in the existence of that link. Few would quarrel with the witness if she had said that ephedra "might" be a human teratogen, or even perhaps that "there is a substantial risk" that it causes birth defects. The problem emerges as she moves up the ladder from *possible* to *probable*, to *very likely*, to *is*.

In the usual case, however, testimony is not excluded because a witness is too certain or too uncertain. Normally, a witness's level of confidence is grist for the adversarial mill: the jury considers it in evaluating the witness's credibility, taking into account any in-roads made in cross-examination, inconsistent evidence on rebuttal, and arguments by counsel. Lots of witnesses are excessively self-assured; many are simply wrong. We do not exclude an eyewitness who says she is 100 percent certain that she can identify the defendant as the man she saw from fifty feet, across a dark parking lot, for five seconds; we impeach her. By the same token, we do not generally exclude testimony from a doctor who says that he *knows* that regular exercise decreases the risk of coronary heart disease. Like the eyewitness, even if this expert is correct on the ultimate issue, she may not be entitled to speak with the level of confidence that she expresses; like the eyewitness, the jury will usually be allowed to hear her testimony, and the opposing party will have a chance to try to get her to weaken or qualify it, or to make it seem unreasonable.

There are, however, several important differences between the lay eyewitness and the expert providing fact instruction. The structural incentives for the expert to distort and overstate are significantly greater. The pressure of litigation, the partisan identification an expert may

develop over time, and the fact that the expert receives payment may spur even a well-intentioned expert to make instructional inferences in a stronger form than a fair-minded reading of the evidence supports. Moreover, it may be precisely the degree of overstatement that makes otherwise plausible testimony unreliable.

When does a general factual proposition become so unreasonable that it may be excluded entirely from evidence? In some tort cases a general factual proposition is a necessary step to establish causation, and substantive tort law requires expert testimony on causation that is phrased in terms of a specified level of confidence. In that situation the admissibility of expert instructional testimony on a question of general fact is outcome determinative, and detailed attention by the parties and the court is worth the candle. For example, in *Joiner*,[62] the plaintiff claimed that he was exposed to PCBs by the defendant, and that the exposure caused the small cell lung cancer from which he suffered. The district court excluded the plaintiff's proffered expert testimony on causation because "the court is not persuaded by a preponderance of proof that the studies support the 'knowledge' the experts purport to have (*i.e.*, that PCBs, 'to a 'reasonable degree of medical certainty," promote small cell lung cancer in humans)."[63] And, since the plaintiff was now without this necessary testimony—"to a reasonable degree of medical certainty"—the court granted the defendant's motion for summary judgment and dismissed the claim.

In *Joiner*, the Supreme Court affirmed the district court's ruling.[64] The main issue in *Joiner* was the standard of review on appeal of a ruling excluding evidence under Rule 702. The Supreme Court concluded that "abuse of discretion is the proper standard by which to review a district court's decision to admit or exclude scientific evidence," and that "the District Court did not abuse its discretion in excluding…[the plaintiff's expert] testimony."[65]

Under *Joiner*, other district courts could reach the opposite conclusion and admit identical expert testimony on the identical issue of general causation. This is a disturbing rule. The trial court's conclusion in *Joiner* that "there is simply too great an analytical gap between the data and the opinion" is really a substantive judgment that the plaintiff's evidence is insufficient masquerading as a procedural judgment that it is inadmissible. It may make perfect sense to give trial courts the authority to decide this issue—but why call it an evidentiary ruling, all but eliminate appellate oversight, and create a situation in which the very evidence that is legally sufficient to prove a fact before Judge Monday is excluded entirely by Judge Friday?

There are bounds on the discretion conferred by *Joiner*. Courts do not like to look (or to be) foolish. If a general proposition becomes widely accepted in a respected community of experts, courts usually fall in line, if sometimes after an uncomfortable gap. In the first part of the twentieth century, most physicians believed that a single traumatic blow could cause a malignant tumor, but by 1940 there was a medical consensus rejecting this theory. It had an afterlife of sorts, however, in court: it was presented with approval in at least a few cases for another fifteen or twenty years[66]—but apparently not beyond the early 1960s. Today, it would be rejected out of hand.

We have seen this process repeatedly in the past several decades in mass toxic tort litigation. In a typical mass tort situation, many separate plaintiffs make the identical argument on general causation, the issue is subject to systematic study and replication, and vast amounts of money are at stake. Given these circumstances, the general causal relationship between the substance and the pathology at issue is studied repeatedly, and eventually everybody (or almost everybody) agrees that there is or is not a causal connection. In *Daubert*, for example, the plaintiff's experts were prepared to testify in 1989 "that to within a reasonable degree of certainty

Bendectin is a teratogen,"[67]—testimony that in form is very similar to that offered in *Joiner*. But the context was different. By 1989, thirty separate epidemiological studies had failed to find a significant relationship between Bendectin and birth defects.[68] Given that body of knowledge, the proposition that Bendectin does not cause birth defects—like the claims that asbestosis causes lung cancer[69] and that silicone-gel breast implants do not cause systemic disease[70]—was no longer open to debate. One way or another, courts now treat these factual statements as rules of law.

Courts do not generally have problems with the expert instruction that sticks close to observable facts. They will not hesitate to admit testimony from a harbor pilot who testifies that he has seen several twelve-meter yachts pass under a particular bridge at high tide, or from a doctor who testifies that he has seen several patients with this sort of injury relearn to walk. Whether we trust doctors more than pilots (or less), the structure of the testimony is the same: "I've seen it happen; draw your own conclusions." For factual generalizations, however, courts are likely to discriminate between fields in a manner that reflects their assessment of the strength of the scientific basis for the discipline. The harbor pilot will be allowed to testify about the height of tides in the harbor he works in, but not about the causes of unusual tidal patterns. Epidemiological evidence that asbestos causes lung cancer is viewed as conclusive, but testimony from a questioned document examiner that forgeries can *always* be detected will meet with skepticism, and testimony from an astrologer that Virgos have controlling personalities would be excluded if it were ever offered.

These sorts of distinctions, of course, are exactly what Rule 702 requires: courts must determine whether expert instruction "is the product of reliable principles and methods."[71] For factual generalizations about the nature of the world, our touchstone is the scientific method of observation, experimentation, and replication. How proficient courts are at applying these principles is another question. Occasionally, they exclude plausible but unproven generalizations (for example, that exposure to PCBs promotes small cell lung cancer). More often, courts admit unreliable generalizations because they are accustomed to evidence of that sort—handwriting expertise, for example[72]—or because the witness has impressive professional credentials—a Ph.D. (like Louise Robbins[73]) or an M.D.

b. Norms

A different kind of instructional information concerns the norms of a field—typically, customs and standards of care. In a malpractice suit, for example, a witness may testify about the appropriate standard of care for treatment of asthma in infants, offering instructional expertise. These claims are also empirical—the witness is describing the custom of the field, her opinion about what is done in ordinary practice—but the empirical facts they concern are socially created. The expert's task is not to describe the *best* treatment for a condition, but rather the *standard* treatment. Ideally, the latter will conform to the former, but not immediately, not always, and not everywhere.

Outside of litigation, this category of expert information is comparatively unimportant, except to the extent that it is used in less formal attempts to assign responsibility or blame. In torts cases involving professionals, however—doctors, lawyers, accountants, psychotherapists—these instructional opinions are often the heart of the case, the key testimony on the key legal question. Because the content of professional norms is central to malpractice litigation, fac-

tual information about these norms is essential. Because this information concerns social norms rather than natural facts, it is treated differently from other types of expert instruction.

The only essential requirement for an expert on the norms of a community is, obviously, familiarity with that community and its behavior. This prerequisite mirrors the knowledge that is required of a translator, except that in this context, the expert must be familiar with the *customs* and *practices* of the culture (or subculture) rather than its *language*. Accordingly, in court, the only requirement for an expert witness who testifies about professional norms is that she qualify as an expert in that profession. Any competent endocrinologist can offer opinions about the ordinary and acceptable practice of endocrinologists, and any experienced criminal defense attorney can talk about the norms of her profession.

Note that the ultimate question in cases involving standards of care or custom is *Frye*-like: What is the *generally accepted* practice of a given community? Here however, the "general acceptability" of a practice is a question for the trier of fact, not a foundational requirement for admissibility. Since the content of the norm *is* the question for the jury—and perhaps also because the issue is so thoroughly one of *social* norms—we are willing to admit even idiosyncratic opinions about community norms, so long as these idiosyncratic opinions come from *bona fide* members of the relevant community. The key for determining admissibility is the speaker, rather than the substance of what is spoken; we evaluate qualifications, rather than the validity of the content of the testimony.

There are, of course, subsidiary issues. How close must the fit be between qualifications and the question at issue? Should a non-specialist physician be able to testify about the custom of a particular medical specialty? Should a family practitioner be allowed to testify about the standard of care in a case involving high-risk obstetrics? Should a properly trained physician who no longer sees patients be allowed to testify about treatment norms? Courts have generally treated these issues as going to weight rather than admissibility, but at some point, a mismatch between the expert's experience and the norms about which she is testifying may become so glaring that exclusion is appropriate.

As always, bias is an important issue for instructional expert testimony about standards and customs. Two kinds of bias are of particular concern. First, there is partisanship—the bias of the professional expert who may be all too willing, consciously or unconsciously, to tailor her testimony to suit her client's needs. Second, there may be a problem of professional solidarity—a tendency to avoid criticism of colleagues. Even if they have no personal relationship, one physician may be reluctant to testify against another, either out of a sense of professional camaraderie, or because the witness recognizes that the mistake the defendant made was one that the witness could well have made herself. In addition, there is the general problem of hindsight bias: it may be too easy after the event to criticize actions that were in fact reasonable in light of what was known at the time. All of these issues of bias typically are understood to affect the weight of the evidence but not its admissibility.

Daubert and its progeny have had little effect on this kind of instructional testimony. While physicians' evidence about causation in toxic torts cases has been curtailed in the last decade,[74] physicians' testimony about standards of care remains substantially unchanged. This lack of substantive scrutiny seems largely appropriate since there is no external standard by which to evaluate evidence on the norms and practices of a social group. For the same reason, however, courts should scrutinize the *fit* between the speaker's experience and the community

about which she speaks. Of course, requiring too close a fit may heighten the danger of bias. If only Charlottesville cardiovascular surgeons can testify about the proper standard of care for a triple-bypass operation in Charlottesville, we might be concerned that the only qualified witnesses would be biased in favor of the defendant because of local professional and community ties; a concern for fit should not be used as an argument for resurrecting the locality rule in medical malpractice cases. On the other hand, it seems reasonable to require that the witness practice cardiovascular surgery *somewhere*. So long as professional communities are defined on a national basis, it ought to be possible to require a close fit for an expert who testifies on professional norms without unduly limiting the pool of potential witnesses.

3. Assessment

When we hire an expert, an architect, or a surgeon, we usually want them to do things for us. One of the most important things we may want them to do is make decisions—where to build, whether to operate—or at least to make specific recommendations. Other than doing the tasks themselves, that is the central function of experts: to provide specific, concrete advice on what course of action to take. Their most conspicuous function in court is much the same, but in a context in which that sort of advice is otherwise prohibited.

Perhaps the most basic distinction in American trials is between the job of the witness and that of the trier of fact. These are the only two absolutely essential speaking roles in court. The parties typically only speak as witnesses, and we can and often do conduct trials entirely without lawyers. But there can be no trial without someone to provide evidence and someone to judge it. At ancient common law, these functions might merge: jurors were sometimes selected because they knew the circumstances of the case. We now require strict separation. Jurors and judges evaluate evidence, but do not produce it. They may only consider information about the case that they hear from witnesses in court; they may not be witnesses themselves;[75] and if they know too much about the facts from extra-record sources, they may not be allowed to serve at all. On the other side, witnesses are supposed to present information, not evaluate it. They are limited by the personal knowledge[76] and lay opinion[77] rules to testimony about matters they perceived ("I saw two men cross the street"), and to low level descriptive inferences based directly on their own perceptions ("The older-looking man was speaking very quickly").

Except for experts. Expert witnesses are not subject to the personal knowledge rule or the lay opinion rule. They are allowed to make wide-ranging evaluations based on many types of evidence, first- and second-hand, admissible and inadmissible; they may express opinions on the precise factual issues the judge or jury must decide; they may evaluate the entire body of evidence before the court and tell the trier of fact what decision to reach. These case-specific evaluations, applying background knowledge and expert skill to the facts at hand, are what we call "assessments." In many cases expert assessments take the form of judgments that can be adopted directly by the trier of fact: "The accident was caused by faulty wiring in the fuselage," or "the defendant is the biological father of the minor child." For some claims, expert assessments are required as elements of proof.

Because of their power and importance, expert assessments have received a great deal of attention. They are the visible high-end of expert evidence. For just that reason, we devote relatively little space to them. Our main agenda is to explore types of expert evidence that have received insufficient attention, and this one does not qualify. Quite the opposite: We could not

begin to do justice to the range of assessments that are offered by expert witnesses, or to the vast literature discussing them. Instead, we will offer a set of observations that connect expert assessments to the other types of expert evidence—observation and instruction—that we have already discussed. Indeed, we have already discussed expert assessments themselves, to some extent, to point out the distinctions between assessment and observation or instruction, and the overlaps between them.

We use a medical malpractice case, *Zuchowicz v. United States*,[78] to illustrate a common and important type of expert assessment and to raise two issues that we will talk about briefly: (1) the relationship between assessments and other types of expert evidence, and (2) the treatment of expert assessments that are not based on "science." While our discussion focuses primarily on assessments of causation, much of the analysis is applicable to the other kinds of assessment in our taxonomy as well.

a. Assessments and other types of expert evidence

The plaintiff in *Zuchowicz* developed the rare disease Primary Pulmonary Hypertension (PPH) shortly after a naval pharmacy instructed her to take over twice the maximum daily dose of the drug Danocrine.[79] The United States as defendant conceded that its doctors and/or pharmacists had been negligent, and that the plaintiff died from this disease. The only triable issues were causation—did the overdose of Danocrine cause Mrs. Zuchowicz's PPH?—and damages. The district court, sitting without a jury, granted judgment to the plaintiff, and the Second Circuit affirmed.[80]

Most of the factual discussion in the Second Circuit opinion is based on expert evidence, although the court does not always seem to be aware of this. We are told that starting in late February 1989,

Mrs. Zuchowicz took the 1600 milligrams of Danocrine each day for the next month. Thereafter, from March 24 until May 30, she took 800 milligrams per day. While taking Danocrine she experienced abnormal weight gain, bloating, edema, hot flashes, night sweats, a racing heart, chest pains, dizziness, headaches, acne, and fatigue.[81]

The symptoms discussed in this section could have been described by lay observers. Undoubtedly, most of the information (hot flashes, night sweats, headaches, dizziness, etc.) originally came from statements by the deceased plaintiff herself. In practice, however, it is likely that all of these symptoms were described in court by experts, the testifying physicians, and mostly on the basis of reports from other physicians or nurses. Certainly few patients would describe a symptom they experience as "edema." We are then told that:

In October 1989, [Mrs. Zuchowicz] was diagnosed with primary pulmonary hypertension ("PPH"), a rare and fatal disease in which increased pressure in an individual's pulmonary artery causes severe strain on the right side of the heart. At the time she was diagnosed with the disease, the median life expectancy for PPH sufferers was 2.5 years. Treatments included calcium channel blockers and heart and lung transplantation.[82]

Most of this section is expert instruction: the nature of PPH, the fact that it is rare, a patient's life expectancy after diagnosis, and standard treatments. The initial statement, however, is an expert *assessment*: that the plaintiff was diagnosed as suffering from PPH. In the context of this case, it is not controversial, but it is an expert assessment of her condition all the same. Inevitably, it will have been based in part on expert observations—blood pressure measurements,

perhaps other tests—by the testifying witnesses or (more likely) by other experts, that may or may not have been described in court.

The opinion next proceeds to a detailed discussion of the PPH:

Pulmonary hypertension is categorized as "primary" when it occurs in the absence of other heart or lung diseases. "Secondary" pulmonary hypertension is diagnosed when the hypertension results from another heart or lung disease, such as emphysema or blood clots. PPH is very rare. A National Institute of Health registry recorded only 197 cases of PPH from the mid-1980s until 1992. It occurs predominantly in young women. Exogenous agents known to be capable of causing PPH include birth control pills, some appetite suppressants, chemotherapy drugs, rapeseed oil, and L-Tryptophan.

According to the district court's findings of fact, the disease involves the interplay of the inner layers of the pulmonary blood vessels known as the endothelium and the vascular smooth muscle. The endothelium releases substances called vasodilators and vasoconstrictors, which dilate and constrict the blood vessels.[83]

The court goes on to present the current theory of the cause of PPH: "If too many vasoconstrictors are released, the blood vessels contract, the endothelial cells die, and the vascular smooth muscle cells proliferate. These actions create increased pulmonary vascular resistance."[84] All of this, of course, is based on expert instruction, which continued in the court's discussion of the drug at issue, Danocrine:

According to the testimony of plaintiff's expert Dr. W. Paul D'Mowski, who personally performed much of the initial research on the drug, Danocrine is safe and effective when administered properly. Based on studies by Dr. D'Mowski and others, Danocrine was approved by the Food and Drug Administration ("FDA") for use in dosages not to exceed 800 mg/day. Mrs. Zuchowicz was accidentally given a prescription instructing her to take twice this amount—1600 mg/day. According to Dr. D'Mowski no formal studies of the effects of Danocrine at such high doses have been performed, and very, very few women have received doses this high in any setting.[85]

This too is expert instruction, and like everything above it, uncontroversial, at least in this case. With all that now on the table, the dispute is narrowed to the two last points the plaintiff must prove: "(a) that defendant's act in giving Mrs. Zuchowicz Danocrine was the source of her illness and death, and (b) that it was not just the Danocrine, but its negligent overdose that led to Mrs. Zuchowicz's demise."[86]

After the discussions of PPH and Danocrine, the opinion turns to a new section entitled "The Expert Testimony" and describes in detail the testimony of two expert witnesses for the plaintiff, a professor of medicine and a professor of pharmacology.[87] The heading is telling, and misleading. Most of the opinion up to this point is based on expert testimony; one of the experts (Dr. D'Mowski) is even named. What distinguishes the issues discussed under the heading "The Expert Testimony" from those discussed above it is that they are in dispute; they define the question the court must answer: did the plaintiff adequately prove causation? This is a difficult, perhaps unanswerable question; it is the focus of the case. But the context of the question is at least equally important; the issues have been narrowed to that one question, and the court has learnt a great deal about what is generally known (or believed) about Danocrine and PPH. All of this is the product of expert testimony.

And what was included in the evidence labeled "Expert Testimony"? Two critical items:

> Dr. Matthay testified that he was confident to a reasonable medical certainty that the Danocrine caused Mrs. Zuchowicz's PPH. When pressed, he added that he believed the *overdose* of Danocrine to have been responsible for the disease.[88]
>
> …
>
> Dr. Tackett testified that, to a reasonable degree of scientific certainty, he believed that the overdose of Danocrine, more likely than not, caused PPH in the plaintiff… .[89]

The court then proceeds to describe the mechanisms by which, according to Dr. Tackett, Danocrine probably caused the plaintiff's illness and death.

These are expert assessments of the most ambitious sort. Drs. Matthay and Tackett each testified, in effect, "If I were judge, I'd find for the plaintiff on the issue of causation." Like any other layperson who consults an expert, the trial judge was free to disregard this advice. In fact—although it is not relevant to the Second Circuit opinion—the defendant almost certainly presented other experts who said the opposite.[90] A plaintiff, however, is required to present expert evidence of this sort to meet her burden of proof in a medical malpractice case,[91] and if it is accepted by the trier of fact, she wins. In this case, as always, the expert assessments were built on a foundation of expert description and instruction. Much of that was uncontroversial: Mrs. Zuchowicz's medical history, the nature of PPH, etc. But each of the doctors also relied, implicitly if not explicitly, on two general factual claims that *were* controversial: that Danocrine can cause PPH, and that an overdose of Danocrine increases the chances of developing PPH.[92] Those propositions were as necessary a step in their reasoning as the parallel claim in *Daubert* that Bendectin can cause birth defects. To the extent that they were discussed explicitly, that too was expert instruction: testimony designed to persuade the trier of fact on an issue of general application.

b. Nonscientific assessments

In *Daubert*, the Court noted that while Rule 702 applies to all expert testimony, "[o]ur discussion is limited to the scientific context because that is the nature of the expertise offered here."[93] That left open the question of whether *Daubert* applies to nonscientific expertise. In *Kumho Tire,* the Court said Yes: "We conclude that *Daubert*'s general holding—setting forth the trial judge's general "gatekeeping" obligation—applies not only to testimony based on 'scientific' knowledge, but also to testimony based on 'technical' and 'other specialized' knowledge."[94] This holding left open the question of how this general gatekeeping obligation should be carried out for nonscientific expertise, an issue that the Court begins to address in *Kumho Tire* itself, and that has attracted a great deal of attention since.

We start with a different question: Why is the expert testimony that was excluded in *Daubert* described as "scientific" (or at least, would-be scientific) evidence, and the expert testimony that was excluded in *Kumho Tire* as "nonscientific" evidence? The evidence at issue in *Daubert*, as we know, was testimony by several qualified medical experts "that Bendectin can cause birth defects."[95] The evidence in *Kumho Tire* was testimony by a qualified engineer with experience as a tire failure consultant, that "a defect in its manufacture or design caused the blow-out" that led to the accident that was the basis for the lawsuit.[96] What makes the evidence in *Daubert* "science" and that in *Kumho Tire* "nonscience"? The distinction does not reflect a difference in the scientific bases of the two disciplines at issue. As the Court notes in *Kumho Tire*,

"[e]ngineering testimony rests upon scientific foundations…"; arguably, engineering is at least as scientific as medicine.[97]

The difference, rather, is the nature of the issue the experts addressed. The question in *Daubert* was general causation: Can this drug ever have the claimed effect? Because it is a general issue it could be, and, as it happens, had been, systematically studied. The question in *Kumho Tire* was the specific historical cause of a single accident: Why did this tire blow out when it did? Science may provide us with tools (microscopes, reagents, etc.) to help answer that question by examining the remains, but the actual event cannot be replicated and studied systematically. The main factual issue in *Daubert* was what to make of the thirty published epidemiological studies that had failed to find a relationship between Bendectin and birth defects. That issue—in our terminology, a question of fact instruction—is properly in the sphere of science. The factual question in *Kumho Tire* was "whether the expert could reliably determine the cause of this tire's separation" on the basis of visual inspection of the tire in question.[98] That question—an issue of expert assessment—was a matter for practical judgment.

There's an irony here. Although the "nonscientific" expertise in *Kumho Tire* came from an engineer, there is no doubt that the bulk of similar testimony comes from medical experts, the same category of witnesses who provided some of the prohibited "scientific" evidence in *Daubert*. A majority of all witnesses in American trials are medical professionals—M.D.'s alone make up half of the total[99]—who testify primarily in personal injury trials of one sort or another. The most common role for a medical witness is to offer an expert assessment of the medical status of an individual—her condition (diagnosis), or its causes, or her future prospects (prognosis). These assessments are often nonscientific, practical clinical judgments in the same way that the engineer's testimony in *Kumho Tire* was nonscientific. The disputed expert testimony in *Zuchowicz* is a good illustration.

As the Second Circuit points out, the nature of the issue of causation in *Zuchowicz* essentially precluded scientific evidence on causation:

> The rarity of PPH, combined with the fact that so few human beings have ever received such a high dose of Danocrine, obviously impacted on the manner in which the plaintiff could prove causation. The number of persons who received this type of overdose was simply too small for the plaintiff to be able to provide epidemiological, or even anecdotal, evidence linking PPH to Danocrine overdoses.[100]

Instead, one of the plaintiff's experts, Dr. Tackett, offered a causal hypothesis: "Danocrine, more likely than not, caused PPH in the plaintiff by producing: 1) a decrease in estrogen; 2) hyperinsulinemia, in which abnormally high levels of insulin circulate in the body; and 3) increases in free testosterone and progesterone."[101] The conclusion of Dr. Matthay, the plaintiff's second expert on causation, "was based on the temporal relationship between the overdose and the start of the disease and the differential etiology method of excluding other possible causes."[102]

There may be nothing wrong with Dr. Tackett's hypothesis, but that's all it is—an untested hypothesis. The hypothesis itself was quite possibly reasonable,[103] and if Dr. Tackett had said that this was a *possible* cause of Mrs. Zuchowicz's illness, that would have been a plausible scientific statement. Instead, he testified "to a reasonable degree of scientific certainty" that it was "more likely than not" that this theoretical process had the effect he hypothesized, and, even more specifically, that it was more likely than not that the *overdose* caused the disease.[104] This

language may have been necessary to meet the substantive legal requirements of the plaintiff's malpractice claim, but asserting, with that level of confidence, that an untested hypothesis is true in general, and explains a particular event, is not science. Dr. Tackett inferred too much on the basis of too little.

Dr. Matthay's testimony sticks closer to an accepted expert methodology: differential diagnosis. But it too is not science. Treating physicians must make judgments about the nature and causes of their patients' pathologies even when they only have limited information, just as judges and juries must make decisions about disputed facts. In the process, doctors (like courts) gather the best information they can—scientific and nonscientific—consider the range of plausible theories, and try to exclude as many as possible. This is probably the best they can do. It may have been reasonable, for example, to conclude that an overdose of Danocrine might have caused Mrs. Zuchowicz's PPH. That may have been the best educated guess a treating physician could produce; it may have justified a therapeutic decision. But in the absence of any other evidence that Danocrine has ever had this effect, the temporal coincidence between the overdose and the disease in this one case cannot justify a high degree of certainty that the former caused the latter. Dr. Matthay's testimony—like much medical testimony based on sound clinical medicine—is far closer to the "nonscientific" expertise in *Kumho Tire* than to the "scientific" expert evidence in *Daubert*. But the problem with this testimony, insofar as there is one, is not that it is not science; it is, once again, the excessive confidence with which the expert's assessment is pronounced.

The expert testimony in *Zuchowicz* was controversial because of the unanswered underlying question of general causation: Can Danocrine, or an overdose of Danocrine, ever cause PPH? This question—a matter that is subject of general expert instruction—is of the sort than could in theory be studied and answered scientifically, although in this instance that might be exceedingly difficult. The district court admitted the testimony over objection, and the circuit court affirmed, essentially on the theory that the plaintiff was presenting the best expert testimony available. Given the discretion conferred by the Supreme Court in *Joiner*, a different trial court could have done the opposite and also have been affirmed.[105] In *Joiner* itself, for example, the trial court excluded specific expert assessments that exposure to PCBs promoted the plaintiff's lung cancer for lack of evidence on general causation: "Plaintiffs have failed to show by a preponderance of proof that their experts' opinions regarding the PCB/lung cancer link are admissible under the standards set out in Rule 702 and explicated in *Daubert*."[106]

In many cases, there is no dispute about general causation. Everybody agrees that exposure to asbestos *can* cause lung cancer. But *did* asbestos cause the lung cancer that killed a particular plaintiff—a fifty-eight year-old worker, for example, who was exposed to a moderate amount of asbestos and who smoked a pack of cigarettes a day since the age of 16? In that situation the only disputed issue is one of clinical judgment. In the usual case, each side will present one or more qualified doctors who will testify that asbestos did or did not cause the disease, and the opposing party will not even bother to object.

We do not mean to say that assessments of specific causation are inherently "nonscientific." The terminology is the Supreme Court's, not ours. In some cases the cause of a person's disease can be determined with a high degree of confidence. Extensive studies have shown that the vast majority, if not all cases of mesothelioma—a rare cancer of the lining of the chest, abdominal cavity, or heart—are caused by exposure to asbestos. Given that body of "scientific" knowledge, an expert could assert with a high degree of confidence that the mesothelioma suf-

fered by a particular patient was caused by exposure to asbestos. And advances in science may improve the ability of experts to determine the cause of a patient's illness in harder cases. New tests could perhaps be developed that will detect subtle morphological differences between lung cancer caused by asbestos and lung cancer caused by smoking. Our point rather is that doctors and other experts must regularly make judgments of this sort with limited information. When experts render opinions in the face of great uncertainty, we should recognize both their expertise and its limits.

In our culture, science is the dominant method for answering general factual questions: Does the consumption of trans-fatty acids increase the risk of heart disease? Can earthquakes be predicted by observing the behavior of animals? What is the relationship between confidence and accuracy for eyewitness identifications of strangers? Does Bendectin cause birth defects? When experts claim to be able to address these questions—to provide general factual instruction—courts should require an appropriate relation between the evidence supporting their claim and the degree of confidence asserted. In many cases the best scientific answer is uncertainty: It's likely, it's possible, it's plausible, it's unlikely—but in any event, we don't know.

The capacity to recognize and acknowledge uncertainty and to postpone judgment is an essential element of scientific inquiry. This is a luxury that does not extend to trials, however, or to any other situation in which decisions about particular cases must be made. Doctor, engineers, investigators, judges and juries all must decide how to act on particular facts in the face uncertainty. There is nothing illegitimate about a doctor's decision to recommend radiation treatment instead of chemotherapy to a cancer patient, even though an equally competent colleague down the hall would have done the opposite. Scientific studies may frame the issue and narrow the options, but if they do not provide a clear answer to a question that cannot be put off the doctor will have to do her best by other means: rely on anecdotal data, reason by analogy, or play a hunch. By the same token, it makes perfect sense for courts to listen to expert evidence about specific factual assessments—the diagnosis of a patient's disease or the cause of an accident—even though assessments are not in themselves scientific statements. If similar assessments are used to decide how to treat patients or design tires, courts should be willing to listen.

From a scientific point of view, the main danger of specific assessments based on limited information is overconfidence. To some extent, this danger may be inherent in the process of making decisions, especially difficult and important ones: those of us who are not plagued by self-doubt tend to become committed to our choices, however arbitrary. A doctor who prefers radiation to chemotherapy for idiosyncratic reasons may come to believe, and tell others, that it is far superior. The use of partisan experts magnifies this effect. Lawyers are likely to encourage their expert witnesses to talk in strong, unambiguous terms and to choose those experts who are likely to do so anyway or willing to take direction. The oncologist who says the plaintiff's lung cancer was "clearly caused by asbestos" may be hired by the plaintiff; the one who says it was "clearly caused by smoking" may be hired by the defendant; but the doctor who says "it looks like ____, but it's hard to tell" will never appear in court, regardless of how he fills in the blank. In addition, as we have seen, the substantive law governing the claim may provide an additional push for the expert to testify that a theory that is reasonable and plausible, but utterly uncertain, is "more likely than not" "to a reasonable degree of scientific certainty."

IV. Conclusion

American evidence law consists primarily of objections to the admission of evidence, so most judicial and academic discussions of expert evidence focus on admissibility. In the wake of *Daubert*, the central issues have been the validity of expert evidence, and to a lesser degree the competence of expert witnesses—both of which can pose problems for admissibility—rather than the bias of expert witnesses or the clarity of their testimony, which are almost always said to go to the weight of the evidence. This is a narrow point of view. For most uses of expert evidence, across topics and fields, admissibility is not a problem. The great majority of experts who are called are deemed competent to testify—"qualified"—without opposition. Many types of expert evidence are routinely admitted without explicit consideration of their validity; and in the uncommon cases in which the validity of expert testimony is genuinely disputed, exclusion is often not the best remedy.

We have tried to explore the entire assorted range of expert evidence that is offered in court, and the full range of issues it presents, by developing a taxonomy of expert information. We hope this scheme has some value; we know that it is a preliminary and incomplete effort. For the most part, we give a descriptive account of those uses of expert evidence that we discuss. We do, however, have a few general recommendations. We do not mean to imply that these ideas are original, or that courts do not (sometimes) act on them. We intend rather to salute those who do and encourage those who as yet do not.

First, attend in detail to the content of expert evidence. Expert witnesses may make many different types of statements. Their testimony will frequently blend observation, description and assessment. Even when part of their testimony is insufficiently reliable, other portions may be admissible. Most qualified experts who are called at trial probably have something to say that the jury should be allowed to hear, even if it is not everything the proponent wants to offer. In deciding on the admissibility of expert evidence, courts should focus on the actual statements the experts intend to make, and exclude only those that are for one reason or another inadmissible, rather than thinking of the expert's testimony as a unit that stands or falls in its entirety. And in *evaluating* expert evidence that is before them, courts should likewise focus on the separate statements rather than the testimony as a whole, and make use of what is reliable and valuable.

Second, pay special attention to the level of confidence the expert witness expresses. In *Daubert* itself the problem was categorical. By 1989, the claim that Bendectin causes human birth defects was simply *bad science*: there had been "more than 30 published studies involving over 130,000 patients," none of which "demonstrated a statistically significant association between Bendectin and birth defects."[107] To assert the opposite, given the strength of the available evidence, was simply not scientifically plausible. More often, however, the real problem with questionable expert testimony is less extreme, that the expert witness has expressed an unjustified degree of confidence. In *Zuchowicz*, for example, it would have been perfectly reasonable to conclude that the plaintiff's primary pulmonary hypertension *might* have been caused by an overdose of Danocrine. That would not have been a scientific *conclusion*—there was too much uncertainty to justify a confident conclusion about causation—but it would have been a plausible *hypothesis*, a clinical judgment on which a doctor or perhaps even a court might reasonably act, especially given the lack of better information. Unfortunately, the informational value of that sort of hypothesis in court is often degraded when it is described as a fact, or even when it is said to be "probable" or "likely" for no better reason than that the party calling the expert would like that

statement to be heard. What could be modest but useful expert evaluation is transformed into misleading pseudo-science.

All of us—doctors, lawyers, teachers, judges—are prone to exaggerate. It would be reassuring if this tendency were mitigated in the solemn precincts of our courts, if witnesses and advocates spoke with unusually measured care. For experts at least, the truth is probably exactly the opposite. Confidence and certainty are traits that lawyers seek when they choose experts—and traits they try to instill as they prepare them for trial—because they are understood to be effective. Worse, in many cases substantive rules of law require expert witnesses to phrase their testimony in terms of a specified level of confidence. It is no accident that Dr. Tackett testified "to a reasonable degree of scientific certainty" that it was "more likely than not" that Mrs. Zuchowicz's death was caused by an overdose of Danocrine; he was required to do so. It may be perfectly reasonable for courts to decide as a matter of law some questions that turn on expert evidence, and to take those issues—or the entire cases—away from the jury. But there is nothing to be gained from rules that distort the evidence that goes into the process, or that require talismanic language that creates special incentives for experts to exaggerate. The experts we call as witnesses will provide better information if we do not encourage them to speak with excessive confidence in order to be heard at all. After they are heard, it may make sense to reject the claims they address by finding insufficient evidence in support as a matter of law—but that judgment should be made on the basis of the best expert evidence we can get, not the most pliable.

Finally, look after the adversarial system. We insist on using partisan, adversarial expertise in court almost exclusively. There are other ways to make use of expert information in litigation—perhaps better ones—but this is ours. As long as we do depend on the adversarial system, we need to have one that actually is what it claims to be: a contest with two sides, where each serves as a check on the other. In many criminal cases, there is only one side on expert issues: the prosecution. The result is a national scandal. We have seen case after case of systematic fraud and incompetence by prosecution experts and police crime laboratories, with no end in sight. *Daubert* and the cases following it were civil cases; the problems they address limit the efficiency of our system of civil litigation, and may result in erroneous judgments at trials where money damages are at stake. The abuses that have been discovered in the use of experts in criminal prosecutions call into question the integrity of our system of determining the guilt or innocence for the most serious crimes, and have produced false convictions that have destroyed the lives of many innocent defendants.

Footnotes

[1] Fed. R. Evid. 702 (emphasis added).

[2] *Id.* (emphasis added).

[3] 509 U.S. 579 (1993).

[4] 522 U.S. 136 (1997).

[5] 526 U.S. 137 (1999).

[6] A number of commentators have addressed this issue. *See, e.g.*, Erica Beecher-Monas, *A Ray of Light for Judges Blinded by Science: Triers of Science and Intellectual Due Process*, 33 Ga. L. Rev. 1047 (1999); Scott Brewer, *Scientific Expert Testimony and Intellectual Due Process*, 107 Yale L.J. 1535 (1998); Judge Harvey Brown, *Eight Gates for Expert Witnesses*, 36 Hous. L. Rev. 743 (1999); Daniel J. Capra, *The* Daubert *Puzzle*, 32 Ga. L. Rev. 699 (1998); David L. Faigman *et al.*, *Check Your Crystal Ball at The Courthouse Door, Please: Exploring the Past, Understanding the Present, and Worrying About the Future of Scientific Evidence*, 15 Cardozo L. Rev. 1799 (1994) [hereinafter Faigman *et al.*, *Check Your Crystal Ball*]; David L. Faigman *et al.*, *How Good is Good Enough?: Expert Evidence under* Daubert *and* Kumho, 50 Case W. Res. L. Rev. 645 (2000) [hereinafter Faigman *et al.*, *How Good is Good Enough*]; Edward J. Imwinkelried, *Should the Courts Incorporate a Best Evidence Ruling into the Standard Determining the Admissibility of Scientific Testimony?: Enough Is Enough Even When It Is Not the Best*, 50 Case W. Res. L. Rev. 19 (1999); *Developments in the Law—Confronting the New Challenges of Scientific Evidence*, 108 Harv. L. Rev. 1481 (1994).

[7] *See Kumho Tire*, 526 U.S. at 158 (upholding trial court's exclusion of plaintiff's tire safety expert as unreliable); *Joiner*, 522 U.S. at 143 (finding trial court's exclusion of plaintiff's expert evidence on causation within judge's discretion); *Daubert*, 509 U.S. at 583–84 (vacating the lower court's exclusion of plaintiff's expert evidence and establishing that "general acceptance" is no longer the proper standard for admissibility); Daubert v. Merrell Dow Pharm., Inc., 43 F.3d 1311, 1322 (9th Cir. 1995) (excluding plaintiff's expert testimony on causation on remand from Supreme Court).

[8] *Daubert* offers a flexible set of guidelines for judges to examine when assessing the reliability of scientific expert evidence—testability, error rate, the existence of standards, peer review—and general acceptance. *Daubert*, 509 U.S. at 593–95. *Kumho Tire* holds that a judge *may* consider these criteria when evaluating nonscientific expert evidence, but emphasizes that this "list of specific factors neither necessarily nor exclusively applies to all experts or in every case." *Kumho*, 526 U.S. at 141. While granting the trial judge "considerable leeway in deciding in a particular case how to go about determining whether particular expert testimony is reliable," the opinion offers no concrete guidance on how to do this apart from using the *Daubert* factors beyond noting that in some cases, "the relevant reliability concerns may focus upon personal knowledge or experience." *Id.* at 150, 152. A number of commentators have begun to address the issue of how to evaluate non-scientific expert evidence. *See, e.g.*, Edward J. Imwinkelried, *Evaluating the Reliability of Nonscientific Expert Testimony: A Partial Answer to the Questions Left Unresolved by* Kumho Tire Co. v. Carmichael, 52 Me. L. Rev. 19 (2000); Edward J. Imwinkelried, *The Next Step After* Daubert*: Developing a Similarly Epistemological Approach to Ensuring the Reliability of Non-Scientific Testimony*, 15 Cardozo L. Rev. 2271, 2292–94 (1994); D. Michael Risinger, *Defining the "Task at Hand": Non-Science Forensic Science After* Kumho Tire Co. v. Carmichael, 57 Wash. & Lee L. Rev. 767 (2000); D. Michael Risinger, *Preliminary Thoughts on a Functional Taxonomy of Expertise for the Post-*Kumho *World*, 31 Seton Hall L. Rev. 508 (2000) [hereinafter Risinger, *Preliminary Thoughts*].

[9] Tacit knowledge and experience matter within science as well. *See, e.g.*, H.M. Collins, *The TEA Set: Tacit Knowledge and Scientific Networks*, *in* The Science Studies Reader 95 (Mario Biagioli ed., 1999). In addition, some have argued that *Daubert's* conception of science is unrealistic and idealized. *See, e.g.*, David S. Caudill, *Ethnography and the Idealized Accounts of Science in Law*, 39 San Diego L. Rev. 269

(2002). Even granting both of these points, developing criteria for evaluating reliability remains at least as difficult for nonscientific expert evidence as it is for scientific expert evidence, and perhaps more so.

[10] In-court experiments provide the limiting case. While it may look like a form of "doing," such an experiment is really a form of "telling" that is gussied up to look like "doing." No experienced attorney would want an expert to perform an "experiment" in court without being quite certain in advance about the outcome. Strictly speaking, this sort of performance is not an experiment but a "demonstration." Its purpose is not to teach the expert anything at all, but to display to the factfinder what the expert already knows.

[11] *See* Fed. R. Evid. 703. Rule 703 permits experts to base their opinions on facts and data not admissible in evidence, so long as they are "of a type reasonably relied upon by experts in the particular field in forming opinions or inferences upon the subject… ." An amendment to the Rule in 2000 restricts the extent to which the expert may testify to the bases for her conclusions when those bases are inadmissible. Such disclosure is permitted only when its probative value "substantially outweighs" its prejudicial effect. *Id.* However, given the existence of hearsay exceptions for statements made for the purpose of medical diagnosis or treatment in Rule 803(4), and for statements in learned treatises in Rule 803(18), it is clear that quite apart from the limited admissibility now available under Rule 703, experts may still inform the jury about a great many hearsay sources for their judgments.

[12] *See, e.g.*, United States v. Vallejo, No. 99–50762, 2001 U.S. App. LEXIS 7367, at *29 (9th Cir. Jan. 16, 2001) (reversing the district court on its admission of expert testimony on drug trafficking and its exclusion of expert testimony by defendant's school psychologist); Jenson v. Eveleth Taconite Co., 130 F.3d 1287, 1295 (8th Cir. 1997) (reversing a successful trial level *Daubert* challenge to plaintiff's psychiatric testimony); United States v. Powers, 59 F.3d 1460 (4th Cir. 1995) (affirming district court's exclusion of expert testimony that defendant did not exhibit characteristics of fixated pedophilia); United States v. Falcon, 245 F. Supp. 1239 (D. Fla. 2003) (excluding defendant's proffer of psychological expert testimony on the personality disorders of a key government witness). The first case to scrutinize handwriting identification evidence with care was *United States v. Starzecpyzal*, 880 F. Supp. 1027 (S.D.N.Y. 1995), in which the court thought the expertise failed *Daubert* but admitted it anyway as nonscientific expert evidence (a move now prohibited by *Kumho Tire*). Some recent cases have rejected handwriting identification expertise. *See, e.g.*, United States v. Brewer, No. 01 CR 892, 2002 WL 596365 (N.D. Ill. Apr. 16, 2002); United States v. Lewis, 220 F. Supp. 2d 548 (S.D.W. Va. 2002); United States v. Saelee, 162 F. Supp. 2d 1097 (D. Alaska 2001). Academic debate about handwriting identification evidence has been similarly vigorous. *See, e.g.*, D. Michael Risinger *et al.*, *Brave New 'Post-*Daubert *World'—A Reply to Professor Moenssens*, 29 Seton Hall L. Rev. 405 (1998); D. Michael Risinger *et al.*, *Exorcism of Ignorance as a Proxy for Rational Knowledge: The Lessons of Handwriting Identification 'Expertise'*, 137 U. Pa. L. Rev. 731 (1989); D. Michael Risinger & Michael J. Saks, *Science and Nonscience in the Courts:* Daubert *Meets Handwriting Identification Expertise*, 82 Iowa L. Rev. 21 (1996); *see also* Andre A. Moenssens, *Handwriting Identification Evidence in a Post-*Daubert *World*, 66 UMKC L. Rev. 251 (1997). *See generally* Jennifer L. Mnookin, *Scripting Expertise: The History of Handwriting Identification Evidence and the Judicial Construction of Reliability*, 87 Va. L. Rev. 1723 (2001) (providing a history of handwriting identification).

[13] Fed. R. Evid. 702 (emphasis added).

[14] Daubert v. Merrell Dow Pharm., Inc., 727 F. Supp. 570, 576 (S.D. Cal. 1989), *aff'd*, 951 F.2d 1128 (9th Cir. 1991).

[15] *Kumho Tire*, 526 U.S. 137 (1999); *Joiner*, 522 U.S. 136 (1997); *Daubert*, 509 U.S. 579 (1993).

[16] *See* Molly Treadway Johnson *et al.*, Expert Testimony in Federal Civil Trials: A Preliminary Analysis (2000) (finding in a 1998 survey of federal judges that fewer judges permitted all proffered expert evidence in their last civil trial than had done so in 1991); Lloyd Dixon & Brian Gill, *Changes in the Standards for Admitting Expert Evidence in Federal Civil Cases since the* Daubert *Decision*, 8 Psychol., Pub. Pol'y & L. 251, 269 (2000) (finding on the basis of an empirical study of district court decisions that "since *Daubert*, judges have examined the reliability of expert evidence more closely

and have found more evidence unreliable as a result"); *see also* 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 (1999); Steve Leben, *In Practice,* Daubert *Raised the Bar*, 37 Court Rev. 37 (Fall 2000); Joseph Sanders & Julie Machal-Fulks, *The Admissibility of Differential Diagnosis Testimony to Prove Causation in Toxic Tort Cases: The Interplay of Adjective and Substantive Law*, 64 Law & Contemp. Probs. 107 (2001).

[17] The empirical data on the effects of *Daubert* are limited, but some findings about how judges perceive *Daubert* are suggestive. *See* Johnson *et al.*, *supra* note 16 (stating that though the number decreased between the years 1991 to 1998, in 1998, 59 percent of judges still allowed all of the expert evidence proffered in their last civil trial); Sophia I. Gatowski *et al.*, *Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-*Daubert *World*, 25 Law & Hum. Behav. 433 (finding in a 1998 survey of state trial court judges that 59 percent thought that the intent of *Daubert* was either not to change the threshold for admissibility for expert testimony, or to lower it).

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

[19] On *Frye* in general, see Paul C. Giannelli, *The Admissibility of Novel Scientific Evidence:* Frye v. United States*, a Half-Century Later*, 80 Colum. L. Rev. 1197 (1980) (describing some of the problems with the *Frye* test). For analysis of the problems with a general acceptance test when the field itself may be invalid, see Jennifer L. Mnookin, *Fingerprint Evidence in an Age of DNA Profiling*, 67 Brook. L. Rev. 13, 63 (2001). *See also* Michael J. Saks, *Merlin and Solomon: Lessons from the Law's Formative Encounters with Forensic Identification Science*, 49 Hastings L.J. 1069, 1078–79 (1998).

[20] Fed. R. Evid. 702 (emphasis added).

[21] 526 U.S. at 141.

[22] Fed. R. Evid. 702 (emphasis added).

[23] Fed. R. Evid. 403.

[24] Fed. R. Evid. 702.

[25] *See* Samuel R. Gross, *Expert Evidence*, 1991 Wis. L. Rev. 1113, 1163–65.

[26] *Id.* at 1187–1208 (describing both benefits and obstacles to use of neutral experts).

[27] *See supra* notes 2–7 and accompanying text.

[28] David H. Kaye & David A. Freedman, *Reference Guide on Statistics*, *in* Reference Manual on Scientific Evidence 83, 102 (2d ed. 2000).

[29] *Id.*

[30] *Daubert*, 509 U.S. at 590 n.9 ("We note that scientists typically distinguish between 'validity' (does the principle support what it purports to show?) and 'reliability' (does application of the principle produce consistent results?)… . [O]ur reference here is to *evidentiary* reliability—that is, trustworthiness.") (emphasis in original) (citations omitted).

[31] Fed. R. Evid. 702.

[32] *See generally* Gross, *supra* note 25, at 1158–62.

[33] *See id.* at 1165–76, 1187–1211.

[34] *See id.* at 1158 n.139 (citing sources).

[35] Some scholars have suggested that courts have implicitly used a 'marketplace test,' inquiring whether the expert could make a living selling his professed expertise. *See* Faigman *et al.*, *Check Your Crystal Ball*, *supra* note 6, at 1803–05; Saks, *supra* note 19, at 1073–74.

[36] Thus, for example, there are reports of widespread expert fraud in comparatively small personal injury cases that settle on the basis of written reports from plaintiffs' experts, with minimal review from the defendants' insurance companies, and very likely none from their lawyers. *See, e.g.*, William K. Rashbaum, *Three Officers Faked Reports, Police Say*, N.Y. Times, Oct. 23, 2002, at B1 (recounting that three police officers were charged with falsifying accident reports along with several other members of a

fraud ring including individuals making false medical and insurance claims); Jerry Urban, *Metro Officials Say Fraudulent Injury Claims are Problem*, Hous. Chron., Aug. 11, 1997, at A15. We suspect that such fraud is far more difficult and less common in cases that stand any chance of going to trial. *But see Husband, Wife Get 7 Years for Fake-Injury Schemes*, Orlando Sentinel Trib., July 17, 1992, at B5. A husband and wife were sentenced after managing to win a $2.5 million malpractice settlement and $900,000 in workers' compensation for a fraudulent injury allegedly arising from a back surgery. *Id.* The scam was revealed when the "injured" husband was video taped by a private investigator walking his dog and climbing stairs. *Id.*

[37] Texas pathologist Ralph Erdmann was eventually convicted of felonies relating to autopsies that he botched and falsified. *See, e.g.*, Richard L. Fricker, *Grave Mistakes*, 79 A.B.A. J., Dec. 1993, at 46 (describing Erdmann's incompetence and quoting a police seargent as stating that Erdman treated autopsies as if they were "kindergarten classes or show and tell"); Richard L. Fricker, *Pathologist's Plea Adds to Turmoil: Discovery of Possibly Hundreds of Faked Autopsies Helps Defense Challenges*, 79 A.B.A. J., Mar. 1993, at 24 (describing Erdmann's conviction and quoting attorney appointed to investigate him as saying, "[i]f the prosecution theory was that death was caused by a Martian death ray, then that was what Dr. Erdmann reported"); Roberto Suro, *Ripples of a Pathologist's Misconduct in Graves and Courts of West Texas*, N.Y. Times, Nov. 22, 1992, at 22 (describing intimations of a police cover-up of wrongdoing and quoting investigator of Erdmann as saying, "We started digging up bodies and when we were seven for seven we decided that in the interests of judicial economy we didn't have to go further to prove that this guy was a liar").

[38] The most notorious example is West Virginia serologist Fred Zain, who even testified to tests his laboratory laced the equipment to perform. *See generally In re* Investigation of the W. Va. State Police Crime Lab., Serology Div., 738 S.E.2d 501, 503–04 (W. Va. 1993) (describing Zain's misconduct and developing procedures and standards for evaluating cases in which Zain's testimony helped procure conviction); Jim Dwyer *et al.*, Actual Innocence: Five Days to Execution, and Other Dispatches from the Wrongly Convicted 109–17 (2000) (describing in detail Zain's mishandling and creation of scientific evidence); Stephanie Martz, *Judge's Report Closes Investigation of Zain*, Charleston Gazette, Feb. 25, 1994, at 1D (reporting that Zain was found to have fabricated blood test results in at least thirty-six cases); Sarah Webster, *Officials Revive Case Against Zain: Another Special Prosecutor Appointed to Pursue Case*, Charleston Daily Mail, Oct. 4, 1996, at 1C (describing incriminating evidence against Zain said to include two coworkers who saw Zain fake data in nearly 100 cases).

[39] *See, e.g.*, Robert Tanner, *Standards, Autonomy Sought for Crime Labs Mishandling of Evidence; Mismanagement Cited*, Seattle Times, July 7, 2003, at A1 (summarizing recent problems with crime labs). Oklahoma City forensic scientist Joyce Gilchrist presents another recent example. *See, e.g.*, *Memo Says Police Lab Manager Botched Evidence*, Tulsa World, May 2, 2001, at A8; Ed Timms & Diane Jennings, *Concern Grows Over Use of Flawed Evidence; Chemist's Case Raises Fears of Problem's Scope*, Dallas Morning News, May 13, 2001, at A43; Jim Yardley, *Inquiry Focuses on Scientist Employed by Prosecutors: Chemist Handled 5,000 Oklahoma Cases*, N.Y. Times, May 2, 2001, at A14 (reporting that the Governor of Oklahoma ordered all felony cases in which Gilchrist was involved to be re-examined); *see also* McCarty v. State, 765 P.2d 1215, 1218, 1222 (Okla. Crim. App. 1988) (reversing and remanding capital conviction on account of Gilchrist's poor lab work). There was also a year-and-a-half long investigation into problems at the FBI crime lab. *See, e.g.*, U.S. Dep't of Justice, Office of the Inspector General, The FBI Laboratory: An Investigation into Laboratory Practices and Alleged Misconduct in Explosives-Related and Other Cases (1997); Richard A. Serrano, *Misconduct Allegations Arise in FBI Lab Probe*, L.A. Times, Jan. 30, 1997, at A1. Most recently, a major scandal has been uncovered at the Houston Police Department crime lab. *See, e.g.*, James Kimberly, *Independent Lab Probe Given Push; City, County Officials Say Inquiry Would Help Rebuild Faith in System*, Hous. Chron., Apr. 30, 2003, at 21, *available at* http://www.chron.com/cs/CDA/ssistory.mpl/special/crimelab/188958

[40] *See generally* Dwyer *et al.*, *supra* note 38; Gary Taylor, *Fake Evidence Becomes Real Problem*, Nat'l L. J., Oct. 9, 1995, at A1. An even more widespread problem may be the biases that result from the close

relationship between forensic experts and the police and prosecution, and the fact that forensic experts frequently know in advance what it is that the investigators hope that they will find. For a general discussion of this issue, see D. Michael Risinger *et al.*, *The* Daubert*/*Kumho *Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion*, 90 Cal. L. Rev. 1 (2002).

[41] By contrast, Risinger uses "translation" in a broader sense to describe almost all expert evaluations about the facts of the particular case that the trier of fact is asked to accept on the basis of the expert's authority. *See* Risinger, *Preliminary Thoughts*, *supra* note 8 at 518–25.

[42] *See* Phylis Skloot Bamberger, *The Dangerous Expert Witness*, 52 Brook. L. Rev. 855, 867 (1986).

[43] On the internally-created dispute resolution mechanisms within the diamond industry, see Lisa Bernstein, *Opting Out of the Legal System: Extralegal Contractual Relations in the Diamond Industry*, 21 J. Legal Stud. 115 (1992).

[44] Sometimes obscure shorthand is used by insiders for the specific purpose of making their views obscure to outsiders. For example, in a public defender's office where one of us worked as a student, "WPD" meant "white punk on dope,"—a shorthand that facilitated messages, on the *outside* of file folders, such as: "D—WPD V—WPD Will settle." Similarly, we're told that in some car dealerships the code "ESO" on a form sent back from a mechanic to a "service representative" means "Equipment Superior to Operator"—*i.e.*, "The idiot customer is complaining because he's too stupid to operate the machine."

[45] Fed. R. Evid. 604.

[46] Court Interpreters Act, 28 U.S.C. §1827(d)(1) (2003); United States v. Armijo, 5 F.3d 1229, 1234–35 (9th Cir. 1993); United States v Taren-Palma, 997 F.2d 525, 531–32 (9th Cir. 1993).

[47] *Taren-Palma*, 977 F.2d. at 532.

[48] *See, e.g.*, United States v. Ben-Shimon, 249 F.3d 98, 101 (2d Cir. 2001) ("If the accuracy of the transcript is contested, competing transcripts may be submitted to the jury."); United States v. Chalarca, 95 F.3d 239 (2d Cir. 1996).

[49] 28 U.S.C. §1827(e)(1).

[50] *See* United States v. Barnett, Nos. 02–4561, 02–4732, 2003 U.S. App. LEXIS 6953, at *3–5 (4th Cir. Apr. 11, 2003) (describing how a special agent was permitted to testify as an expert witness with regard to drug code); United States v. Ceballos, 302 F.3d 679, 685 (7th Cir. 2002) (recounting that the government presented two agents as experts on drug code language who testified as to the content of several taped conversations played for the jury); United States v. Dukagjini, 326 F.3d 45, 53 (2nd Cir. 2001) (allowing an agent to serve as a dual witness testifying as a case agent and as an expert on drug code); United States v. Evans, 272 F.3d 1069, 1096 (8th Cir. 2001) (upholding the admittance of a sergeant's testimony regarding prostitution, including jargon); Bamberger, *supra* note 42.

[51] *See Dukagjini*, 326 F.3d at 53–56; Bamberger, *supra* note 42, at 866–69 (discussing *United States v. Young*, 745 F.2d 733 (2d Cir. 1984)). Though courts express concern about expert evidence bolstering the lay evidence, the inverse is also possible: if the factfinder believes the lay witness's eyewitness account, she may be more inclined to trust her expert judgment as well.

[52] For a detailed analysis of experts as educators, see Ronald J. Allen & Joseph S. Miller, *The Common Law Theory of Experts: Deference or Education*, 87 Nw. U. L. Rev. 1131 (1993).

[53] Edward Imwinkelried has made a similar distinction between what he terms the expert's major and minor premises: "The major premise is a principle, procedure, or explanatory theory derived by the inductive, scientific technique. The [expert] applies that major premise to the facts of the case" and these specific facts are the witness's minor premise. Edward J. Imwinkelried, *The "Bases" of Expert Testimony: The Syllogistic Structure of Scientific Testimony*, 67 N.C. L. Rev. 1, 2–3 (1988); *see also* Edward J. Imwinkelried, *The Educational Significance of the Syllogistic Structure of Expert Testimony*, 87 Nw. U. L. Rev. 1148 (1993).

[54] The case of *Treadwell v. Nickel*, 228 P. 25 (Cal. 1924), describes an extreme example: an eighty-three page hypothetical question, followed by a fourteen page objection.

[55] Some judges believe that even permitting expert testimony on the lack of relationship between confidence and accuracy invades the province of the jury. In *Garden v. State*, 815 A.2d 327, 338 (Del. 2003), the trial court rejected the portion of the eyewitness identification expert's testimony dealing with the confidence/accuracy relationship "because such opinion would amount to a comment on the veracity of the witnesses who testified and thus would invade the province of the jury." The Delaware Supreme Court found the exclusion to be error, albeit harmless. *Id.*

[56] Fed R. Evid. 201.

[57] *See* Gross, *supra* note 25, at 1134. With most witnesses, parties are constrained by circumstances: they have no choice but to use those who, because of their physical relationship to the facts, have uniquely valuable information. (The personal knowledge rule operates as a proxy for information quality, supplemented by extensive (though not necessarily effective) mechanisms for testing witness credibility to weed out percipient but unreliable observers.) In the context of factual instruction, where the issues are ones of general knowledge, parties are not limited by the happenstance of percipient witnesses. Since they can call any qualified expert who is willing to cooperate, they can, in principle, get the best available evidence in every case. In practice, their dominant incentive is to get the most useful expert testimony to support their position, not the most accurate.

[58] Some courts permit jurors to ask questions of witnesses through the judge, but this varies by jurisdiction and the temperament of the judge. Judges can, of course, question witnesses.

[59] Fed. R. Evid. 702. Rule 702 was revised in 2000 in the wake of *Daubert* and *Kumho Tire* to emphasize the reliability requirement for expert testimony.

[60] *Id.*

[61] *See* Paul C. Giannelli, *The Abuse of Scientific Evidence in Criminal Cases: The Need for Independent Crime Laboratories*, 4 Va. J. Soc. Pol'y & L. 439, 457–62 (1997) (discussing Louise Robbins's so-called expertise).

[62] 522 U.S. 136 (1997).

[63] Joiner v. Gen. Elec. Co., 864 F. Supp. 1310, 1326 (N.D. Ga. 1994) (internal citation omitted).

[64] *Joiner*, 522 U.S. 136 (affirming the district court's exclusion and reversing a decision by the Eleventh Circuit).

[65] 522 U.S. at 146–47.

[66] Gross, *supra* note 25, at 1184 n.217.

[67] *Daubert*, 727 F. Supp. at 574 (emphasis added).

[68] *Daubert*, 951 F.2d at 1129.

[69] *See, e.g.*, Norfolk & W. Ry. v. Ayers, ___U.S. ___, 123 S. Ct. 1210, 1222 (2003) ("There is an undisputed relationship between exposure to asbestos sufficient to cause asbestosis, and asbestos-related cancer.").

[70] *See, e.g.*, Meister v. Med. Eng'g Corp., 267 F.3d 1123, 1132 (D.C. Cir. 2001) ("[T]he defendants introduced expert testimony that was supported by a uniform body of evidence including epidemiological studies failing to establish a causal link between silicone breast implants and connective tissue disease… . Hence, the district court could reasonably conclude that reasonable people could not differ as to the import of the epidemiological evidence."); Pozefsky v. Baxter Healthcare Corp., 92-CV-0314 (LEK)(RWS), 2001 U.S. Dist. LEXIS 11813, at *20 (N.D.N.Y. Aug. 16, 2001) ("Because Plaintiff has failed to establish the scientific reliability of his theories on causation and the overwhelming weight of scientific authority to the contrary, the Court will join these courts and exclude the testimony of [Plaintiff's expert] as to the causation of Plaintiff's alleged [systemic] condition by her silicone breast implants.").

[71] Fed. R. Evid. 702.

[72] On the history of handwriting identification, see Mnookin, *supra* note 12.

[73] Giannelli, *supra* note 61.

[74] On the transformation of standards for proving causation, see Finley, *supra* note 16, and Sanders & Machal-Fulks, *supra* note 16.

[75] *See* Fed. R. Evid. 605 (prohibiting judges from testifying as witnesses); Fed. R. Evid. 606 (prohibiting jurors from testifying as witnesses).

[76] Fed. R. Evid. 602.

[77] Fed. R. Evid. 701.

[78] 140 F.3d 381 (2d Cir. 1998).

[79] *Id.* at 383.

[80] Zuchowicz v. United States, 870 F. Supp. 15 (D. Conn. 1994), *aff'd*, 140 F.3d 381 (2d Cir. 1998).

[81] *Zuchowicz*, 140 F.3d at 384.

[82] *Id.* (citation omitted).

[83] *Id.* at 384–85 (internal citations omitted). In the context of this case, the instructional testimony about the causal relation between PPH and other drugs was uncontroverted, part of the generally accepted background information about the condition. If one of these other causal relations were at issue in the case, what was taken as true in *Zuchowicz* might instead have been challenged as insufficiently established.

[84] *Id.* at 385.

[85] *Id.*

[86] *Id.* at 389.

[87] *Zuchowicz*, 140 F.3d at 385.

[88] *Id.* (emphasis in original).

[89] *Id.* at 386.

[90] The trial court's opinion mentioned one of the defendant's experts in passing, noting that this expert acknowledged one of the plaintiff's experts to be an expert in pulmonary hypertension. *Zuchowicz*, 870 F. Supp. at 18.

[91] *See* Aspiazu v. Orgera, 535 A.2d 338, 342 (Conn. 1987); Shelnitz v. Greenberg, 509 A.2d 1023, 1027 (Conn. 1986). *Aspiazu* was cited in *Zuchowicz* by the Second Circuit, 140 F.3d at 389, and by the trial court, 870 F. Supp. at 16.

[92] If the plaintiff would have been equally likely to get PPH if she had received the proper dose, then the pharmacy's negligence would not have been the cause of her illness.

[93] 509 U.S. at 590 n.8.

[94] 526 U.S. 137, 141 (1999).

[95] 509 U.S. at 583.

[96] 526 U.S. at 143.

[97] *Id.* at 150.

[98] *Id.* at 154 (emphasis omitted).

[99] Gross, *supra* note 25, at 1119; Samuel R. Gross & Kent D. Syverud, *Don't Try: Civil Jury Verdicts in a System Geared to Settlement*, 44 UCLA L. Rev. 1, 31–32 (1996).

[100] *Zuchowicz*, 140 F.3d at 385.

[101] *Id.*

[102] *Id.*

[103] As the trial court notes, another one of the plaintiff's experts deemed the theory "very plausible." *Zuchowicz*, 870 F. Supp. at 20. Obviously questions of partisan solidarity—a form of bias—arise when assessing the weight due to corroboration by another expert testifying for the same side.

[104] *Zuchowicz*, 140 F.3d at 385.

[105] There are no other reported cases regarding a causal relation between Danocrine and PPH. For an example of a disputed issue of causation on which courts are currently reaching contradictory conclusions about admissibility under *Daubert*, see Mark Hansen, *"When Expert Testimony Fails the Test: District Courts Disagree on what Defines Causation Evidence in Drug Disability Cases,"* 88 A.B.A. J. 22 (2002) (examining the cases considering whether Parlodel, an anti-lactation drug, causes stroke, and finding that plaintiff's causation evidence has been excluded under *Daubert* in half of the cases and permitted in the other half).

[106] *Joiner*, 864 F. Supp. at 1326.

[107] *Daubert*, 951 F.2d at 1129.

Does *Frye* or *Daubert* Matter?
A Study of Scientific Admissibility Standards

Edward K. Cheng

Assistant Professor of Law
Brooklyn Law School

Albert H. Yoon

Associate Professor of Law and
Associate Professor of Political Science (by courtesy)
Northwestern University School of Law

We would like to thank Margaret Berger, Joe Cecil, Jenny Diamond Cheng, Ted Eisenberg, Jesse Rothstein, Aaron Twerski, Nicole Waters, and the Vanderbilt Dean's Lunch Workshop for helpful comments and discussions. Melissa Ballard, Jackie Debs, and Monica Falcone provided research assistance, and Victoria Szymczak provided library assistance. We would like to give special thanks to Stanley Drosky of the New York State Unified Court System Division of Technology; Linda Dykman and Jacek Aleksandrowicz of the Judicial Information Systems division of the Connecticut Judicial Branch; the National Center for State Courts (through ICPSR Study No. 9266), and the Federal Judicial Center (through ICPSR Study No. 8429) for providing us with the relevant data.

EDWARD K. CHENG is an Assistant Professor of Law at Brooklyn Law School. He teaches evidence, torts, and criminal law, and his scholarship focuses on the legal dimensions of scientific and technological developments. He (along with co-author Albert Yoon) was recently awarded a grant from the Project on Scientific Knowledge and Public Policy for a study on whether scientific admissibility standards make a difference in practice. The results were published in the April 2005 issue of the *Virginia Law Review* and previewed in a January 2005 article in the *American Lawyer*. Professor Cheng is also the author of *Changing Scientific Evidence*, 88 Minn. L. Rev. 315 (2003). Professor Cheng holds a B.S.E. in Electrical Engineering from Princeton University, an M.Sc. in Information Systems from the London School of Economics, where he was a Fulbright Scholar from 1997–98, and a J.D. from Harvard Law School, where he was the Articles, Book Reviews & Commentaries Chair of the *Harvard Law Review*. Following law school, he clerked for Judge Stephen F. Williams of the United States Court of Appeals for the D.C. Circuit, and he was the Searle Fellow at Northwestern University School of Law.

ALBERT H. YOON is an Associate Professor of Law at Northwestern University School of Law. Professor Yoon's scholarship focuses primarily on the empirical analyses of various tort reform measures. In 2001, Professor Yoon wrote an article entitled, *Damage Caps and Civil Litigation: An Empirical Study of Medical Malpractice Litigation in the South*, 3 Am. L. & Econ. Rev. 199-227. More recently, Professor Yoon has examined mandatory arbitration and wrote an article entitled *Mandatory Arbitration and Civil Litigation: An Empirical Study of Medical Malpractice Litigation in the West*, 6 Am. L. & Econ. Rev. 95-134 (2004). In addition to his work on tort reform, Professor Yoon has written on the federal judiciary, political parties, and voting rights, with his work appearing in the *California Law Review, Northwestern University Law Review*, *Vanderbilt Law Review* and *American Law & Economics Review*, among other journals. Professor Yoon holds a B.A. from Yale University, and a J.D. and Ph.D. in political science from Stanford University. He also served as the Senior Articles Editor of the *Stanford Law Review*. He clerked for Judge R. Guy Cole, Jr. of the United States Court of Appeals for the Sixth Circuit after graduation, and was a Robert Wood Johnson Scholar at the University of California, Berkeley prior to joining the faculty at Northwestern. At Northwestern, Professor Yoon also has an appointment in the Political Science department (by courtesy), and is a Faculty Associate at both the Institute for Policy Research and the Center for Legal Studies.

Does *Frye* or *Daubert* Matter?
A Study of Scientific Admissibility Standards

Table of Contents

Does *Frye* or *Daubert* Matter?
A Study of Scientific Admissibility Standards

This study was funded by a grant from the Project on Scientific Knowledge and Public Policy (SKAPP). Major support for SKAPP is provided by the Common Benefit Trust, a fund established pursuant to a court order in the Silicone Gel Breast Impact Products Liability Litigation, with additional support from the Alice Hamilton Fund and the Bauman Foundation. Ed Cheng would also like to thank the Brooklyn Law School Dean's Summer Research Fund for additional support.

Since it was announced by the Supreme Court in 1993, *Daubert v. Merrell Dow Pharmaceuticals, Inc.*[1] has become the foundational opinion in the modern law of scientific evidence and arguably one of the most important decisions in the area of tort reform. Over the years, the *Daubert* test for scientific admissibility has spawned countless articles, symposia, and informal discussions about its merits and drawbacks, particularly in contrast to its principal rival, the *Frye* "general acceptance" test.[2] Commentators have extensively debated which test is the stricter standard,[3] and whether either standard places decisionmaking power in the proper institution (*Frye* in the scientific community, *Daubert* in the judiciary). In addition, state supreme courts have repeatedly grappled with whether to adopt *Daubert* or maintain *Frye*.[4]

Although the practical effects of *Daubert* were initially ambiguous,[5] the enduring legacy of the *Daubert* decision is now relatively clear.[6] In federal courts, where the decision is legally binding, *Daubert* has become a potent weapon of tort reform by causing judges to scrutinize scientific evidence more closely.[7] Tort reform efforts often focus on medical malpractice, products liability, and toxic torts—all cases in which scientific evidence is likely to play a decisive or at least highly influential role. The resulting effects of *Daubert* have been decidedly pro-defendant. In the civil context, *Daubert* has empowered defendants to exclude certain types of scientific evidence, substantially improving their chances of obtaining summary judgment and thereby avoiding what are perceived to be unpredictable and often plaintiff-friendly juries.

The big question, however, is how the *Daubert* decision has affected state courts, since state courts provide the fora for the vast majority of tort litigation. As Figure 1 shows, over the years a number of states have formally adopted the *Daubert* standard. In those states, one might expect results similar to those observed in the federal context. As Figure 1 also details, however, many states have expressly rejected *Daubert* and chosen to retain the *Frye* standard. What influence, if any, has the *Daubert* decision had on these states despite being formally rejected? Does a state's adoption of a *Daubert* or *Frye* test make any difference in the way scientific evidence is handled in practice?

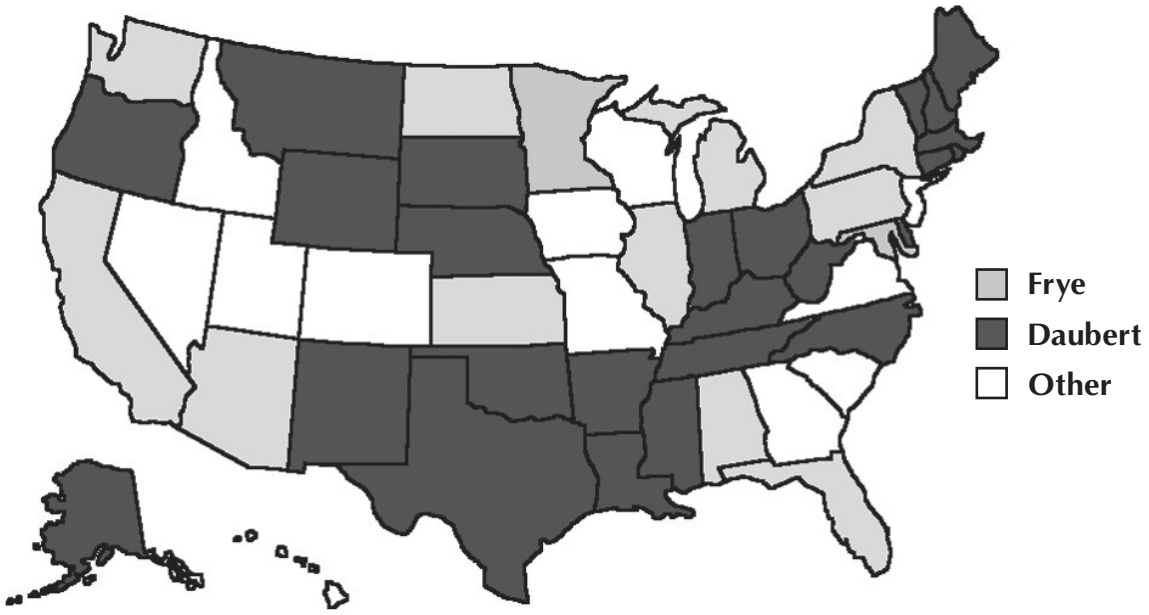

Figure 1: Geographical Map of *Frye* Versus *Daubert* States[8]

Among some commentators, there has been growing suspicion that whether a state adopts *Daubert* or *Frye* does not ultimately affect how courts handle scientific evidence. As the authors of the leading treatise on scientific evidence suggest: "Arguably…relatively few toxic tort case admissibility rulings actually turn on the difference between *Daubert* and *Frye*. *Daubert*'s shadow now casts itself over state court opinions even in jurisdictions that have not formally adopted the *Daubert* test."[9] Under this view, the real contribution of the *Daubert* decision was not in creating a new doctrinal test, but rather in raising the overall awareness of judges—in all jurisdictions—to the problem of unreliable or "junk" science.[10] Therefore, whether a jurisdiction nominally follows *Frye* or *Daubert*, the practical results are essentially the same.

This theory, if true, could have important ramifications for both the field of scientific evidence and for tort reform more generally. If courts are making scientific admissibility decisions based not on doctrinal tests but rather on other extralegal views, then the traditional focus on the merits of *Frye* versus *Daubert* may be largely misguided. Instead of debating *Frye* versus *Daubert*, perhaps research should concentrate on these "softer" extralegal mechanisms that judges use in their decisionmaking process, as well as on how best to educate the judiciary further about scientific methods and the interaction between law and science.

In addition, such a theory would caution against tort reform efforts centered on purely doctrinal changes to procedural (or evidentiary) standards. Although announcing new tests seems to be a straightforward method of changing court behavior, the judicial decisionmaking process in some cases may be too complex to be significantly affected by a vague and indeterminate standard. This is particularly true in the procedural or evidentiary context, in which trial judges tend to have broad discretion and are less subject to appellate scrutiny.

This Essay tests that theory and provides evidence on whether state court adoption of *Frye* or *Daubert* matters. Part I will begin with some background on the *Frye* and *Daubert* standards. We will then review the existing scholarship on scientific admissibility standards and discuss how it informs our research question.

Part II will explain the general difficulties of empirically testing the effect of an evidentiary standard such as *Daubert*, and will conclude that the limitations of traditional methods such as surveys or case analyses make them undesirable tools in this context. Part II will then suggest a new and potentially interesting metric based on the rates at which defendants remove cases from state to federal court. Using a removal metric enables researchers to harness the vast datasets produced and made available by the Federal Judicial Center, the National Center on State Courts, and various state court information systems departments.

Part III will present our research design and interpret the results from an initial, more limited comparison of removal rates between one *Daubert* state (Connecticut) and a geographically similar region of a *Frye* state (New York). Part IV will expand this preliminary study to encompass a much broader swath of the country, limited only by the availability of data and the determinacy of a state's admissibility standard.

Both Parts III and IV will offer strong support for the theory that the choice between a *Frye* and *Daubert* standard does not make any practical difference. Part V will discuss the ramifications and limitations of the results and will touch upon two areas for future study.

I. Background

A. Scientific Admissibility Standards

Conceptually, the admissibility requirements for scientific evidence are the same as those imposed on any type of evidence: the evidence must be both reliable and relevant. As a practical matter, however, courts have scrutinized scientific evidence more carefully, revisiting what the appropriate standards should be and who should be making that determination.

For most of the twentieth century, pursuant to *Frye v. United States*,[11] courts evaluated scientific evidence under a "general acceptance" standard.[12] In affirming the trial court's decision to exclude expert testimony regarding a lie detector test based on changes in systolic blood pressure, the D.C. Circuit held that scientific findings must "be sufficiently established to have gained general acceptance in the particular field in which it belongs."[13] The court rejected the testimony because the lie detector test "ha[d] not yet gained such standing and scientific recognition among physiological and psychological authorities."[14] Although *Frye* did have its detractors, who thought it imposed an unreasonably high standard and would serve to exclude information that jurors would find otherwise helpful to deciding cases,[15] *Frye*'s "general acceptance" emerged as the standard at trial for determining the reliability of scientific evidence.

In 1993, the Supreme Court directly addressed the reliability of scientific evidence in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*[16] In reversing the trial court's decision to preclude expert testimony on the health risks of Bendectin pursuant to *Frye's* "general acceptance" standard, the Court adopted a new framework for evaluating the reliability of scientific evidence, based on four considerations:[17] falsifiability, peer review, error rates, and "acceptability" in the relevant scientific community.[18] While not meant to be exhaustive, these factors were intended to provide guidance to the judge. Perhaps most importantly, *Daubert* established the role of the judge as a "gatekeeper" in the scientific evidence context, requiring trial courts to scrutinize the reliability of any expert evidence offered by the parties.[19]

Since *Daubert*, the Supreme Court has strengthened and broadened the gatekeeping role of the trial judge regarding scientific evidence. In *General Electric Co. v. Joiner*, the Court held that a trial judge's determinations regarding the admissibility of expert testimony were to be reviewed only for abuse of discretion by appellate courts.[20] Most recently, the Court in *Kumho Tire Company v. Carmichael* extended the four-factor test and the court's gatekeeping role to encompass all expert testimony, whether scientific or otherwise.[21]

B. Existing Scholarship

The everyday practice of law suggests that a state's adoption of *Frye* or *Daubert* should make at least some practical difference. Doctrine provides the framework by which judges analyze facts and decide cases, so changing that framework should presumably change outcomes. Nevertheless, a number of recent studies have provided some cause to believe that a state's choice of *Frye* or *Daubert* has no effect in tort cases.

1. Application of the Tests in Practice

A few studies suggest that *Daubert* courts in practice perform what is essentially a *Frye* analysis. In a 2001 study analyzing federal district court decisions, Lloyd Dixon and Brian Gill found that the "general acceptance" prong played a critical role in *Daubert* admissibility determinations in federal court.[22] Similarly, a 2001 survey by Sophia Gatowski and others reported that state court judges not only found general acceptance to be the most useful *Daubert* factor, but that state judges also had a strikingly poor understanding of other *Daubert* factors such as falsifiability and error rate.[23] Thus, while the *Daubert* decision itself may have raised judicial scrutiny of scientific evidence across the board, courts in practice engage in essentially the same analysis regardless of whether their jurisdiction is formally *Frye* or *Daubert*. Courts either do not understand the additional *Daubert* factors or simply do not find them useful. This result is particularly understandable given that many states have adopted deferential standards of appellate review for scientific admissibility determinations. The resulting discretion given to trial courts may undermine any constraints that formal evidentiary doctrine purports to impose.

2. Studies of Criminal Cases

Other evidence that *Frye* and *Daubert* may not matter in tort cases is found in a study conducted by Professor Jennifer Groscup and others in 2002. The Groscup study involved a case analysis of 372 federal and 321 state criminal appellate decisions on scientific admissibility from 1988 to 1999.[24] The study found that in criminal cases, the adoption of the *Daubert* test, whether in state or federal court, had no statistically significant effect on admission rates.[25] A more limited case analysis by Pamela Jensen published in 2003 reported similar results.[26]

Both the Groscup and Jensen studies certainly made important contributions to our understanding of the practical implications of adopting *Frye* over *Daubert*. Both studies, however, were limited to criminal cases,[27] making their results difficult to generalize to the tort context, since courts are motivated by different considerations and biases in criminal cases. In addition, both studies performed case analyses of appellate decisions, which are limited by possible selection effects and other drawbacks that Part II will discuss. Our study therefore seeks to address these limitations, as well as to fill a significant gap in the literature by looking at the effect of *Frye* versus *Daubert* in tort cases.

II. Research Metric

A. The Problems of Measurement

How does one determine whether the adoption of a *Frye* or *Daubert* standard makes a difference? Unfortunately, the effect of a scientific admissibility standard can be extremely difficult to measure. Traditional methods such as case analyses, surveys, or various other quantitative measures, while helpful and informative, have significant limitations, and so we ultimately employed a different methodology for this study.

1. Case Analyses

Case analyses, while often a powerful tool for observing and interpreting the behavior of appellate courts, face significant difficulties in the scientific admissibility context. Scientific admissibility determinations are evidentiary rulings, and so unlike most other important legal decisions, their primary forum is the trial court. In addition, because the vast majority of tort litigation occurs in state courts, research therefore must focus on *state trial courts*. Unfortunately for researchers, however, very few state trial court opinions are published or are available on electronic database services such as Westlaw or Lexis-Nexis.[28]

Case analyses could instead focus on state appellate decisions, but those studies would necessarily suffer from potential selection bias. Furthermore, to the extent that appellate courts focus on establishing bright-line rules regarding the admissibility of broad types of evidence, reading those opinions alone may neglect more subtle influences that admissibility standards can have, such as their effect on the level of scrutiny that trial judges impose on an everyday basis. Case analyses also cannot switch to federal district court decisions because all federal courts operate on the *Daubert* standard, eliminating any basis for comparison.

Case analyses also do not observe instances in which the parties settle early in the litigation process.[29] The vast majority of civil cases never go to trial,[30] and a substantial number never proceed to a stage at which formal opinions are likely to be written, creating more potential bias effects.

Finally, case analyses involve reading published decisions, which are stylized communications that may not necessarily provide an accurate, unadulterated look into the actual judicial decisionmaking process. Courts may—consciously or unconsciously—fail to discuss certain considerations in their opinions, leaving researchers only with a sanitized view. Case analyses also often require the subjective interpretation and coding of decision texts.

2. Surveys

Surveys offer the potential advantage of richer responses and discussions, but are of limited use because they rely heavily on the respondents' ability to recall past experiences truthfully and accurately. Our research question requires assessments of a somewhat vague concept (the scrutiny given to scientific evidence), over the significant period of time before and after a state's adoption of a new scientific standard, in a great variety of cases. It is therefore unclear how much knowledge could be gained from a survey. While a survey could certainly be helpful for ascertaining an impressionistic view of whether attorneys perceive a difference between *Frye* and

Daubert jurisdictions, its usefulness is limited. Surveys naturally also suffer some selection bias effects based on the willingness of participants to respond.

3. Basic Quantitative Measures

Beyond case analyses or surveys, one could track various quantitative measures to study the effect of a switch from *Frye* to *Daubert*. For example, changes in the number of favorable or unfavorable admissibility rulings in a jurisdiction could suggest a tightening or loosening of scrutiny. The problem, however, is that admissibility rulings are dependent not only on the governing admissibility standard, but also on the perceived validity or strength of the scientific evidence in question. Whether certain types of evidence are found admissible or inadmissible can therefore be significantly time-dependent, because the underlying scientific basis can improve (or decline) over time as new studies are conducted.[31]

Observing changes in final damage awards presents similar problems. Damage award data is subject to the censoring effects of settlements, which are generally sealed.

B. The Removal Metric

In an attempt to address some of the above concerns, our study develops what we hope to be a promising new metric for understanding the effect of scientific admissibility standards. Rather than observing case decisions, tallying admissibility determinations, or conducting surveys, our study measures the effect of *Frye* versus *Daubert* by using the rate at which defendants choose to remove cases from state to federal court.[32]

1. Review of How Removal Works

Generally speaking, tort claims are only actionable under state law and therefore must be litigated in state courts. However, when the parties in a lawsuit are citizens of different states, either party has the option of forcing the lawsuit into federal court under diversity jurisdiction.[33] The reasons for litigating in federal court vary—for example, preference for federal procedural rules, concern about out-of-state bias, perceived quality of the federal judiciary, and so forth.[34]

If the plaintiff chooses to file her claim in federal court, the procedural issues are straightforward. However, if the plaintiff files the claim in state court, and the defendant wishes to litigate in federal court, the defendant must remove the case. In order to remove, the defendant must file a motion with the appropriate federal court within thirty days of being served with process.[35] The federal court then transfers the case from the state court and asserts jurisdiction over it.

2. Admissibility Standards and Removal

A change in scientific admissibility standards is likely to affect removal rates considerably. First, scientific admissibility determinations are not sporadic or isolated instances, but rather are implicated in the vast majority of tort cases.[36] Thus, while not all tort cases are removable because of the diversity-of-citizenship requirement, among those that are, scientific evidence is likely to be involved.

Second, in cases that involve scientific evidence, the governing standard is likely to play a major role in defendants' decisions to remain in state court or remove to federal court. Under the

Supreme Court's well-established doctrine in *Erie Railroad Co. v. Tompkins*, parties litigating tort claims in federal court are governed by the same substantive tort law as those in state court.[37] The primary potential legal advantages of litigating in federal court are therefore procedural.[38] Scientific admissibility standards, however, while technically procedural, have a significantly substantive cast, since the inability to introduce certain types of scientific evidence can severely undermine a litigant's substantive case and result in an adverse judgment.[39] Consequently, one would expect parties, particularly those in products liability and similar tort litigation—where scientific evidence often plays a major role—to care deeply about the governing scientific admissibility standard.

3. Advantages

The removal metric also provides several advantages over other methods. While the removal metric is by no means perfect, it has different attributes that, in combination with previous case and survey studies, will help to produce a more comprehensive picture.

For example, unlike case analyses and other outcome-dependent measures, the removal metric observes cases at a much earlier stage of the litigation process. As mentioned previously, defendants must remove within thirty days of being served. Removal thereby captures a larger and more representative sample of cases. It suffers less from the selection bias of appeal or publication and is also sufficiently early to avoid most of the censoring caused by sealed settlements.

Additionally, the removal metric measures the effect of admissibility standards by what attorneys do, rather than what they say. As a result, it avoids the concerns in surveys about inaccurate recall, and the problem in case analyses of less-than-candid judicial opinions. The removal metric measures the law in action, rather than the law on the books.

Finally, unlike studying actual admissibility decisions, the removal metric minimizes concerns about the effect of changes in the strength of scientific evidence over time. We can expect removal decisions to be made with minimal regard to the underlying facts in a case. Litigants are likely to seek favorable forums regardless of the strength of their specific case.[40]

III. Preliminary Study: New York and Connecticut

A. Scope

As a preliminary study, we analyzed removal rates in tort cases filed from 1994 to 2000 in the geographical areas defined by the federal district courts for the Eastern District of New York ("EDNY") and the District of Connecticut. The EDNY region is comprised of Kings, Nassau, Queens, Richmond, and Suffolk counties in New York. The District of Connecticut encompasses the entire state of Connecticut. Figure 2 shows both districts.

The EDNY and the District of Connecticut were chosen because they set up somewhat of a "natural experiment" for studying scientific admissibility standards.

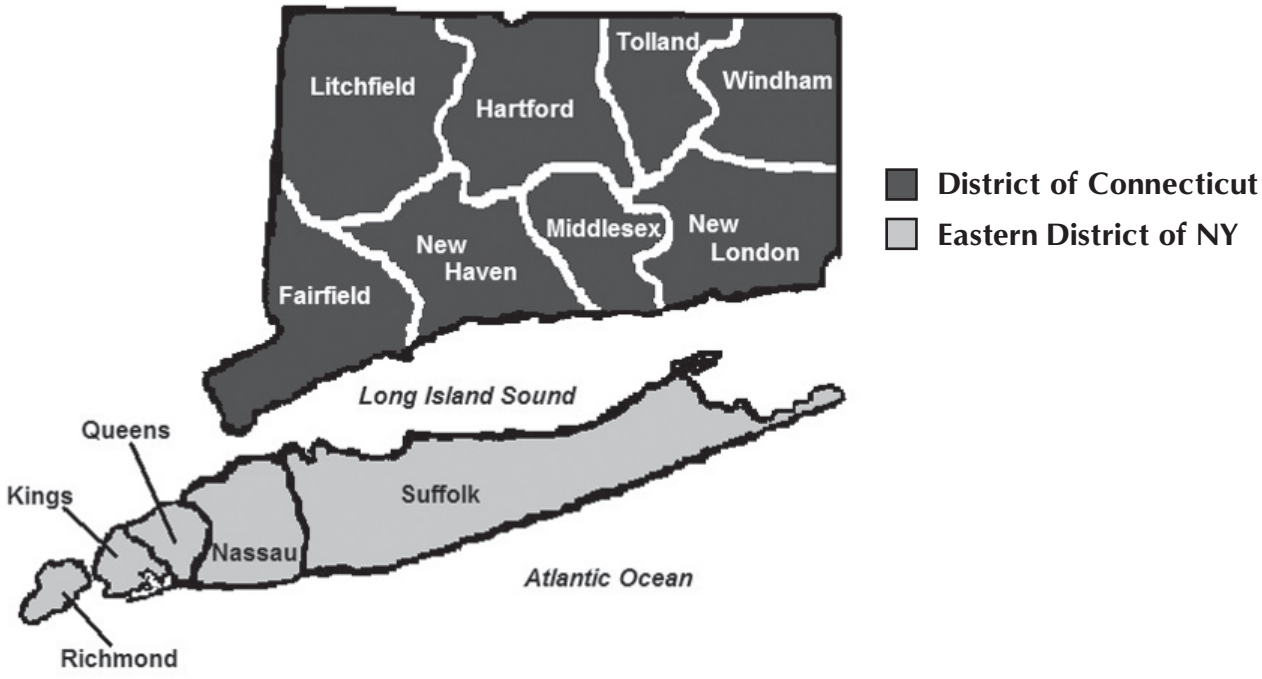

Figure 2: Counties Chosen for Preliminary Study

New York state courts consistently adhered to the *Frye* standard throughout the entire period from 1994 to 2000 (and indeed, continue to adhere to *Frye* today).[41] Connecticut state courts, however, followed *Frye* until May 1997, when they switched to the *Daubert* standard.[42] Federal courts in both states, of course, have applied the *Daubert* standard since 1993. Consequently, the removal rates in Connecticut serve as the treatment group (that is, the group affected by the policy change), while the removal rates in EDNY serve as a convenient control group (that is, a comparable group not affected by the policy change). If a state's adoption of *Frye* and *Daubert* has a practical impact, all else being constant, we would expect removal rates to change in Connecticut after 1997 because defendants would have significantly different incentives to remove.[43] The scientific admissibility standards for the various jurisdictions during the dates studied are shown in Figure 3.

	Period	
Jurisdiction	1994–1997	1997–2000
New York State Courts	*Frye*	*Frye*
Connecticut State Courts	*Frye*	*Daubert*
Federal Courts (both in New York and Connecticut)	*Daubert*	*Daubert*

Figure 3: Scientific Admissibility Standards by Jurisdiction

EDNY and Connecticut actually provide a rather compelling comparison because of the similarities between the regions. For one thing, Connecticut and EDNY are geographically proximate, essentially comprising the northern and southern shores of Long Island Sound. Both are well-connected to New York City, generally considered to be part of the New York metropolitan area, and demographically similar. Thus, exogenous factors, such as political, economic, or social

changes affecting one region are likely (or as likely as can be found for any two federal districts) to affect the other. Other states or federal districts encompass much larger geographic regions, creating problems such as multiple metropolitan areas with different political or economic environments, demographic variations, and so forth.

B. Calculation and Data Collection

1. Definition of Removal Rate

As seen in Figure 4, removal rates were defined to be the ratio of the number of tort cases removed to federal court under diversity jurisdiction in a given year and geographical area to the total number of tort cases filed in the state courts of that area.[44]

$$\text{Removal Rate} = \frac{\text{Number of tort cases removed}}{\text{Total number of tort cases}}$$

Figure 4: Removal Rate Formula

This definition of "removal rate" does not describe the rate at which "removable" cases in fact remove. Many of the cases counted in the denominator are not removable, often because the parties fail the diversity-of-citizenship requirement. What is critical, however, is that the denominator accounts for relative changes in caseloads from year to year or from jurisdiction to jurisdiction, which would otherwise skew a metric based on raw numbers alone.

The definition of "year" also necessarily involves some imprecision. The denominator (total number of tort cases) is determined by the number of cases filed in state court during the given calendar year. The numerator (number of tort cases removed) is determined by the number of cases removed to federal court during the given calendar year. Because there is some delay between state court filing and removal to federal court, some cases filed in state court during one year will be removed to federal court during the next. There appears to be no reason, however, why these numbers would not average out over the long term, or why the method would exert any biasing effect.

2. Data Collection

To calculate the removal rates, we gathered data for each jurisdiction. For removed cases in the two federal district courts, data collection was simplified by using the Federal Court Cases: Integrated Data Base created by the Federal Judicial Center and publicly available through the Inter-University Consortium for Policy and Social Research ("ICPSR").[45] The full ICPSR database provides information on every civil and criminal case filed in federal court between 1970 and 2002. Given our research design, we extracted only tort cases filed in the EDNY and the District of Connecticut for the period of 1994 to 2000.

To determine the total number of cases filed in New York state courts in the EDNY region, we obtained data from the Technology Division of the New York State Unified Court System. Extracting only tort cases filed in the five counties associated with EDNY provided the required information. We acquired similar data from the Judicial Information Systems division of the Connecticut Judicial Branch to determine the relevant Connecticut figures.

C. Results

Figure 5 shows the raw numbers and the calculated removal rates for Connecticut and the EDNY. Figure 6 graphs the two removal rates. The dotted line in Figure 6 represents the year (1997) in which Connecticut switched from the *Frye* to the *Daubert* standard.

If Connecticut's change from *Frye* to *Daubert* had an impact, we would expect a change in Connecticut's removal rate relative to EDNY's removal rate after 1997. (Recall that EDNY is acting as the control group, so all assessments of Connecticut's removal rate must be made relative to it.) Removal rates for both Connecticut and EDNY, however, appear to move in lockstep between 1997 and 1998.

	Connecticut			**EDNY**		
Year	**Filed in State Court**	**Removed to Federal Court**	**Removal Rate**	**Filed in State Court**	**Removed to Federal Court**	**Removal Rate**
1994	16172	56	0.35 %	42120	207	0.49 %
1995	18417	64	0.35 %	46199	237	0.51 %
1996	20165	48	0.24 %	47711	333	0.70 %
1997	20295	49	0.24 %	47235	263	0.56 %
1998	20054	63	0.31 %	46808	288	0.62 %
1999	18845	52	0.28 %	45838	310	0.68 %
2000	18201	56	0.31 %	43964	362	0.82 %

Figure 5: Removal Rates for D. Conn. and EDNY, 1994–2000

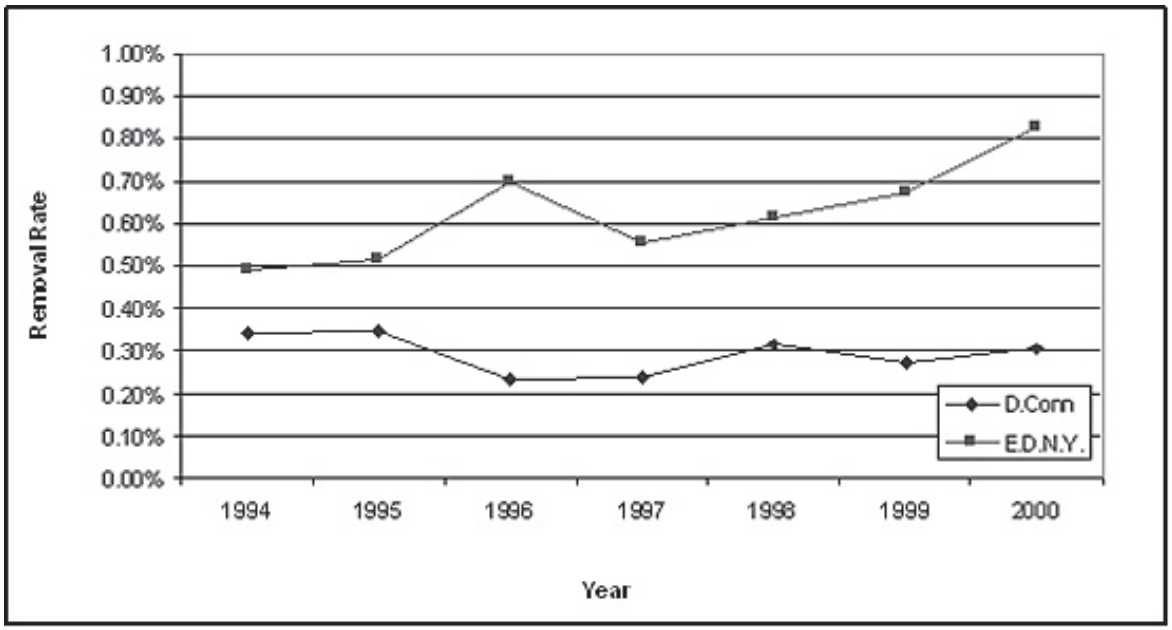

Figure 6: Removal Rates for D. Conn. and EDNY, 1994–2000

Looking more broadly, the removal rates in both states appear relatively stable over the entire period. The removal rate in EDNY does have a slight upward trend, but the difference between Connecticut's removal rates (again, relative to EDNY) in the pre-*Daubert* period (1994–

1996) and the post-*Daubert* period (1998–2000) is not statistically significant.[46] This result suggests that the change in Connecticut from *Frye* to *Daubert* did not have any significant effect detectable by this model.

D. Possible Refinements

1. Types of Torts

A natural extension of this study would involve breaking down the removal rates into smaller subsets. For example, because one might expect removal rate to vary by type of tort (medical malpractice, automobile accidents, products liability, and so forth), separating the aggregate data could shed further light on the effect of scientific admissibility standards. After all, one would expect expert evidence to be given far greater weight in a products liability case than any other tort case. Products liability litigation may also offer more opportunities for removal, since the defendant is often an out-of-state manufacturer; automobile accidents, in contrast, are ordinarily between two in-state drivers.

Unfortunately, our datasets were ultimately too inconsistently coded at the "tort-type" level to enable further analysis. While the state and federal data all had a products liability category, the jurisdictions appeared to either define "products liability" differently, or use the category irregularly. As a result, attempts to measure removal rate often yielded percentages above 100 %, suggesting either coding errors or different approaches in coding by state and federal data compilers.

2. Other States

As previously mentioned, studying two well-matched jurisdictions such as Connecticut and the EDNY has a number of advantages, especially the presumed presence of a control group. One difficulty, however, is that there are relatively few data points. While the number of cases involved in constructing the removal metric is enormous, ultimately there is only one removal rate per year for each jurisdiction. This small number of data points inhibits our ability to control precisely for single-year variations.

Another limitation of a two-state comparison is that removal rates may be affected by unobserved variables that differ between Connecticut and EDNY. For example, perhaps Connecticut's adoption of *Frye* over *Daubert* does make a difference, but Connecticut experienced some political or legal change at approximately the same time that cancelled out any accompanying removal rate effect.[47] We selected Connecticut and the EDNY specifically to minimize these types of asymmetric changes, but unfortunately the study's construction can only do so much.

The best method to remedy both of these deficiencies is to expand the inquiry to include as many states as possible. Such a study would generate more data points for econometric analysis, control for regional variations, and reduce the likelihood that some unique political or other change in a particular jurisdiction distorts the results. The next Part does just that.

IV. National Study, 1990–2000

In this Part, we expand the scope of our study to look across the country. By increasing the number of states (and therefore the number of data points), we are able to control for varia-

tions from year to year as well as from state to state, enabling us to better isolate the effect of the doctrinal admissibility standard. Other unobserved variables average out over the various states located in different geographic regions. This broader approach provides more definitive conclusions on whether a state's choice of admissibility standard has any practical effect.

A. Data Sources and Selection

1. Data Sources

To obtain data on the total number of cases removed to federal court, we once again relied on the Federal Court Cases database created by the Federal Judicial Center and publicly available through ICPSR.[48] Given our research goals in this national study, however, we retained tort cases from all jurisdictions. We combined figures wherever appropriate—for example, to ascertain the number of removed cases in Pennsylvania, we combined data from the Western and Eastern Districts of Pennsylvania. In addition, because the state court data discussed below was available for 1985 to 2001, we were able to expand the time period to encompass 1990 through 2000 for observational purposes, although, for reasons discussed below, the econometric analysis was still confined to 1994 through 2000.

For state court data, the study relied on the State Court Statistics, 1985 to 2001 dataset created by the National Center for State Courts and also publicly available through ICPSR.[49] This dataset includes summary statistics on all state court systems from 1985 to 2001 whenever such statistics are available. Therefore, unlike the Connecticut, EDNY, and federal datasets, the state court dataset does not have case-level information. For our purposes here, however, the aggregate level data was sufficient.

2. States Selected

The states ultimately included in the national study are presented in Figure 7. Not all states were appropriate for use in this national study. While we kept as many states as possible, we had to exclude states on the basis of two criteria: data availability and the existence of a clear *Frye* or *Daubert* standard. Many states did not have statistics on total tort cases filed for the entire period, or the National Center for State Courts reported that the statistics were incomplete or overinclusive. Using a state that had only reported some cases or only had complete data from 1997 to 2000 could undesirably skew results. As a result, we required that any state used in the national study have complete data going back at least as far as the *Daubert* decision in 1993.

Perhaps even more importantly, some states did not have a clearly defined scientific admissibility standard for some portion of the period, while others had a standard that was neither *Frye* nor *Daubert*. Since these states would invariably lead to coding errors if

State	Starting Year of Data	Standard	Year of Change to Daubert
Alaska	1990	Daubert	1999
Arizona	1990	Frye	N/A
Arkansas	1990	Daubert	1990
Connecticut	1990	Daubert	1997
Florida	1990	Frye	N/A
Indiana	1990	Daubert	1995
Kansas	1990	Frye	N/A
Michigan	1990	Frye	N/A
Minnesota	1990	Frye	N/A
Missouri	1990	Frye	N/A
New Mexico	1993	Daubert	1993
New York	1990	Frye	N/A
North Carolina	1990	Daubert	1995
Oregon	1991	Daubert	Pre-1990
Tennessee	1990	Daubert	1997
Washington	1990	Frye	N/A

Figure 7: States Included in National Study

characterized as either a "Frye" or "Daubert" state, the study excluded these as well. It was not necessary for a state to specifically adopt Frye or Daubert by name, but the test adopted had to be substantially equivalent. Hence, for the purposes of this study, states such as Arkansas,[50] Indiana,[51] North Carolina,[52] and Oregon[53] were considered to be *Daubert* states, and Missouri[54] was considered a *Frye* state.

B. Results

1. Removal Rates

Removal rates for all the included states were calculated using the same methodology described in the preliminary study. Removal rate was once again defined to be the total number of cases removed to federal court during the calendar year divided by the total number of tort cases filed during the same year. For reference purposes, Appendix A contains the resulting data.

2. Graphical Trends

As one might imagine, a graph of the removal rates of all sixteen states is noisy and too difficult to interpret without further analysis. Examining smaller geographic regions, however, can illuminate interesting trends. For example, examining Midwestern and Plains states (Indiana, Kansas, Michigan, Minnesota, and Missouri) together yields the graph in Figure 8.

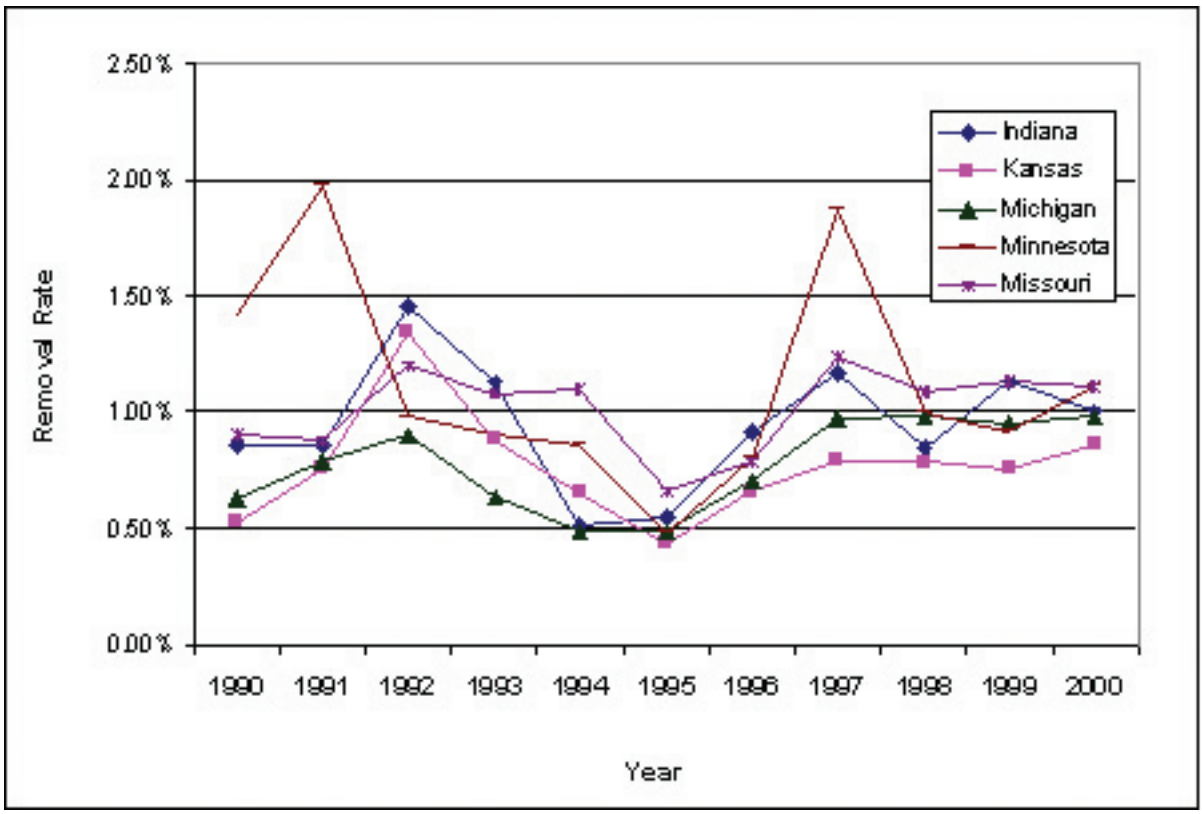

Figure 8: Removal Rates in Midwest and Plains States

All of the states in Figure 8 seem to demonstrate the same general trend, even though Indiana switched from *Frye* to *Daubert* in 1995, and the remainder of the group remained *Frye* throughout. The graph shows a relatively high removal rate until around 1993, then a steady decrease in removal rate until a nadir in 1995, and then a steady return climb in removal rate until 1997.

3. Basic Econometric Model

Ultimately, a regression best controls for year-to-year and state-to-state variations analytically, and ensures that a switch to *Daubert* is not playing some small but heretofore undetected role. For the regression analysis, we use a difference-in-differences approach, a common econometric method of program or policy evaluation.[55] This approach measures the effect of a policy when one group (treatment) is exposed to the policy and another group (control) is not. The central assumption underlying difference-in-differences is that, in the absence of the policy change, the trend in the treatment group relative to the control group would have remained the same over time. Hence, the absence of any relative change in the treatment group on the variable of interest following the policy change is probative to show that the policy did not have an effect. Econometrically, the basic model looks like the following:

$$RR = \alpha + \sum_i \beta_i YEAR_i + \sum_j \gamma_j STATE_j + \delta * DAUBERT + \varepsilon$$

Equation 1

RR represents the rate of removal (in percentage points) and is the dependent variable. The y-intercept is measured by α. $YEAR_i$ represents a series of dummies for the year *i* in which the removal rate is measured, normalizing for year-to-year effects.[56] $STATE_j$ represents a series of dummies for the various states *j* included in the study, normalizing for state effects.[57] DAUBERT is an indicator variable for whether or not the jurisdiction was following the *Daubert* standard at the time. The error term is captured by ε.

For the econometric model, the data was limited to the period from 1994 to 2000. Because *Daubert* was decided in 1993, removal decisions prior to 1994 would be based on federal courts operating under a *Frye* standard, disrupting the analysis. Furthermore, our study only investigates whether *Frye* or *Daubert* makes a difference in state courts, making pre-*Daubert* removal rates unnecessary for the regression analysis.

Running the basic econometric model in Equation 1 yields the results in Figure 9.

Regression with robust standard errors				**Number of obs = 110** **F(22, 87) = 24.11** **Prob > F = 0.0000** **R-squared = 0.8262** **Root MSE = .23301**		
	Coefficient	**Robust Std. Err.**	**t**	**P>|t|**	**[95 % Conf. Interval]**	
DAUBERT	-.0053476	.0863988	-0.06	0.951	-.1770746	.1663793
Year 1995	-.2732086	.096789	-2.82	0.006	-.4655872	-.0808301
Year 1996	-.0237572	.0765122	-0.31	0.757	-.1758335	.1283191
Year 1997	.2221668	.0975977	2.28	0.025	.0281809	.4161528
Year 1998	.1584701	.1106303	1.43	0.156	-.0614196	.3783598
Year 1999	.1317904	.0972652	1.35	0.179	-.0615347	.3251154
Year 2000	.2278502	.0826192	2.76	0.007	.0636357	.3920647
Arizona	-1.554228	.188518	-8.24	0.000	-1.928928	-1.179528
Arkansas	-.4039759	.2148389	-1.88	0.063	-.8309915	.0230398
Connecticut	-1.587598	.1885287	-8.42	0.000	-1.962319	-1.212876
Florida	-1.335077	.1859045	-7.18	0.000	-1.704582	-.9655713
Indiana	-1.012043	.1787954	-5.66	0.000	-1.367418	-.6566681
Kansas	-1.184077	.1828137	-6.48	0.000	-1.547439	-.8207152
Michigan	-1.094621	.1876164	-5.83	0.000	-1.467529	-.721713
Minnesota	-.8849699	.2208067	-4.01	0.000	-1.323847	-.4460927
Missouri	-.8732507	.1858851	-4.70	0.000	-1.242717	-.5037839
New Mexico	-.9864639	.163869	-6.02	0.000	-1.312171	-.6607566
New York	-1.095563	.1863652	-5.88	0.000	-1.465984	-.7251422
N. Carolina	-1.346426	.1681778	-8.01	0.000	-1.680698	-1.012155
Oregon	-1.153108	.214477	-5.38	0.000	-1.579404	-.7268117
Tennessee	-.6373306	.1758412	-3.62	0.000	-.9868339	-.2878272
Washington	-1.588858	.1922537	-8.26	0.000	-1.970983	-1.206733
Constant	1.827034	.1837049	9.95	0.000	1.461901	2.192168

Figure 9: Regression Results from Basic Model (1) [58]

The most important result in Figure 9, of course, is that after year and state effects have been accounted for, the DAUBERT variable—whether a state follows the *Daubert* standard in the year in question—has a vanishingly small effect on removal rate. DAUBERT contributes only five-thousandths of a percentage point to a state's removal rate, and the result is statistically insignificant. This result suggests that, in making removal decisions, defendants place little weight on whether a state follows *Frye* or *Daubert*.

A few other observations can be made about the results using the basic model. First, the various state variables all have large coefficients that are statistically significant. This result makes sense, because each state is likely to have an average baseline rate of removal, depending

on how comfortable defendants are with the state's rules of procedure, the perceived quality of the state's judiciary, and so forth. Second, most of the year variables are also statistically significant, suggesting that there are indeed year-to-year variations. Again, this finding is intuitive. Tort litigation can be trendy, focusing on a particular industry or defendant for a time and then moving on to new pastures. Removal rates for a given year may therefore reflect the particular trend in tort litigation for that year.

4. Weighted Econometric Models

One possible concern about the basic model presented in Equation 1 is that it fails to distinguish light-caseload jurisdictions from heavy-caseload jurisdictions. The removal rates observed for smaller jurisdictions will be more sensitive to small, random fluctuations in the number of removed cases (conversely, jurisdictions with heavier caseloads experience greater averaging effects). Without some weighting mechanism, swings in light-caseload jurisdictions are thus inappropriately valued equally to swings in heavy-caseload jurisdictions, where such swings are far more difficult to achieve randomly.

Weighting within the econometric model can be achieved via two alternatives. The first method is not to look at removal rate, but rather at the number of cases removed, as seen in Equation 2:

$$C_{removed} = \alpha + \sum_i \beta_i Year_i + \sum_j \gamma_j State_j + \delta * DAUBERT + \varphi * C_{filed} + \varepsilon$$

Equation 2

$C_{removed}$ represents the total number of cases removed to federal court in the jurisdiction for a given year. The remainder of the model is identical to Equation 1, except that the number of tort cases filed in a given year is controlled by using C_{filed}.

Figure 10 shows the results from a regression based on the model in Equation 2. The fit of this model appears to be far better than the basic model in Equation 1 (Equation 2 has an R2 of 0.96, while Equation 1 has an R^2 of 0.82).[59] As expected, the number of cases filed, represented by the variable TORTFILE in the regression, is a statistically significant predictor of the number of cases removed: the more cases that get filed, the more cases that will get removed.[60] Once again, many of the state and year variables are statistically significant, but, importantly, the DAUBERT variable is once again relatively small and not statistically significant.

Regression with robust standard errors				Number of obs = 110 F(23, 86) = 60.35 Prob > F = 0.0000 R-squared = 0.9617 Root MSE = 34.252		
	Coefficient	**Robust Std. Err.**	**t**	**P>|t|**	**[95 % Conf. Interval]**	
DAUBERT	-9.829792	11.56274	-0.85	0.398	-32.81577	13.15618
TORTFILE	.004749	.001473	3.22	0.002	.0018208	.0076773
Year 1995	-27.52213	15.7583	-1.75	0.084	-58.84859	3.804323
Year 1996	13.71642	11.03207	1.24	0.217	-8.214613	35.64746
Year 1997	39.86707	13.17015	3.03	0.003	13.68567	66.04847
Year 1998	26.14495	10.67405	2.45	0.016	4.925644	47.36425
Year 1999	29.08516	11.40336	2.55	0.013	6.416019	51.75429
Year 2000	44.10236	16.12342	2.74	0.008	12.05006	76.15467
Arizona	-37.98712	24.01074	-1.58	0.117	-85.7189	9.744657
Arkansas	40.64963	15.28405	2.66	0.009	10.26593	71.03332
Connecticut	-44.07844	30.34544	-1.45	0.150	-104.4032	16.2463
Florida	8.941038	67.51588	0.13	0.895	-125.2761	143.1582
Indiana	43.62383	22.26032	1.96	0.053	-.6282314	87.87589
Kansas	-2.18437	12.50121	-0.17	0.862	-27.03595	22.66721
Michigan	69.55586	38.48628	1.81	0.074	-6.952324	146.064
Minnesota	20.03074	14.83085	1.35	0.180	-9.45201	49.51349
Missouri	87.5226	29.69488	2.95	0.004	28.49113	146.5541
New Mexico	15.92776	15.00656	1.06	0.291	-13.90428	45.75981
New York	242.367	125.9804	1.92	0.058	-8.073721	492.8078
N. Carolina	-1.758002	17.80038	-0.10	0.922	-37.14398	33.62798
Oregon	11.06637	20.14878	0.55	0.584	-28.98807	51.12081
Tennessee	89.91687	20.94021	4.29	0.000	48.28911	131.5446
Washington	-36.9327	20.14384	-1.83	0.070	-76.97732	3.111927
Constant	-2.469264	11.47271	-0.22	0.830	-25.27626	20.33773

Figure 10: Regression Results from Weighted Model Two

The second method for weighting is to include analytic weights to the initial regress. These weights are the total number of tort cases filed in the jurisdiction for the year that give rise to that state's respective removal rate, and are inversely proportional to the variance of each observation. Results from this weighting method are shown in Figure 11.

Under this second method, we see that the results largely remain unchanged. As in the basic model, nearly all of the state and year variables are statistically significant. Again, the DAUBERT variable is small, contributing two hundredths of a percentage point, and not statistically significant.

Regression with robust standard errors				Number of obs = 110 F(22, 87) = 33.20 Prob > F = 0.0000 R-squared = 0.8395 Root MSE = .14857		
	Coefficient	**Robust Std. Err.**	**t**	**P>|t|**	**[95 % Conf. Interval]**	
DAUBERT	.0233748	.0684189	0.34	0.733	-.1126151	.1593647
Year 1995	-.1629444	.0543312	-3.00	0.004	-.2709335	-.0549552
Year 1996	.034418	.0542082	0.63	0.527	-.0733267	.1421627
Year 1997	.2023167	.0641396	3.15	0.002	.0748323	.3298011
Year 1998	.1223483	.0503039	2.43	0.017	.0223639	.2223327
Year 1999	.1651497	.0526122	3.14	0.002	.0605772	.2697221
Year 2000	.2663094	.0510324	5.22	0.000	.1648771	.3677418
Arizona	-1.561249	.1997495	-7.82	0.000	-1.958273	-1.164226
Arkansas	-.464657	.2549951	-1.82	0.072	-.9714875	.0421734
Connecticut	-1.615328	.2111751	-7.65	0.000	-2.035061	-1.195594
Florida	-1.338229	.1993651	-6.71	0.000	-1.734489	-.9419693
Indiana	-1.039019	.2095735	-4.96	0.000	-1.455569	-.6224686
Kansas	-1.192793	.1979847	-6.02	0.000	-1.586309	-.7992768
Michigan	-1.125059	.2034454	-5.53	0.000	-1.529429	-.7206891
Minnesota	-.8814852	.2462549	-3.58	0.001	-1.370944	-.3920269
Missouri	-.8827942	.203519	-4.34	0.000	-1.28731	-.4782783
New Mexico	-1.025804	.2094682	-4.90	0.000	-1.442145	-.6094637
New York	-1.103796	.199178	-5.54	0.000	-1.499684	-.7079083
N. Carolina	-1.375954	.2017611	-6.82	0.000	-1.776977	-.9749323
Oregon	-1.209415	.2471505	-4.89	0.000	-1.700653	-.7181765
Tennessee	-.6647566	.2060349	-3.23	0.002	-1.074273	-.2552399
Washington	-1.595988	.2031096	-7.86	0.000	-1.99969	-1.192285
Constant	1.808944	.2004615	9.02	0.000	1.410505	2.207383

Figure 11: Regression Results Using Analytic Weights

5. Summary

Figure 12 summarizes the results for the various econometric models. It also includes the results from a basic regression relating the DAUBERT variable to removal rates in which no state or year effects are controlled. Notably, DAUBERT is a statistically significant predictor of removal rate in this crude model, but any predictive effect it has disappears once year and state effects are controlled for.

	Model			
	No Controls	Equation 1	Equation 1 using Analytic Weights	Equation 2
Location of results	—	Figure 9	Figure 11	Figure 10
DAUBERT	0.1949*	-0.0053	0.0234	-9.83
	(1.99)	(0.06)	(0.34)	(0.85)
Constant	0.7782**	1.8270**	1.8089**	-2.47
	(13.29)	(9.95)	(9.02)	(0.22)
Control for Number of Cases Filed in State in Year	No	No	Yes	Yes
Control for State Effects	No	Yes	Yes	Yes
Control for Year Effects	No	Yes	Yes	Yes
N	110	110	110	110
R^2	.0348	0.8262	0.8395	0.9617

* Statistically significant at 5 % level

** Statistically significant at 1 % level

Figure 12: Summary of National Study Results[61]

V. Discussion

The results from the national study suggest that a state's choice of scientific admissibility standard does not have a statistically significant effect on removal rates (or number of cases removed). This finding may support the broader theory that a state's adoption of *Frye* or *Daubert* makes no difference in practice. Graphically speaking, as shown in Figure 9, removal rates seem to follow the same trends regardless of whether a state retains the *Frye* standard during the entire period, or as in the case of Indiana, switches to *Daubert* in the middle of the period. More importantly, however, using econometric techniques and controlling for year-to-year and state-to-state variations, the data shows that whether or not a jurisdiction follows the *Daubert* standard has no statistically significant effect on the removal rate.

A. Ramifications

1. Daubert's Influence

The results of this study are consistent with the theory that the power of the Supreme Court's *Daubert* decision was not so much in its formal doctrinal test, but rather in its ability to create greater awareness of the problems of junk science. This suggests that courts apply some generalized level of scrutiny when considering the reliability of scientific evidence, regardless of

the governing standard. If accepted, this thesis suggests that debates about the practical merits and drawbacks of adopting a *Frye* versus a *Daubert* standard are largely superfluous.[62]

One basic policy recommendation arising out of this result is that state courts should consider uniformly adopting *Daubert* as their scientific admissibility standard.[63] If *Frye* and *Daubert* do not make a difference, then the skirmishing between the champions of *Frye* and *Daubert* yields few benefits and creates more confusion than anything else. Certainly, states should feel free to experiment with entirely different standards of admissibility—for example, Utah maintains a more rigorous test for scientific admissibility.[64] These alternative standards, however, must be sufficiently different and well understood to have any hope of achieving different results. If the states are to be the laboratories of legal progress, variation and experimentation should be embraced, but having doctrinal differences in name only provides little benefit.[65]

In addition, the findings suggest that future attempts to improve the handling of scientific evidence in the courts could be more effective if advocates for rigorous use of scientific evidence shifted their focus away from tinkering with doctrinal tests and instead toward "softer" solutions that increase the judiciary's understanding of scientific concepts and processes.[66] For example, reformers instead might pay greater attention to judicial education programs and help develop official literature such as the acclaimed *Reference Manual on Scientific Evidence*.[67]

Alternatively, if reformers believe that doctrinal tests are important and would like these tests to have greater traction, they may want to concentrate on advocating for stricter appellate review standards. Although by no means universal, many state courts, perhaps following the lead of the federal courts, review scientific admissibility decisions under an abuse-of-discretion standard.[68] This deferential standard of review empowers trial judges with substantial discretion, making the exact contours of the governing doctrinal test less important than it might otherwise be. Increasing appellate scrutiny would allow finer variations in admissibility standards to have more bite.

2. Tort Reform Through Procedure

More broadly, our study suggests some caveats when implementing substantive tort reform through changes in procedural rules. Placing procedural limits on tort litigation has been quite popular of late, ranging from new scientific admissibility rules, to mandatory arbitration, to limits on class certification. Studies of these procedural mechanisms, however, suggest that they often have little or no effect on ultimate outcomes. For example, a recent study of mandatory arbitration rules has shown that they have no observable effect on plaintiff awards or on litigation time.[69] Another study on the effect of two recent Supreme Court decisions limiting class settlements has reported that the decisions have had no clear effect on class action filings in federal courts.[70]

Our results shed further light on this issue, though read in light of the existing literature on *Daubert*, our findings are more subtle. In combination with previous studies that have shown that the *Daubert* decision itself had a substantial effect on the treatment of scientific evidence in federal courts and beyond, our study suggests that *Daubert's* influence was not from its doctrinal reform, but from its educative function. Therefore, any subsequent state tort reform effort that focused on doctrinal shifts from *Frye* to *Daubert* was ineffective, because any potential benefits from *Daubert* had already been realized.

3. Removal as a Metric

From a research methodology perspective, our study also suggests that the removal metric holds much promise as a research design, particularly for those studying the effect of procedural reforms. As with scientific admissibility, many procedural reforms do their work very early in the litigation process and have subtle effects. They are therefore difficult to study through case analyses or other metrics, which, as discussed previously, suffer from censoring effects and are perhaps suited for studying more substantive legal reforms. The removal metric offers an important, useful, and much-needed alternative.

By taking advantage of the dual federal-state system, as well as the *Erie* doctrine, the removal metric also offers a method of tapping the quantitative data readily available from state and federal court systems. That data often consists primarily of basic filing information and lacks descriptive richness, necessitating the use of more complex phenomena such as removal to extract as much information as possible.

B. Limitations

In interpreting the results, one must bear in mind some of our study's limitations. Valid use of the removal metric rests on a number of significant assumptions, which we discuss below.

1. The Effect of Evidentiary Standards on Removal Rates

An assumption necessarily made when using a removal metric is that procedural rule changes will affect defendants' decisions to remove. This assumption seems reasonable. As explained previously, the *Erie* doctrine limits defendants' ability to gain substantive legal advantages by transferring to federal court. The only advantages are therefore either procedural or intangible (judiciary quality, bias, and so forth). Given that the scientific admissibility rulings have substantive effects and may themselves be outcome determinative, one would expect that they would receive serious consideration by defense attorneys.

A serious concern would arise, however, if defendants automatically removed every case they could to federal court. In other words, if the incentives for defendants to litigate in federal court are already enormous, and the percentage of *removable* cases that are being removed is near 100 %,[71] then it may be difficult to detect the effect of *Daubert* or any other procedural rule.[72]

This scenario, however, seems highly unlikely. While it is unfortunately very difficult to empirically determine the removal rate for *removable* cases,[73] survey data suggests that removal is not an automatic decision. Most notably, a 1981 survey investigating attorney forum selection strategy showed that if offered a choice, 55 % of out-of-state defense attorneys preferred to litigate in federal court and 45 % preferred to litigate in state court.[74] Other studies on diversity jurisdiction and forum selection, though not as directly relevant, similarly imply that attorneys consider removal an open question.[75] Furthermore, the fluctuations in removal rates observed in our study in and of themselves may suggest that removal is not automatic. While year-to-year changes in the number of eligible cases may account for some of the variation in removal rates seen, it seems unlikely that those random changes alone could account for all of the increases or decreases in removal observed.[76]

2. Attorney Perception

This study's ability to use changes in removal rates to measure the effect of scientific admissibility standards also critically relies on the ability of defense counsel to accurately judge the practical ramifications of the scientific admissibility standard adopted in his or her jurisdiction. Just as case analyses depend on judicial pronouncements and surveys depend on the perception of respondents, the removal metric depends on defense counsel's judgment.

This information source, however, is not only reasonable, but arguably better than court decisions or survey responses. Courts often have incentives to obscure their decisional processes; survey respondents are impressionistic and have no incentives to be accurate. In contrast, attorneys operating in their professional capacity—*i.e.*, making tactical legal decisions—have huge incentives to make accurate choices. They are paid largely for their ability to ascertain the practical effect of legal rules and to predict future court behavior, and, most importantly, they want to win their case. For instance, a recent study on Supreme Court decision forecasting showed that appellate attorneys had a 92 % accuracy rate, whereas academics were right only 53 % of the time.[77]

3. Amount in Controversy Requirement

Finally, we should note that the removal metric is limited by the "amount in controversy" requirement.[78] As is well known, the requirements for diversity jurisdiction (and hence removal) not only require that the parties be from different states, but also that the amount in controversy be greater than a certain value. This requirement necessarily limits the scope of our inquiry to cases in which the damages claimed are greater than the statutory requirement (currently $75,000).[79] While one would obviously prefer not to be so limited, the limitation should not affect the validity of the results. Due to the high price of experts, scientific evidence battles generally surface in cases involving high damage claims.[80] The focus of the *Daubert* decision, products liability cases, also typically have high damage claims. And finally, to the extent that *Daubert* is viewed as an element of tort reform, it is those high-claim cases that are of particular interest to policymakers, practitioners, and scholars.

C. Future Areas of Research

Looking forward, two areas of future study are particularly noteworthy:

1. Removal Rate Trends

The removal rate trends seen in Figure 8 may have broader significance than as a graphical example of the non-effect of Indiana's change to a *Daubert* standard. Although Figure 8 is rather noisy, one might speculatively tell a story about the history of the treatment of scientific evidence in the federal courts. Prior to the *Daubert* decision in 1993, each jurisdiction had some average removal rate based on considerations such as the perceived quality of the state judiciary, concerns about out-of-state bias, and so forth. After federal courts switched to *Daubert*, removal rates plummeted for one of three reasons. First, the uncertainty of the new standard encouraged many defendants to remain in state court, where at least the results were more predictable (the "devil you know" phenomenon). Second, much of the early commentary on the *Daubert* decision hailed it as a defeat for defendant corporations and a victory for plaintiffs,[81] which may have deterred some defendants from litigating in federal court. Third, and relatedly, plaintiffs in diver-

sity cases may have increasingly filed in federal court to start, eliminating the need for defendants to remove.

Around 1995, as practitioners gained experience in both state and federal courts, however, the removal calculus began to change. *Daubert*, at least as practiced in the federal courts, turned out to be a defendant-friendly decision, and so those initial disincentives for defendants to litigate in federal court gradually vanished. Why did the removal rates not ultimately end up at levels higher than the baseline pre-1993? One guess is that by 1997 or 1998, the softer effects of *Daubert* emphasized throughout this study had already taken hold in state courts, negating any doctrinal advantage of switching from a *Frye*-governed state court to a *Daubert*-governed federal one.

A future study could explore this hypothesis further. For example, if the story is indeed accurate, then the rate of original diversity filings should increase from 1993 to 1997, corresponding to the general decline in removal rates during the same period. Research on original filings could offer some verification of the tentative hypothesis.

2. Joiner

Some recent scholarship argues that the crucial decision in the *Daubert* trilogy is not the *Daubert* decision itself, but the Supreme Court's decision in *General Electric Company v. Joiner*, which established an abuse-of-discretion standard for reviewing of scientific admissibility determinations.[82] A future study could investigate whether state adoption of a *Joiner*-type standard—and not the difference between *Frye* and *Daubert*—is positively correlated with harsher scrutiny of scientific evidence.

VI. Conclusion

The overarching goal of this Essay was to determine whether formal, doctrinal standards have any effect on scientific admissibility determinations. Nearly every discussion of scientific evidence begins with a treatment of the differences between the *Frye* and *Daubert* standards. This Essay asked candidly whether a state's adoption for *Frye* or *Daubert* has any practical impact.

Using both a preliminary study of Connecticut and the EDNY, as well as a national study of all available and relevant states, we found no evidence that *Frye* or *Daubert* makes a difference. In the preliminary study, removal rates in EDNY and Connecticut remained relatively stable from 1994 to 2000, despite Connecticut's change from a *Frye* to *Daubert* standard in 1997. In the national study, the econometric model established that the governing scientific admissibility standard was not a significant factor in determining removal rates after appropriately controlling for year-to-year and state-to-state variations.

The results of this study have both immediate and broader ramifications. For the scientific evidence field, the results suggest that debates about the practical merits and drawbacks of *Daubert* versus *Frye* may be largely superfluous, and that that energy should be refocused. In addition, our findings lend support to those scholars advocating for the uniform adoption of *Daubert* by the states. Perhaps it is time to move away from debating the merits of *Frye* versus *Daubert* and toward a broader focus on how judges actually make decisions about science.

More broadly, this study has made the first steps in developing removal as a method for measuring the effect of changes in procedural or evidentiary rules. The results suggest that doc-

trinal reforms do not always directly correlate with substantive changes in practice. Sometimes the power of a court decision or even a piece of legislation comes more from its underlying idea than from its technical legal effect.

Appendix A

Removal Rates in States Used in National Study

Year	Alaska	Arizona	Arkansas	Connecticut	Florida
1990	1.69 %	0.30 %	1.27 %	0.16 %	0.58 %
1991	2.39 %	0.36 %	1.12 %	0.23 %	0.73 %
1992	5.15 %	0.51 %	1.35 %	0.29 %	0.61 %
1993	1.39 %	0.61 %	1.42 %	0.24 %	0.54 %
1994	1.49 %	0.26 %	1.28 %	0.36 %	0.43 %
1995	1.17 %	0.25 %	0.57 %	0.36 %	0.31 %
1996	1.79 %	0.31 %	1.45 %	0.25 %	0.46 %
1997	2.39 %	0.40 %	1.35 %	0.25 %	0.76 %
1998	2.83 %	0.34 %	1.87 %	0.31 %	0.58 %
1999	1.62 %	0.31 %	2.09 %	0.28 %	
2000	1.94 %	0.48 %	1.75 %	0.30 %	

Removal Rates (A–F)

Year	Indiana	Kansas	Michigan	Minnesota	Missouri	NewMexico
1990	0.86 %	0.52 %	0.63 %	1.42 %	0.90 %	
1991	0.86 %	0.76 %	0.79 %	1.97 %	0.88 %	
1992	1.45 %	1.34 %	0.90 %	0.98 %	1.20 %	
1993	1.13 %	0.89 %	0.64 %	0.90 %	1.08 %	0.35 %
1994	0.51 %	0.65 %	0.49 %	0.86 %	1.10 %	0.93 %
1995	0.55 %	0.43 %	0.49 %	0.48 %	0.66 %	0.37 %
1996	0.91 %	0.66 %	0.71 %	0.80 %	0.78 %	0.81 %
1997	1.17 %	0.79 %	0.97 %	1.87 %	1.24 %	1.03 %
1998	0.85 %	0.79 %	0.98 %	0.99 %	1.09 %	1.01 %
1999	1.13 %	0.76 %	0.96 %	0.92 %	1.14 %	0.99 %
2000	1.01 %	0.86 %	0.98 %	1.12 %	1.11 %	1.15 %

Removal Rates (I–N)

Year	New York	North Carolina	Oregon	Tennessee	Washington
1990	0.46 %	0.31 %		0.97 %	0.61 %
1991	0.56 %	0.49 %	0.88 %	0.84 %	0.24 %
1992	0.55 %	0.47 %	0.97 %	1.15 %	0.43 %
1993	0.61 %	0.56 %	1.08 %	1.24 %	0.38 %
1994	0.71 %	0.42 %	1.46 %	1.25 %	0.28 %
1995	0.57 %	0.37 %	0.61 %	0.76 %	0.16 %
1996	0.83 %	0.44 %	0.48 %	1.03 %	0.38 %
1997	0.80 %	0.52 %	0.86 %	1.39 %	0.23 %
1998	0.80 %	0.62 %	0.41 %	1.28 %	0.25 %
1999	0.85 %	0.64 %	0.56 %	1.36 %	0.32 %
2000	1.01 %	0.77 %	0.74 %	1.67 %	0.48 %

Removal Rates (N–W)

Footnotes

[1] 509 U.S. 579 (1993).

[2] See D.H. Kaye, Choice and Boundary Problems in *Logerquist*, *Hummert*, and *Kumho Tire*, 33 Ariz. St. L.J. 41, 42 (2001) ("Much has been written about the merits, pedigree, and operation of these standards. Each has its strengths and weaknesses, its friends and foes.").

[3] Erica Beecher-Monas, Blinded by Science: How Judges Avoid the Science in Scientific Evidence, 71 Temp. L. Rev. 55, 75–76 (1998) (describing the two sides of the debate but arguing that the issue of whether the *Daubert* standard is more strict than the *Frye* standard is a "red herring").

[4] See, *e.g.*, Logerquist v. McVey, 1 P.3d 113, 125–29 (Ariz. 2000); People v. Leahy, 882 P.2d 321, 327–31 (Cal. 1994); State v. Porter, 698 A.2d 739, 749–52 (Conn. 1997). Numerous articles have also peppered the literature advocating for a particular state to adopt *Daubert* or maintain *Frye*. See, *e.g.*, Mary Gaston, Note, *State v. Gentry*: The Washington Supreme Court Opens the Door for Unreliable Scientific Evidence, 31 Gonz. L. Rev. 475, 498–99 (1995–96) (proposing either modifications to *Frye* or an adoption of *Daubert*); Penelope Harley, Comment, Minnesota Decides: *Goeb v. Tharalson* and the Admissibility of Novel Scientific Evidence, 24 Hamline L. Rev. 460, 463 (2001) (summarizing argument that Minnesota should have switched from *Frye* to *Daubert*); Andrew R. Stolfi, Note, Why Illinois Should Abandon *Frye's* General Acceptance Standard for the Admission of Novel Scientific Evidence, 78 Chi.-Kent L. Rev. 861, 862–63 (2003).

[5] Compare Jeffry D. Cutler, Implications of Strict Scrutiny of Scientific Evidence: Does *Daubert* Deal a Death Blow to Toxic Tort Plaintiffs?, 10 J. Envtl. L. & Litig. 189, 191–92 (1995) (suggesting a "grim overall outlook for toxic tort plaintiffs which could result from strict interpretation of *Daubert*"), with Paul M. Barrett, Justices Rule Against Business in Evidence Case—Restrictive Standard for Use of Scientific Testimony in Trials Is Struck Down, Wall St. J., June 29, 1993, at A3 (characterizing *Daubert* as a pro-plaintiff decision). Some of the commentary immediately following the *Daubert* decision expressed skepticism that the *Daubert* rule would change outcomes very much. See, *e.g.*, Richard D. Friedman, The Death and Transfiguration of *Frye*, 34 Jurimetrics J. 133, 143 (1994) (arguing that little would change under the new standard); Barbara Frederick, Note, *Daubert v. Merrell Dow Pharmaceuticals, Inc.*: Method or Madness?, 27 Conn. L. Rev. 237, 270 (1994) (predicting early on that "while *Daubert* will change the language of admissibility decisions, it will have little impact on their outcome").

[6] Joseph Sanders, Shari S. Diamond & Neil Vidmar, Legal Perceptions of Science and Expert Knowledge, 8 Psychol., Pub. Pol'y & L. 139, 141 n.13 (2002) (noting that early on, both plaintiffs and defendants attempted to spin *Daubert* in their direction, but that ultimately "in practice the *Daubert* test has been more restrictive than *Frye*").

[7] Lloyd Dixon & Brian Gill, Changes in the Standards for Admitting Expert Evidence in Federal Civil Cases Since the *Daubert* Decision xv (2001) (reporting that after *Daubert*, "[federal] judges scrutinized reliability more carefully and applied stricter standards in deciding whether to admit expert evidence"); 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, 330–31 (2002) (reporting results from judge and attorney surveys that suggest greater scrutiny of scientific evidence in the wake of *Daubert*).

[8] *Daubert-Frye* surveys have become rather popular contributions to the scholarly literature. See, *e.g.*, David E. Bernstein & Jeffrey D. Jackson, The *Daubert* Trilogy in the States, 44 Jurimetrics J. 351 (2004); Clifton T. Hutchinson, *Daubert* in State Courts, 9 Kan. J.L. & Pub. Pol'y 15 (1999); Manuel L. Real, *Daubert*—A Judge's View—A Reprise *in* ALI-ABA Course of Study Materials: Civil Practice and Litigation Techniques in Federal and State Courts, 411, 450 (2004). This map is based on those summaries as well as independent research and verification. The diagram is necessarily somewhat of a simplification. First, some states, such as Maine, have adopted *Daubert* in all but name. The map classifies these states as "*Daubert*" since they are doctrinally very similar. Second, some states have adopted *Daubert* but not subsequent, related Supreme Court decisions, such as *Kumho Tire Co. v. Carmichael*, 526 U.S.

137 (1999), and *General Electric Co. v. Joiner*, 522 U.S. 136 (1997); the map does not capture these nuances. For a breakdown among the states along these more complex lines, see Bernstein & Jackson, *supra*, at 357–61. Nevertheless, Figure 1 hopefully provides a convenient snapshot of the geographic distribution of the different standards.

[9] 4 David Faigman *et al.*, Modern Scientific Evidence §35-1.3, at 150–51 (2d ed. 2002); see also, *e.g.*, Paul C. Giannelli, Admissibility of Scientific Evidence, 28 Okla. City U. L. Rev. 1, 11 (2003) (noting that the language of *Daubert* has "crept into the *Frye* lexicon").

[10] See David E. Bernstein, *Frye*, *Frye*, Again: The Past, Present, and Future of the General Acceptance Test, 41 Jurimetrics J. 385, 388, 404 (2001) (observing that the "case law under *Frye* is slowly converging with *Daubert* jurisprudence"). The term "junk science" is commonly attributed to Peter Huber. Peter W. Huber, Galileo's Revenge: Junk Science in the Courtroom 2 (1991).

[11] 293 F. 1013, 1014 (D.C. Cir. 1923).

[12] See Bernstein, *supra* note 10, at 388–89 (noting that although there were few citations to *Frye* through the 1960's, it remained influential nonetheless).

[13] *Frye*, 293 F. at 1014.

[14] Id.

[15] For example, Judge Harvey Brown wrote

> [T]he *Frye* test was criticized because the newness of a scientific theory does not necessarily reflect its unreliability, "nose counting" of the scientific community could be difficult and unhelpful, and the standard delays the admissibility of new evidence simply because the scientific community has not had adequate time to accept the new theory.

Harvey Brown, Eight Gates for Expert Witnesses, 36 Hous. L. Rev. 743, 779 (1999).

[16] 509 U.S. 579 (1993). This issue of reliability had, in a sense, been percolating since 1975, when Congress codified the Federal Rules of Evidence. In particular, Congress included Federal Rule of Evidence 702 to address the admissibility of expert testimony: "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." Fed. R. Evid. 702. Neither the rule nor the commentary notes refer to Frye, emphasizing relevance rather than reliability or "general acceptance." Not surprisingly, many federal courts continued to follow Frye in evaluating scientific evidence.

[17] The Court specifically held that Frye was superseded by Rule 702 (notwithstanding the fact that the rule made no mention of Frye or "general acceptance"). See Daubert, 509 U.S. at 587.

[18] Id. at 593–94.

[19] Id. at 592–93.

[20] 522 U.S. 136, 142 (1997).

[21] 526 U.S. 137, 141 (1999).

[22] Dixon & Gill, *supra* note 7, at 41. Dixon and Gill analyzed 399 district court opinions from January 1980 to June 1999. Id. at 15–18.

[23] Sophia I. Gatowski *et al.*, Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-*Daubert* World, 25 Law & Hum. Behav. 433, 444–48 & tbl.1, 452–53 (2001).

[24] Jennifer L. Groscup *et al.*, The Effects of *Daubert* on the Admissibility of Expert Testimony in State and Federal Criminal Cases, 8 Psychol. Pub. Pol'y & L. 339, 342, 344 (2002).

[25] Id. at 345, 363. Professor Groscup also observed that while *Daubert* courts discussed reliability issues at greater length, discussion about the three new *Daubert* factors was scant, a result consistent with the Dixon and Gatowski studies. Id. at 365 (concluding that while judges understood the import of the *Daubert* decision and cited to it accordingly, they did not apply the criteria in any meaningful way).

[26] Pamela J. Jensen, Note, *Frye* Versus *Daubert*: Practically the Same?, 87 Minn. L. Rev. 1579, 1581 (2003). Jensen considered all relevant state appellate decisions on three forms of scientific evidence used in criminal cases (32 in total) and found no support for "the idea that *Frye* and *Daubert* admissibility standards lead to distinct practical outcomes." Id. at 1611–12 & tbl.1; see also id. at 1619 (commenting that "[a]lthough states vary widely in how they treat certain types of scientific evidence, this variation does not correlate with the adherence to *Frye* or *Daubert* admissibility standards").

[27] Groscup, *supra* note 24, at 344 (limiting study to criminal cases only); Jensen, *supra* note 26, at 1585–90 (describing the types of expert evidence studied, which are primarily found in criminal cases).

[28] See Bernstein, *supra* note 10, at 389 (recognizing that "most state court opinions, particularly at the trial court level, are unpublished").

[29] See Krafka, *supra* note 7, at 331 ("To determine how *Daubert* and its associated cases have affected judicial and attorney practices in the majority of cases that never go to trial, further research is needed.").

[30] See, *e.g.*, Theodore Eisenberg *et al.*, Litigation Outcomes in State and Federal Courts: A Statistical Portrait, 19 Seattle U. L. Rev. 433, 442, 443 tbl.4 (1996) (reporting an overall state trial rate of 2.9 % for all torts for 1991–1992).

[31] See Edward K. Cheng, Changing Scientific Evidence, 88 Minn. L. Rev. 315, 333–35 (2003) (discussing the scientific life cycle).

[32] We are aware of only one previous study that has used removal rates to measure the effect of a legal change—a 2002 Federal Judicial Center study of the effect of two Supreme Court decisions on federal class actions. Bob Niemic & Tom Willging, Federal Judicial Center, Effects of *Amchem*/*Ortiz* on the Filing of Federal Class Actions: Report to the Advisory Committee on Civil Rules, Sept. 9, 2002, at 12–13. This study, however, looked only at the number of removals, not the removal rate as defined below. See *infra* note 44 and accompanying text. Additionally, the Niemic study used different data sources than we used. Niemic & Willging, *supra*, at 4–5.

[33] 13B Charles Alan Wright *et al.*, Federal Practice & Procedure §3602 (2d ed. 1984 & Supp. 2004) (noting that a plaintiff may seek federal diversity jurisdiction even when litigating in his or her home state). But see 28 U.S.C. §1441(b) (2000) (barring removal if any of the defendants is a citizen of the state in which the action was brought).

[34] See, *e.g.*, Neal Miller, An Empirical Study of Forum Choices in Removal Cases Under Diversity and Federal Question Jurisdiction, 41 Am. U. L. Rev. 369, 400–23 (1991) (describing the various factors that attorneys consider in making forum choices).

[35] 28 U.S.C. §1446(b) (2000); see also 14C Wright *et al.*, *supra* note 33, at §3732.

[36] See Samuel R. Gross, Expert Evidence, 1991 Wisc. L. Rev. 1113, 1118–19 (reporting that in a sample of California civil trials from 1985 to 1986, 86 % involved expert testimony). Professor Gross further found that experts were involved in 97 % of medical malpractice trials (at an average of five experts per trial), and in 100 % of products liability trials. Id. Naturally, these rates may suffer from selection bias because they only describe cases that went to trial, but they nonetheless support the general proposition that the use of expert testimony is widespread in tort litigation.

[37] 304 U.S. 64, 78 (1938).

[38] Scholars such as Judge Richard Posner also argue that there is a qualitative difference between state and federal judges due to selection effects and institutional incentives. See Richard Posner, The Federal Courts 142–45 (1985) (discussing factors lawyers consider in choosing between federal and state court).

[39] 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 & Contemp. Probs. 289, 290 (2001).

[40] To the extent that litigants must trade off aspects of litigating in federal court, strength of evidence may affect removal decisions. For example, a state-court-inclined defendant may be more inclined to remove to federal court if the admissibility of plaintiff's scientific evidence is debatable. In those instances, the

criticality of having the stricter *Daubert* standard in federal court may outweigh any preferences the defendant would normally have for state court.

[41] See People v. Wernick, 674 N.E.2d 322, 324 (N.Y. 1996); Marsh v. Smyth, 785 N.Y.S.2d 440, 441 (N.Y. App. Div. 2004) (applying *Frye* standard to medical expert's testimony).

[42] See State v. Porter, 698 A.2d 739, 746 (Conn. 1997).

[43] Econometrically, this comparison sets up a "difference-in-differences" approach. See *infra* notes 55–57 and accompanying text.

[44] This definition of "removal rate" differs from the one used in the Niemic & Willging study, which is the only previous study to have used removal rates to investigate legal changes. See Niemic & Willging, *supra* note 32, at 12. Niemic & Willging measured removal rates by comparing the ratio of cases originally filed in federal court with cases removed from state court. Id.

[45] Federal Judicial Center, Federal Court Cases: Integrated Data Base, 1970–2000, ICPSR Study No. 8429 (2001), at http://webapp.icpsr.umich.edu/cocoon/ICPSR-STUDY/08429.xml (on file with the Virginia Law Review Association).

[46] The change in Connecticut's removal rate between the pre- and post-periods was assessed using a difference-in-difference model. For an explanation of the difference-in-difference approach, see *infra* notes 55–57 and accompanying text.

[47] One can argue that a change in scientific admissibility standard is an exogenous determination that is minimally correlated with short-term changes in demographic or other unobserved variables. The length of judicial tenure and the independence of the judiciary make judicial decisions by and large independent of any short-term demographic or political change. Furthermore, given that the precise admissibility standard for scientific evidence is a specialized evidentiary issue unlikely to attract significant public attention, one would expect it to play a sharply limited role, if any, in the judicial appointment process.

[48] See Federal Judicial Center, *supra* note 45.

[49] National Center for State Courts, State Court Statistics, 1985–2001, ICPSR Study No. 9266 (2002), at http://webapp.icpsr.umich.edu/cocoon/ICPSR-STUDY/09266.xml (on file with the Virginia Law Review Association).

[50] Arkansas followed a multifactor, *Daubert*-like test from 1990 to 2000, and then explicitly adopted *Daubert* in 2000. See Farm Bureau Mut. Ins. Co. v. Foote, 14 S.W.3d 512, 519 (Ark. 2000). Arkansas was considered a "*Daubert*" state for purposes of this study.

[51] Indiana followed *Frye* until 1994, had a somewhat unclear standard from 1994 to 1995, and then began relying on *Daubert* to guide evidentiary expert rulings beginning in 1995. See Steward v. State, 652 N.E.2d 490, 498 (Ind. 1995) ("[A]lthough not binding upon the determination of state evidentiary law issues, the federal evidence law of *Daubert* and its progeny is helpful to the bench and bar in applying Indiana Rule of Evidence 702(b)."). For purposes of this study, we considered Indiana to be a *Daubert* state.

[52] From 1984 to 1995, North Carolina refused to adopt *Frye* but adopted its principles. See State v. Peoples, 319 S.E.2d 177, 187 (N.C. 1984). In 1995, the North Carolina Supreme Court expanded its standard to look at factors beyond general acceptance, in essence adopting a *Daubert* standard. See State v. Goode, 461 S.E.2d 631, 639 (N.C. 1995); see also Taylor v. Abernethy, 560 S.E.2d 233, 273 (N.C. Ct. App. 2002) (noting the North Carolina Supreme Court's citation of *Daubert*); Bernstein & Jackson, *supra* note 8, at 358 & n.41 (classifying North Carolina as a *Daubert* state). The North Carolina Supreme Court later held its standard to be distinct from *Daubert* in 2004, see Howerton v. Arai Helmet, Ltd., 597 S.E.2d. 674, 689 (N.C. 2004), but from 1995 to 2000, appellate courts often thought otherwise. See, *e.g.*, *Taylor*, 560 S.E.2d at 273; State v. Bates, 538 S.E.2d 597, 600 (N.C. Ct. App. 2000) (interpreting *Goode* as adopting *Daubert*).

[53] Oregon adopted a multi-factor standard of admissibility similar to *Daubert* in 1984. See State v. Brown, 687 P.2d 751, 759 (Or. 1984).

[54] At first glance, Missouri's standard during the study period is somewhat ambiguous. Prior to 1993, the Missouri Supreme Court clearly followed the *Frye* standard. See State v. Davis, 814 S.W.2d 593, 600 (Mo. 1991) (stating that *Frye* had "been adopted and regularly applied in a variety of Missouri decisions"). After the *Daubert* decision in 1993, confusion initially arose among Missouri courts because Missouri's statutory rule governing expert evidence, §490.065 of the Missouri Revised Statutes, was modeled after Federal Rule of Evidence 702. See Long v. Mo. Delta Med. Ctr., 33 S.W.3d 629, 642 (Mo. Ct. App. 2000). A consensus soon formed among the appellate courts, however, that *Frye* should continue to be applied. See id.; M.C. v. Yeargin, 11 S.W.3d 604, 619 (Mo. Ct. App. 1999) ("The Missouri Supreme Court continues to apply the *Frye* test of the admissibility of expert testimony in criminal cases and in civil cases."). But *cf.* Lasky v. Union Elec. Co., 936 S.W.2d 797, 802 (Mo. 1997) (remanding case for an admissibility determination under the statutory language itself). In light of this apparent consensus, we have classified Missouri as a *Frye* state for the study. See also Bernstein & Jackson, *supra* note 8, at 355 n.25 (defining Missouri as a *Frye* state). The Missouri Supreme Court subsequently adopted a more *Daubert*-friendly approach in 2003. See State Bd. of Registration for Healing Arts v. McDonagh, 123 S.W.3d 146, 155 (Mo. 2003).

[55] See Orley Ashenfelter, Estimating the Effects of Training Programs on Earnings, 84 Rev. Econ. & Stat. 47 (1978); David Card, The Impact of the Mariel Boatlift on the Miami Labor Market, 43 Indus. & Lab. Rel. Rev. 245 (1990); David Card & Alan B. Kreuger, Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania, 84 Amer. Econ. Rev. 772 (1994); Jonathan Gruber, The Incidence of Mandated Maternity Benefits, 84 Amer. Econ. Rev. 622 (1994); Albert Yoon, Damage Caps and Civil Litigation: An Empirical Study of Medical Malpractice Litigation in the South, 3 Amer. L. & Econ. Rev. 199 (2001).

[56] Given the econometric model, all year coefficients are in relation to 1994. The year 1994, which was chosen arbitrarily, is accounted for when all the state variables are zero.

[57] Just as with the year variables, all state coefficients are in relation to Alaska. Alaska, which was chosen arbitrarily, is accounted for when all the state variables are zero.

[58] In reading the regression results, variable names are in the leftmost column. The correlation of each variable on the removal rate is measured by the coefficient in the second column. Thus, for example, in Figure 9, the fact that the year is 1995 tends to lower removal rates, whereas the fact that the year is 1997 tends to raise removal rates. A coefficient of zero, or relatively close to zero, suggests that a variable either has no effect on removal rates (in the case of a single variable like DAUBERT), or no effect relative to the omitted baseline dummy variable. See *supra* notes 57–58.

> Statistically, the confidence we have that any coefficient is different from zero is found in the P>|t| (or p-value) column. A small p-value means that there is a very small probability that the coefficient is actually zero (but calculated to be non-zero because of random variations). "Statistical significance" is generally set at the 5 % level, or p=0.05, which means that there is only a 5 % chance that the coefficient is actually zero.

[59] R^2 is a statistical measure of "fit"—that is, how well the model and its variables predict removal rate. Generally, the higher the R^2 the better the predictive force of the model.

[60] Compared to the other coefficients, the coefficient for TORTFILE may appear small, and therefore insignificant, but only deceptively so. Unlike the other variables, which have values of only zero or one, the TORTFILE variable generally has values in the thousands, which means that a small TORTFILE coefficient can still suggest a strong relationship between the number of cases filed and the number of cases removed.

[61] Again, in reading Figure 12, the number in the DAUBERT row is the coefficient describing the effect that adoption of the *Daubert* standard has on removal rate (or number of cases removed). The parenthetical number is the t-value, a statistical measure of confidence. N is the number of observations of removal rate considered in the analysis (one observation per state per year). R^2 is a statistical measure of "fit," or how well the model and its variables predict removal rate. Note that without the state and year

controls, R^2=0.03, suggesting that a model based purely on DAUBERT cannot predict removal rate at all.

[62] Of course, theoretical discussions about the differences between *Frye* and *Daubert* remain important, since those debates may ultimately lead to a greater understanding of how to assess science.

[63] Bernstein, *supra* note 10, at 404–07 (arguing that *Frye* jurisdictions should adopt *Daubert*, not necessarily because *Daubert* is a different or better rule, but merely to eliminate confusion).

[64] See State v. Crosby, 927 P.2d 638, 641–42 (Utah 1996) (suggesting that Utah's test is similar to *Daubert*, but is more rigid).

[65] See New State Ice Co. v. Liebmann, 285 U.S. 262, 311 (1932) (Brandeis, J., dissenting) ("It is one of the happy incidents of the federal system that a single courageous State may, if its citizens choose, serve as a laboratory; and try novel social and economic experiments without risk to the rest of the country.").

[66] See Gatowski *et al.*, *supra* note 23, at 454–55 (advocating for "more science-based judicial education").

[67] Federal Judicial Center, Reference Manual on Scientific Evidence (2d ed. 2000).

[68] See General Electric Co. v. Joiner, 522 U.S. 136, 142 (1997) (holding that abuse of discretion is the proper standard of review for district court's scientific evidentiary rulings); Bernstein & Jackson, *supra* note 8, at 352–66 (summarizing state adoption or rejection of *Joiner*).

[69] Albert Yoon, Mandatory Arbitration and Civil Litigation: An Empirical Study of Medical Malpractice Litigation in the West, 6 Am. L. & Econ. Rev. 95, 118–27 (2004) (finding that the implementation of mandatory arbitration in Nevada did not lead to a statistically significant effect on how much plaintiffs recovered or the duration of their litigation, but did have a small but statistically significant downward effect on the probability that the judicial system would resolve the dispute).

[70] Niemic & Willging, *supra* note 32, 1–2, 23–24 (studying the effect of *Amchem Products, Inc. v. Windsor*, 521 U.S. 591 (1997) and *Ortiz v. Fibreboard Corp.*, 527 U.S. 815 (1999)).

[71] Note that this rate is quite different from the removal rate used in the study. The study defined "removal rate" as the percentage of all tort cases filed in state court that are ultimately removed to federal court. This figure is far lower than the rate at which removable cases are removed, because many of the state court cases lack diversity of citizenship.

[72] Another concern about using a removal metric is the apparent recent increase in the abuse of removal motions to increase costs and delays. See Theodore Eisenberg & Trevor Morrison, Forum Manipulation by Defendants: The Growth of Wrongful Removal to Federal Court (2004) (unpublished manuscript) (on file with author). This phenomenon, however, should not pose any problems because our study includes *Frye* states as controls, and the model controls for year-to-year variations.

[73] Calculating the removal rate for removable cases would require information on the citizenship of the parties involved in the litigation. Unfortunately, this information is typically not available in most state judicial databases.

[74] Jolanta Juszkiewicz Perlstein, Lawyers' Strategies and Diversity Jurisdiction, 3 Law & Pol'y Q. 321, 328 tbl.1 (1981) (showing baseline forum preferences of out-of-state defense attorneys as a "control group").

[75] Miller, *supra* note 34, at 400–23 (describing the various factors that attorneys consider in making forum choices); *cf.* Jerry Goldman & Kenneth S. Marks, Diversity Jurisdiction and Local Bias: A Preliminary Empirical Inquiry, 9 J. Legal Stud. 93, 100 tbl.4 (1980) (ranking various reasons why attorneys in the Chicago area prefer filing in federal court over state court).

[76] This second inference, however, has a weakness arising from the earlier insight that toxic tort litigation follows trends in which certain defendants or industries are targeted in a given year. Thus, even if the rate of removal for removable cases was 100 %, there could be significant fluctuations in observed removal rates on the basis of the number of suits filed against the targeted defendant and whether removal was available to that defendant.

[77] See Theodore W. Ruger *et al.*, The Supreme Court Forecasting Project: Legal and Political Science Approaches to Predicting Supreme Court Decisionmaking, 104 Colum. L. Rev. 1150, 1177–79, 1178

tbl.5 (2004) (cautioning, however, against reading too much into the results given the small sample size and potential selection effects).

[78] 28 U.S.C. §1332 (2000).

[79] Id.

[80] See Samuel R. Gross & Kent D. Syverud, Getting to No: A Study of Settlement Negotiations and the Selection of Cases for Trial, 90 Mich. L. Rev. 319, 354 tbl.5 (1991) (showing trends suggesting a correlation between higher settlement offers, higher trial awards, and the greater use of experts).

[81] See, *e.g.*, Paul M. Barrett, Justices Rule Against Business in Evidence Case—Restrictive Standard for Use of Scientific Testimony in Trials Is Struck Down, Wall St. J., June 29, 1993, at A3 (characterizing *Daubert* as a pro-plaintiff decision).

[82] 522 U.S. 136, 138–39 (1997).