University of Michigan Law School

University of Michigan Law School Scholarship Repository

Book Chapters

Faculty Scholarship

1996

Substance and Form in Scientific Evidence: What Daubert Didn't Do

Samuel R. Gross University of Michigan Law School

Available at: https://repository.law.umich.edu/book_chapters/286

Follow this and additional works at: https://repository.law.umich.edu/book_chapters



Part of the Civil Procedure Commons, and the Evidence Commons

Publication Information & Recommended Citation

Gross, Samuel R. "Substance and Form in Scientific Evidence: What Daubert Didn't Do." In Reforming the Civil Justice System, edited by L. Kramer, 234-79. New York: New York Univ. Press, 1996.

This Book Chapter is brought to you for free and open access by the Faculty Scholarship at University of Michigan Law School Scholarship Repository. It has been accepted for inclusion in Book Chapters by an authorized administrator of University of Michigan Law School Scholarship Repository. For more information, please contact mlaw.repository@umich.edu.

Reforming the Civil Justice System

Edited by Larry Kramer



Substance and Form in Scientific Evidence: What Daubert Didn't Do

Samuel R. Gross*

On its face, Daubert v. Merrell Dow Pharmaceuticals¹ was about as easy a case as the Supreme Court gets. The plaintiffs claimed that their birth defects were caused by the anti-nausea drug Bendectin, which their mothers had used during their gestation. In response to a motion for summary judgment by the defendant, the plaintiffs presented affidavits of eight expert witnesses who offered their opinions—based on a variety of studies—that Bendectin was indeed the culprit. The federal district court that heard the motion granted summary judgment to the defendant, and the Ninth Circuit affirmed. Both lower courts held that critical portions of the plaintiffs' evidence were inadmissible, and that without that evidence the plaintiffs had not met their burden of producing sufficient evidence to present a genuine factual dispute. The first holding—that this expert evidence was inadmissible—was the sole issue in the Supreme Court.

The question is governed by Rule 702 of the Federal Rules of Evidence:

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.

^{*}Professor, University of Michigan School of Law. This chapter has benefited from advice and comments by David Garabrandt, Heidi Feldman, Richard Friedman, Robert Harris and Richard Lempert, and from excellent research assistance from Nancy Laetham, Kristina Maritezae and Nancy Vettorello. The research was supported by the Cook Funds of the University of Michigan School of Law.

Was this a case in which scientific knowledge would "assist the trier of fact"? Certainly. Were the plaintiffs' witnesses qualified by "knowledge, skill, experience, training, or education" to testify as experts? That was uncontested. Were there any other requirements that these witnesses had to satisfy? Arguably—plausibly—the term "knowledge," scientific knowledge, implies that the experts' testimony must be based on scientific information—which it was. Arguably, also, their evidence was subject to the requirement in Rule 703 of the Federal Rules of Evidence that the information on which an expert opinion is based, if it is not itself admissible in evidence, must be "of a type reasonably relied upon by experts in the particular field in forming opinions or inferences upon the subject." Would a reasonable scientist who was assessing the possibility that a drug produces birth defects rely on the sort of information these experts used? That too was not in dispute; in fact, the defendant's experts used much of the same data themselves.

Why then did two lower courts erroneously exclude this evidence? Why was Supreme Court review necessary? Why did this case attract 22 briefs *amicus curiae* and an unusual amount of publicity, even by Supreme Court standards?

For the most part, the defendant in Daubert, the lower courts, and the amici curiae, were all concerned with issues other than whether this sort of evidence is admissible under the Federal Rules of Evidence. There were several additional questions hovering around the case, of greater or lesser importance and generality, and much of the debate and the anxiety were directed at these other questions: (1) Does Bendectin in fact cause birth defects in human beings? The scientific consensus on this question is that Bendectin probably does not cause birth defects, and that certainly there is no substantial evidence that it does. (2) Do juries answer such questions correctly? Many people contend that juries sometimes, perhaps frequently, make terrible mistakes in dealing with scientific issues in general, and with causation in toxic torts cases in particular. Juries have obviously been wrong on Bendectin at least several times, if for no other reason because they have reached inconsistent verdicts in different cases. (3) Is the danger of such errors - and the resulting risk of legal liability-wreaking havoc in the American pharmaceutical industry or in American manufacturing generally? Some people contend that product liability litigation based on bogus science is destroying our international competitiveness. (4) Can we do anything to help juries get scientific questions right? Perhaps some new rule of evidence or procedure would improve matters. (5) Should juries decide such issues at all? Does it make sense to entrust scientific questions to ad-hoc tribunals of lay people? And if not juries, who should do it? Judges? Medical or scientific experts? Administrative bodies?

These issues were not before the Supreme Court, but they might have been affected by its opinion. Supreme Court opinions often make law well beyond the specific questions at hand. In this case that seemed particularly likely, since the Court rarely addresses the problems of scientific evidence. It didn't happen.

The good news about *Daubert* is that the Supreme Court got it right. It understood the technical issue presented by the case, and it wrote about as useful an opinion as possible given the structure of the Federal Rules of Evidence. The bad news is that (as with most easy cases) the Court did little. Its decision is not likely to change practice much, and it will not have any major effect on the important side issues that generated most of the heat.

The problem with *Daubert* is that it deals with form rather than content, the admissibility rather than the sufficiency or accuracy of the scientific evidence. The underlying substantive question is simple: How strong is the evidence that Bendectin causes birth defects? On that factual issue as well *Daubert* was an easy case, but in the opposite (anti-plaintiff) direction. By the time the District Court issued its opinion there was a clear scientific consensus that Bendectin is safe, or in any event that there is no substantial evidence to the contrary.

The lower courts that have handled Bendectin cases have been influenced by the substance of the evidence, and their decisions have come to reflect the consensus among scientists. In recent years they have uniformly judged the evidence against Bendectin lacking, and have held for the defendant. But they have done so by subterfuge, claiming that the plaintiff's evidence is legally insufficient under the standard that applies to motions for summary judgment, or that it is inadmissible altogether. Neither rationale is consistent with the usual rules that apply to these procedural issues.

I have no quarrel with the courts' judgment in the Bendectin cases, but their methods are unfortunate. Instead of distorting procedural law they ought to say directly what they are really doing: ruling on the factual merits of a scientific claim. In an easy case—and Daubert certainly qualifies—this would amount to nothing more than an application of the well established doctrine of "legislative fact finding." On other scientific issues the evidence is often less one-sided, and the outcome less obvious. Even so, courts may do well to extend the practice of substantive judicial decision making to some harder scientific issues, provided they do it explicitly and with great care. In all cases, courts will do better by adding reliable scientific evidence from trustworthy sources than by excluding evidence they consider unreliable.

I. Daubert and the Bendectin Cases

Daubert was a late Bendectin cases. Bendectin is the brand name for an anti-nausea drug that was marketed by the Richardson-Merrell pharmaceutical company ("Merrell") from 1956 until 1983.² It was widely prescribed for morning sickness; in the 1960s and '70s Bendectin was taken by perhaps as many of 25% of all pregnant women in the United States.³ In January 1977 a lawsuit was filed on behalf of David Mekdeci, who was born with missing and malformed fingers and a missing pectoral muscle. The suit alleged that these birth defects were caused by the Bendectin that was prescribed to his mother while she was carrying him. This was the first Bendectin case; it ended in 1981, after a retrial, with a defense verdict that was upheld on appeal.⁴ Between 1980 and 1991 about 1,700 additional Bendectin cases were filed, some thirty trials were held, and several dozen decisions were published by state and federal courts.⁵ By December 1991, however, when the Ninth Circuit issued its Daubert opinion, this spate of litigation was all but completely over.

The Bendectin cases, as a group, constitute one of several recent examples of mass toxic tort litigation. As mass toxic tort litigation goes, Bendectin is a neat and inviting subject for study, because the entire lifespan of the issue covered a mere dozen years. Asbestos litigation, by contrast, has been going on for decades, with no end in sight. Three inter-related reasons made it possible to dispose of the Bendectin claims with such relative dispatch. First, the alleged harm has a short latency; it is apparent at the birth of the plaintiff, which is necessarily within nine months of the use of the drug. For other toxic substances - asbestos, for example, or the drug DES6-ill effects may not be apparent until ten or twenty or fifty years after exposure. Therefore it was comparatively easy to identify potential plaintiffs quickly. Second, because of this short latency, the claim that Bendectin causes birth defects was also comparatively easy to examine; the studies can be completed in months rather than years or decades. As a result, and also because of the intense legal interest in the issue, Bendectin was studied with uncommon intensity.7 Third, the clear conclusion from the large body of scientific evidence that had accumulated by the mid-1980's was that Bendectin does not cause birth defects.8

There is no need for an extensive history of the Bendectin litigation in this context. It is not important to the issues in *Daubert*, and a detailed and excellent description of the life cycle of this set of cases has been already published by Professor Joseph Sanders, on whose work this section is largely based.⁹ A synopsis of the highlights, however, will be useful.

The next Bendectin trial (after Mekdeci) was Oxendine v. Merrell Dow Pharmaceuticals, 10 which was tried in the Superior Court for the District of Columbia in May 1983. In Oxendine the jury awarded the plaintiff \$750,000, the trial judge entered a judgment notwithstanding the verdict in favor of the defendant, and the District of Columbia Court of Appeals reversed the j.n.o.v. and reinstated the jury's verdict. Oxendine remains the only Bendectin case in which a verdict against Merrell has been

sustained on appeal. In the meantime, starting in February 1982, all Bendectin cases that were pending in federal courts were consolidated for pretrial discovery before Judge Carl A. Rubin of the Southern District of Ohio (the home district of Merrell). Eventually most of these cases –818 out of 1,186—were consolidated before Judge Rubin for trial as well as discovery. Judge Rubin ordered a trifurcated joint trial, which began in February 1985. The first part was restricted to general causation: does Bendectin cause any birth defects in humans? A jury found that it does not, which ended the proceedings. In September 1985, Judge Rubin denied the plaintiffs' motions for j.n.o.v. or a new trial; three years later the judgment for the defendant was upheld on appeal.

There have been some 24 trials and retrials of Bendectin cases since the joint trial before Judge Rubin. 14 They have produced two hung juries, one mistrial, fourteen defense verdicts, and seven plaintiff's verdicts - of which five have been reversed and two are pending on appeal. In three of the reversals, a federal circuit court granted a j.n.o.v. to the defendant, or upheld a j.n.o.v. by the trial court, on the grounds that the plaintiff's expert evidence on the issue of general causation was either inadmissible or insufficient as a matter of law, 15 and in a fourth reversal a j.n.o.v. was remanded for reconsideration in light of Daubert. 16 By one route or another, several federal circuits in succession decided that on the existing state of the scientific evidence, Bendectin claims could not be maintained: the First Circuit and the D.C. Circuit in 1987. 17 the Fifth Circuit in 1989, 18 the Sixth Circuit in 1992. 19 As this pattern emerged, trial court judges became more willing to short-circuit the process by granting pre-trial motions for summary judgment on the same grounds. 20 Not surprisingly, trials became increasingly rare. There was one Bendectin trial in 1985 (in addition to the joint trial), seven in 1986, eleven in 1987, three in 1988, none in 1989, two in 1990, one in 1991 and none since.²¹

Initially, Daubert fit right into this pattern. Judge Earl Gilliam of the United States District Court for the Southern District of California granted a defense motion for summary judgment in November 1989.²² Like other late Bendectin opinions, Judge Gilliam's consists primarily of a review of the evidence on general causation, in the context of earlier opinions on the same issue dealing with much the same evidence.

Briefly: There are four types of scientific studies that are commonly used to determine whether a substance has toxic consequences:²³ analyses of the chemical structure of the compound that focus on similarities between it and known toxins; *in vitro* tests that examine its effects on human or animal cells that are cultivated in the laboratory for this purpose; *in vivo* studies that test its effects on laboratory animals; and *epidemiological* studies that examine its effects on human beings. There is general agreement that epidemiological studies are the

best and most informative, since only they provide direct evidence on the occurrence of pathologies in people. Unfortunately, epidemiological studies are also considerably slower and more expensive than in vitro or in vivo studies. Our advanced industrial economy produces thousands of potentially toxic substances; for many (perhaps most) there are no epidemiological studies at all. Bendectin is different. It has been heavily studied, for several reasons-because it's a regulated drug that required FDA approval before marketing, because the effect at issue is comparatively easy to test, and because of the interest generated by litigation. By 1985, thirty-seven epidemiological studies of the relationship between Bendectin and human birth defects had been published; in 1989 one more appeared.²⁴ In none of these studies did the researchers conclude that Bendectin is a teratogen, a substance that causes birth defects. In six the authors reported some relationship between the drug and one or more types of birth defects, but concluded that their evidence was insufficient to show causal connection. In the remaining thirty-two studies, they drew no conclusion about the teratogenic effects of Bendectin or affirmatively concluded that it has none.25

This state of the evidence was presented to Judge Gilliam in affidavits from expert witnesses submitted by Merrell in support of its motion for summary judgment. The plaintiffs responded with affidavits from eight experts of their own. The District Court held, in two stages, that the evidence offered by the plaintiffs was inadmissible. (1) Several experts concluded that Bendectin causes birth defects on the basis of their evaluations of its chemical structure, and on in vitro and in vivo studies. Judge Gilliam, following cases in other circuits, concluded that in the absence of "statistically significant epidemiological evidence" in support, such expert opinions were inadmissible because they "lack[] the sufficient foundation necessary under FRE 703"26in other words, because experts in their fields would not "reasonably rely" on such data to reach a conclusion such as this. (2) One of the plaintiffs experts, however, also based his opinion on a reanalysis of previously published epidemiological data. Despite the fact that the expert claimed to have found a statistically significant relationship between Bendectin and birth defects, Judge Gilliam held that this evidence was "insufficient to take this matter to the jury"27 because the claim was too vague, and because it was "never published or subjected to peer review."28 Judge Gilliam seemed to say that this final expert opinion, like the others, was inadmissible under FRE 703, but he could also be read to mean that although it was admissible, it was insufficient to raise a question for the jury to decide.29

II. Daubert and the Frye Rule

A. The Ninth Circuit Opinion

The District Court opinion in *Daubert* was in line with the trend in Bendectin cases. Judge Gilliam granted summary judgment to the defendant on narrow grounds. His opinion could conceivably mean that statistically significant epidemiological evidence is required in every toxic tort case. Or it could simply mean that the evidence offered by the plaintiff was inadmissible *in this case* because, given the wealth of negative epidemiological data on Bendectin, no expert could reasonably conclude that Bendectin is a teratogen on the basis of the limited data these experts relied on. Either way, the impact of the opinion, like that of other trial-court and appellate opinions in Bendectin cases, was limited to one claim, or to a single class of claims. It did nothing to restrict the use of expert evidence in general.

The Ninth Circuit opinion in *Daubert* is a sharp departure from this pattern.³⁰ The court (in an opinion by Judge Alex Kozinski) begins its legal analysis by restating and relying on a general evidentiary rule, the rule that the Circuit Court for the District of Columbia announced in 1923 in *United States v. Frye*:³¹ The issue in *Frye* was the admissibility of a "systolic blood pressure deception test," a primitive precursor of the modern polygraph. The District of Columbia Circuit laid down the standard:

Just when a scientific principle or discovery crosses the line between the experimental and the demonstrable is difficult to define. Somewhere in this twilight the evidential force of the principle must be recognized . . . [but] the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.³²

Although the *Frye* rule had been a controversial element of American evidence law on expert evidence for nearly seventy years, it played essentially no role in the Bendectin litigation up to that point. *Frye* was cited only once in the half-dozen circuit court Bendectin opinions preceding *Daubert*: in *DeLuca v. Merrell-Dow Pharmaceuticals, Inc.*, ³³ the Third Circuit noted that it had rejected the *Frye* test. This is hardly surprising. Whatever the cryptic "general acceptance" test means, its range was always understood to be restricted to the type of issue that was presented in *Frye* itself, the admissibility of *novel* scientific evidence. But there was nothing novel about the plaintiffs' expert evidence in *Daubert*, or any of the other Bendectin cases; on the contrary, both the data and the modes of analysis were tried and true, old hat. The problem was

a different one-simply that it seemed quite clear that the plaintiffs'

experts were wrong.

The Ninth Circuit begins its discussion of the evidence in *Daubert* by adopting a principle it derives from other circuit court Bendectin cases: "animal and chemical studies . . . [are] insufficient to establish a link between Bendectin and birth defects." That leaves the reanalysis of published epidemiological studies as the plaintiffs' last hope. For that item, the court invokes the *Frye* standard, so with a twist. As a mode of scientific investigation, reanalysis of published data is neither novel nor controversial; but that is not enough:

[T]he reanalysis of epidemiological studies is generally accepted by the scientific community only when it is subjected to verification and scrutiny by others in the field. . . . Plaintiffs' reanalyses do not comply with this standard: they were unpublished, not subjected to the normal peer review process, and generated solely for use in litigation.³⁶

Does this mean that under the Ninth Circuit's interpretation of Frye, peer review and publication are general prerequisites for the admissibility of scientific evidence? The opinion doesn't quite say so, but a couple of passages do suggest that reading. The court says that to be admissible the evidence "must meet all of the essential requirements" "imposed by the scientific community," and it quotes a passage from a muckraking book on "junk science in the courtroom:"

[T]he best test of certainty we have is good science—the science of publication, replication, and verification, the science of consensus and peer review.³⁷

Judge Kozinski's opinion in *Daubert* is poorly reasoned and carelessly written. If it means what it seems to mean, it might have changed the use of scientific evidence in a big way. *Frye* said that evidence based on *novel* scientific *techniques* is inadmissible until the basic principles and technology have become "generally accepted." Kozinski suggests that every scientific *finding* that an expert witness relies on—and perhaps every *conclusion* the expert testifies to—must be "generally accepted," as demonstrated by publication in a peer-reviewed journal. This would have been a vast extension of *Frye*, and it would have been utterly unworkable.

Peer review is the procedure by which articles are evaluated for possible publication in scientific journals. It is an extremely useful institution for the purposes for which it was designed, but not as a test for admissibility in court. Given the Supreme Court decision in *Daubert*, I will mention only some of the objections to the peer-review version of *Frye*, and only in the briefest passing.³⁸

(1) Peer review is not a substantive evaluation of the truth of the findings reported in the scientific research. The reviewers, however pos-

itive their evaluations, do not attest to the accuracy of the study; their assignment is simply to evaluate the appropriateness of the methodology, the completeness of the reporting, and the importance and novelty of the topic.

- (2) Studies that are rejected for publication by peer review are not necessarily wrong or misleading. In fact, there are famous examples of articles that were turned down at first and went on the become classics in their fields, (and famous examples of published peer-reviewed nonsense).
- (3) Peer review is a slow process that is designed for substantial works of original research. Most scientific testimony in court cases consists of specific applications of established principles. Such work is usually unsuitable for publication because it has little or no general scientific interest, and publication is often impractical because it must be completed and presented quickly to meet the timetable of litigation.
- (4) Courts could not successfully apply a standard based on an editorial process with which they are so thoroughly unfamiliar. Worse, making peer review an issue in legal proceedings would have predictable and unfortunate effects on the operation of scientific journals. Judge Kozinski's opinion contains no definition of a "peer-reviewed journal." The world of science already includes a large and varied collection of publications that fit that description, ranging from those at the top that are extremely selective (e.g., The New England Journal of Medicine), to some at the bottom that are not so choosy. A peer review-based Frye test would create an incentive to send any research that might be the subject of testimony straight to the bottom, in order to avoid the danger of rejection. Worse, it might spark the creation of ersatz "peer reviewed journals." If they printed their purpose on the masthead - The Annals of Impending Testimony or the Journal of Unpublishable Results - they might be spotted. But more realistic ventures, with respectable shells and mixed purposes, would get by.

B. The Supreme Court Opinion

The Supreme Court used *Daubert* as a vehicle to review the viability of the *Frye* test, and to reject it.³⁹ Given that decision, it was unnecessary to consider the peculiar additions embroidered on to *Frye* by Judge Kozinski—extending *Frye* from scientific techniques to the conclusions of expert witnesses, and (apparently) equating "general acceptance" with peer review.⁴⁰

Frye was widely followed by American courts, but never universally endorsed. For decades, it was the dominant authority on the admissibility of scientific evidence, ⁴¹ but by 1992 its heyday had passed. It was the rule in seven federal circuits, ⁴² four circuits had rejected it, ⁴³ and the First Circuit—along with the Supreme Court—had not spoken on the issue. ⁴⁴ Many state courts followed Frye, but some did not. ⁴⁵ Almost

all cases applying *Frye* were criminal prosecutions; ⁴⁶ *Daubert* was only the third federal civil case to use this rule. ⁴⁷ In the two decades before *Daubert*, *Frye* had been widely attacked as outmoded, unmanageable, and unduly restrictive. ⁴⁸ Critics argued that the "general acceptance" test required judges to make decisions they were not competent to make, imposed an excessively high burden on the proponents of novel scientific evidence, excluded valuable information at the frontiers of knowledge, and produced arbitrary results depending on how a court defined the "field" in which the evidence had to attain acceptance. Moreover, since the adoption of the federal rules of evidence in 1975, the legal status of *Frye* (in federal courts at least) had been under a cloud. The petitioners in *Daubert*, and many evidence scholars before them, argued that *Frye* had been repealed by Rule 702. The Supreme Court agreed—correctly, I believe—and reversed.

The question turns on Federal Rule of Evidence 402, which states:

All relevant evidence is admissible, except as otherwise provided by the Constitution of the United States, by Act of Congress, by these rules, or by other rules prescribed by the Supreme Court pursuant to statutory authority. Evidence which is not relevant is inadmissible.

In other words, Rule 402 abolishes common law objections to evidence. Since Frye's "general acceptance" test is not mentioned anywhere in the Federal Rules that govern expert evidence, Frye could only survive if it is a reasonable and useful interpretation of the relevant portion of Rule 702: "If scientific knowledge will assist the trier of fact . . . a witness qualified as an expert . . . may testify thereto. . . ."

Nothing about Rule 702 itself suggests that the unstated qualification "generally accepted" should be read into it. The best argument that could be made is that "general acceptance" is an implied quality of "scientific knowledge." This would have been a strained reading, unless the Court were willing to derive the insertion from the historical context: Frye was widely applied in 1975, and since the drafters of the Federal Rules of Evidence did not explicitly reject it, they must have endorsed it. In fact, the context suggests the opposite, since (as the Court noted in Daubert) it was the clear intention of the drafters to liberalize the use of expert testimony. 49 But the strongest arguments against Frye are not historical or textual but practical. A majority on the Court undoubtedly believes that federal courts have the authority to devise judge-made rules to fill in gaps, resolve ambiguities, and help provide consistency in the application of Federal Rules generally. They certainly do it themselves. But why this rule? It's arbitrary, it was not much good to begin with, by now its problems are famous, and Daubert presented a clear occasion to reconsider and reject it.

Having rejected *Frye*'s "general acceptance" test for scientific testimony, the Supreme Court proceeded to devise one of its own. Under Rule 702, the trial judge must decide two preliminary issues⁵⁰ in order to admit any scientific evidence:⁵¹ "whether the expert is proposing to testify to (1) *scientific knowledge* that (2) will *assist the trier of fact* to understand or determine a fact in issue."⁵² Or, to phrase it differently, judges must decide "whether the reasoning or methodology underlying the testimony is scientifically *valid* and whether that reasoning or methodology can be *properly applied* to the facts at issue."⁵³

The Court recognizes that this is a "flexible" inquiry,⁵⁴ and that "many factors will bear on it."⁵⁵ It then goes on to list four that are likely to become the focus of future litigation: (1) Is the theory or technique testable, and has it been tested? If not, it probably does not qualify as science.⁵⁶ (2) Has the theory or technique "been subjected to peer review and publication"?⁵⁷ This ought to be "a relevant though not dispositive consideration in assessing the scientific validity" of the evidence.⁵⁸ (3) If the evidence concerns a particular scientific technique, what is its "known or potential rate of error"?⁵⁹ (4) Finally, "general" or "widespread" acceptance, while not a requirement, "can be an important factor in ruling particular evidence admissible. . . . "⁶⁰

C. Did Frye Matter?

So Frye has been rejected; so what? At first blush, I'm tempted to say "Not much"—and that may prove to be so. Frye had two aspects. It embodied a recognition that courts must (in some manner) screen purportedly scientific evidence that is presented to them. That is an enduring and uncontroversial precedent. But Frye also announced a specific test for admissibility, the test that was rejected in Daubert. As a specific test, Frye may never have mattered nearly as much as all the fuss suggests.

As I've mentioned, several states never adopted Frye, and others explicitly rejected it within the past ten years. Has practice in these states differed markedly from that in states that follow Frye? Not in any major way. For example, most states continue to exclude polygraph evidence—at least in the absence of a stipulation by the parties—regardless of whether they follow Frye. Among the few state decisions admitting polygraph evidence without a stipulation, the two earliest were in New Mexico, which followed Frye, and Ohio, the two earliest were in New Mexico, which followed Frye, and Ohio, the two earliest were in New Mexico, swhich followed Frye, and Ohio, the two earliest were in New Mexico, which followed Frye, and Ohio, the two earliest were in New Mexico, which followed Frye, and Ohio, the two earliest were in New Mexico, which followed Frye, and Ohio, the two earliest were in New Mexico, which followed Frye, and Ohio, the two earliest were in New Mexico, which followed Frye, and Ohio, the two earliest were in New Mexico, which followed Frye, and Ohio, the two earliest were in New Mexico, and later explicitly rejected it. The followed Frye and Ohio, the two earliest were in New Mexico, and later explicitly rejected it. The followed Frye and Ohio, the two earliest were in New Mexico, and later explicitly rejected it. The followed Frye and Ohio, the followed Frye and I at the followed Frye and I

the technique had no scientific basis. ⁶⁹ In that case the admissibility of the evidence did bear *some* relationship to the test applied, since the admissibility of voice print evidence was the context in which a few courts rejected *Frye*. ⁷⁰ Even so, admissibility could not be predicted from the local status of *Frye*, especially not among the more recent cases. ⁷¹ Finally, on the most hotly debated scientific evidence issue of the decade—the admissibility of DNA identification evidence—the courts are nearly unanimous in letting it in, regardless of the test they use. ⁷²

Daubert is not the first federal case that both rejects Frye and directs judges to screen scientific evidence by reference to a set of several factors. Eight years earlier, in United States v. Downing, 73 the Third Circuit produced its own alternative to the "general acceptance" test: a list of eight factors, several of which overlap with those later listed by the Supreme Court. 74 And yet, there is no particular reason to believe that practice in the Third Circuit with regard to scientific evidence has diverged from that in other circuits; Bendectin claims, for instance, have fared about as poorly there as elsewhere. 75

It should not be much of a surprise that the formal test for admission of scientific evidence seems to make little difference. For one thing, the two types of tests are more similar than they may seem. *Daubert* and *Downing* both include "general acceptance" as a factor to be considered among others, and while *Frye* says that "general acceptance" is the only criterion, there is rarely direct evidence on this point, so judges must look for circumstantial indicia of "general acceptance"—including, conspicuously, reliability. As the Sixth Circuit put it:

[W]e deem, general acceptance as being nearly synonymous with reliability. If a scientific process is reliable, or sufficiently accurate, courts may also deem it "generally accepted."⁷⁶

More important, *Downing* and *Daubert* promise more than they can deliver. They avoid some of the arbitrary consequences (read: arbitrary exclusions) that may follow from mindless reliance on "general acceptance," if judges are knowledgeable and make good use of their new freedom in conducting the "flexible inquiries" that are now required—a significant if. *Frye* was criticized because the "general acceptance" of a principle or technique is a scientific issue that courts are not competent to evaluate. *Daubert* and *Downing* require them to make that same evaluation and a good deal more.

In most cases, of course, judges don't do any of this. They decide whether to admit scientific evidence not by determining global facts (is the polygraph a "generally accepted" test for deception?), nor by weighing factors pro and con (Testability?—yes; tested—yes; peer review—no; and so forth), but by referring to specific governing precedents: has DNA identification been admitted in this state? Has that decision been affirmed by the State Supreme Court? Ultimately, it seems, the major precedential

decisions on which trial courts rely tend to converge on the same results, regardless of the test used.

Nonetheless, Daubert may have substantial impact after all. On its face, the Supreme Court's decision loosens the standard for admitting scientific evidence by overruling the restrictive test announced by the District of Columbia Circuit seventy years earlier. In theory, such a ruling could produce a sharp increase in the quantity of questionable scientific evidence that is admitted in federal courts. In practice, not only is such a shift unlikely, but the decision is more likely to cause a modest shift in the opposite direction. The Daubert opinion recognizes - no, celebrates - the role of the trial judge as a gatekeeper who screens scientific evidence on its way to the jury. The discretionary power of the trial court judge is broadcast at every turn. The Court removes any ambiguity that the decision is for the judge, not the jury.⁷⁷ It lists multiple statutory bases for decisions to exclude - Rule 702, Rule 703, and Rule 403, the catch-all provision that gives judges discretion to exclude relevant evidence if it is unduly confusing, wasteful or prejudicial.78 The Court allows judges to base their decisions on a wide range of factors rather than on a single and apparently specific test, 79 and it explicitly recognizes and accepts the costs of giving them this power:

We recognize that in practice, a gatekeeping role for the judge, no matter how flexible, inevitably on occasion will prevent the jury from learning of authentic insights and innovations. That, nevertheless, is the balance that is struck. . . . 80

Paradoxically, the main effect of *Daubert* may be to embolden some judges to exclude more purportedly scientific evidence than they would have under the "inflexible" and "austere" *Frye* test⁸¹ that *Daubert* rejects as unduly restrictive.⁸²

III. What's Really at Stake?

As we have seen, the real claim in *Daubert* was not that the data or the method of the plaintiffs' experts were unacceptable, but that the conclusion they reached was false. This concern surfaces time and again in the lower court opinions in *Daubert*, in the Supreme Court briefs submitted by Merrell and its *amici*, and in other Bendectin cases. The First Circuit's opinion in *Lynch v. Merrell-National Laboratories Division of Richardson-Merrell* is a good illustration.⁸³ After brief review of the studies, the court concludes:

[O]n the basis of the epidemiological evidence to date, Bendectin is as likely as aspirin to cause limb reduction [birth defects]. . . . The association of Bendectin with limb reduction is in the opinion of the health-care community an instance of popular delusion and error. . . .

247

A new study coming to a different conclusion and challenging the consensus would be admissible in evidence. Without such a study there is nothing on which expert opinion on Bendectin as a cause [of birth defects] may be based.⁸⁴

In Richardson v. Richardson-Merrell Inc., 85 the District of Columbia Circuit followed Lynch. In the process it had to distinguish an earlier toxic tort case, Ferebee v. Chevron Chemical, 86 in which it had affirmed a judgment based on evidence no stronger than that offered by the plaintiff in Richardson. The issue in Ferebee was whether exposure to the herbicide paraquat can cause pulmonary fibrosis. The evidence in support consisted of the opinions of two of the deceased plaintiff's treating physicians, who testified that they believed that had happened in this case. Apparently, no systematic studies of any sort were introduced, and even one of the plaintiff's experts concluded "that cases like Ferebee's were rare." Chevron argued that this was not enough, but the court disagreed: "The short answer to Chevron's argument is that two expert witnesses refuted it and the jury was entitled to believe those experts."

In Richardson, however, the D.C. Circuit explains that paraquat is one thing and Bendectin something else entirely. The relationship between paraquat and lung disease was untested, it stood "at the frontier of current medical and epidemiological inquiry." With Bendectin, however,

we are at the other end of the spectrum. . . . And far from a paucity of scientific information on the oft-asserted claim of causal relationship of Bendectin and birth defects, the drug has been extensively studied and a wealth of published epidemiological data has been amassed, none of which has concluded that the drug is teratogenic.⁹⁰

Any ambiguity about the meaning of *Richardson* was removed in *Ealy v. Richardson-Merrell*, *Inc.*, 91 a later District of Columbia Circuit case following it:

[A]n expert opinion that Bendectin is a human teratogen which caused the plaintiff's birth defects is without scientific foundation under Federal Rule of Evidence 703 in the face of "a wealth of published epidemiological data" to the contrary. . . . Accordingly such expert opinion is inadmissible. 92

The underlying rule that emerges is simple: "expert opinion evidence that Bendectin causes birth defects is inadmissible because it is wrong." It's also an oxymoron. This cannot be a test for admissibility. A court could hardly apply such a test and then require the jury to answer the question: "Does Bendectin cause birth defects?" A court that employs this test is obviously using "admissibility" as a disguise to determine the merits of the claim; it has already decided that Bendectin does not cause birth defects. In Daubert itself, the Supreme Court (without citing Lynch

or *Richardson*) disapproves this sort of evaluation: "The focus [of the judge's inquiry], of course, must be solely on principles and methodology, not on the conclusions they generate." 93

Why use such an internally inconsistent rule? The impulse to find for the defendant in Bendectin cases is perfectly understandable: the evidence is overwhelmingly in its favor. What's more, since Merrell keeps facing the same evidence in case after case and winning (at trial or on appeal), it's also understandable that judges would want to short-circuit the process and rule for the defendant as a matter of law. But why not do it the obvious way—why not simply rule (on a motion for judgment as a matter of law) that the evidence against Bendectin, while admissible, is insufficient to create a genuine factual dispute? Two federal circuits have done just that: the Fifth Circuit in Brock v. Merrell Dow Pharmaceuticals, Inc., 94 and the Sixth Circuit in Turpin v. Merrell Dow Pharmaceuticals, Inc., 95 both on records similar to that in Daubert. But three other circuits—the First in Lynch, the District of Columbia in Richardson and Ealy, and the Ninth in Daubert itself—all relied on admissibility rather than sufficiency to achieve the same result.

A scientific question is never presented to the courts in the abstract. It comes up in the context of a particular dispute between specific parties that must end with a binding final judgment. In the process, all factual issues must somehow be resolved, even if essential questions cannot be answered with confidence. This is true for specific historical facts (Was the light red or green? Did the defendant see something that looked like a knife before he shot the deceased?) as least as often as it is for general scientific issues (Does Bendectin cause birth defects in human fetuses? Does non-ionizing radiation from electrical power lines and substations cause cancer?). But there are two differences. (1) The exact same scientific question can come up repeatedly, in dozens or in thousands of cases; similar patterns are rare for ordinary historical issues, and when they do occur (e.g., common factual issues in airline crash cases), the number of trials involved is likely to be comparatively small. (2) There is a well developed and highly successful method for investigating such questions - the scientific method - that it is at odds with the requirements of legal fact finding. Scientists are interested in truth, which is elusive; therefore they try to be comfortable with uncertainty. They work with hypotheses and theories that can be disproven but not directly proven. The most interesting and innovative scientific thinking is usually provisional, and even established scientific findings - in principle, all scientific findings-are subject to revision. Courts are happy to find truth, if they can, but (truth or no truth) they are required to decide cases by judgments that cannot normally be revised if better information becomes available.

All legal systems need some method or another for reaching final judgments on uncertain facts. In our system the archetypal procedure for doing so is trial by jury. In fact many legal disputes—probably most—are tried by judges without juries, 96 and some of those include tricky scientific questions. But the prominent cases with problematic scientific evidence seem to occur in categories that are generally tried by juries: criminal prosecutions and personal injury suits of one sort or another. As a result this issue has been framed in terms of fact finding by juries.

A jury is a one-time, ad-hoc, non-professional decision-making group. Juries are probably more unpredictable than other fact finders because each jury is constituted for one case only. A jury has no collective experience of its own to learn from, no track record for litigants to rely on, and almost no opportunity to learn from the experience of other juries. Juries are also delphic. They are never asked to justify or explain their decisions, or to describe the evidence on which they relied. A typical jury verdict consists of very few words-the name of the winning side, and (if it's the plaintiff in a civil case) the amount of the judgment. In unusual cases, civil juries are sometimes asked to answer one or several factual interrogatories; for example, the jury in the consolidated Bendectin trial answered "no" to the question: "Have the plaintiffs established by a preponderance of the evidence that ingestion of Bendectin at therapeutic doses during the period of fetal organogenesis is the proximate cause of human birth defects?"97 But even this unusual procedure makes the jury's judgment only slightly less opaque.

Finally, and most important, jury decisions are substantively unreviewable. This is an overstatement, but only a slight one. Judges do occasionally grant motions for new trial or for judgment notwithstanding the verdict, but these are limited restraints. They are infrequently exercised, and they require the judge to decide that the jury's final verdict is unreasonable or inconsistent with the evidence. There is no procedure for a court to review the quality of a jury's reasoning or the accuracy of the factual premises on which it relied, since these intermediate steps are neither recorded nor expressed, and the jurors themselves are not allowed to "impeach their verdict" by telling the court how it was reached.98 Indeed, it is hard to avoid the impression that on the whole judges not only accept the fact that jury verdicts are impenetrable, they relish it. Faced with what looks like an unanswerable question-does paraquat cause pulmonary fibrosis?-they seem happy to say: "Given the conflicting evidence, the jury was entitled to find. . . . " Or, to put it unkindly: "Following a proper ceremony, the question was put to the oracle. . . . " This argument sells. It's not exactly a religion, but the jury may be the most popular and trusted institution we use to exercise governmental authority.

Judges, of course, are not mindless jury fans. Quite the contrary. They worry that juries are too readily confused, too easily swayed by passions, too willing to ignore fixed rules of law out of sympathy or anger. Much of the elaborate common law of evidence is based on this systemic

mistrust of juries. Some judges also complain that juries are unsuitable for "complex" cases, especially those that depend on expert and technical evidence — although it's not at all obvious that judges themselves handle such cases better. Finally, juries cost a lot. It is expensive, in both cash and opportunity costs, to run trials in the dramatic, over-prepared, do-or-die, one-time continuous performance mode that is required by this awkward fact finder.

These are general characteristics of juries. There is also a particular problem that looms large in the context of Bendectin and other toxic tort cases: Juries tend to be inconsistent in the sense that different juries reach different judgments on the identical issue. This has been shown repeatedly by experimental studies of juries 100 - but it is also necessarily true of jurors and of individual decision makers, including judges. Indeed, jury research shows that the collective process of deliberation reduces interpersonal variation in judgment.¹⁰¹ If each case presented a new issue, juries might have no disadvantage in this respect over judges or any other type of tribunal, or they might be better. But when the same issue is tried repeatedly jury decision making suffers by comparison not because juries are stupid but because they are ephemeral. The jurors in a Bendectin case are unlikely to be influenced by the decisions of other juries dealing with the identical issue for the simple reason that they won't know about them. The judgments of other juries are inadmissible as evidence that what those juries found is true-that would be hearsay and inadmissible opinion evidence-and (in the usual case) there will be no other jury issue on which those judgements are relevant. Equally important, jurors will typically have no information about other similar cases through informal channels.

A judge, however, will know. Many of the earlier decisions (or appeals from them) are published in case reports. The judge may have seen them on her own, heard about them from colleagues, or even read about them in a law review article; if not they will be brought to her attention in motions to dismiss or for summary judgment, or informally in pretrial discussions of motions, dates, and possible settlements. Trial judges, therefore, are able to make their findings consistent from case to case, if they want to. If they don't want to, consistency can be more easily achieved on review. Appellate judges may be more willing to reverse the verdict of other judges than those of juries, and they certainly have more to work with, since judges must justify their decisions with written findings of fact and conclusions of law.

In the Bendectin cases, most juries came to the correct conclusion, assuming (as I do throughout) that Bendectin is not in fact a teratogen. Of 22 jury trials that produced verdicts, the defendant won 14, for an overall success rate of 64%. (There were also 2 hung juries and one mistrial.)¹⁰² But that also means that juries were wrong in 36% of the cases, and wrong in the face of overwhelming scientific evidence for the

251

defense. It is tempting to conclude that judges would have done better. but it's not obvious that they are any more skillful than juries at evaluating evidence of this sort. For example, Wells v. Ortho Pharmaceutical Co. 103 was another product liability law suit in which the plaintiff claimed that her birth defects were caused by a product that her mother used during the plaintiff's gestation. The product in Wells was orthogynol spermicidal jelly, a popular contraceptive that is widely used in the United States in conjunction with diaphragms. As with Bendectin, the epidemiological evidence in Wells was strong and one-sided: several studies had shown that orthogynol spermicidal jelly is not a teratogen. 104 Wells was tried without a jury, and despite this evidence the judge awarded the plaintiff a judgment of \$5.1 million. 105 Wells was the first orthogonal trial. (In the only other case of which I am aware, the defendant won a motion for summary judgment. 106) If the Bendectin litigation had consisted of 25 first trials, the outcomes may have been roughly the same whether tried by judges or by juries.

There was, of course, only one first trial, which was retried, followed by a second, followed by the consolidated trial in the Southern District of Ohio, and so forth. By the end of 1985 the defendant had won 3 of the 4 cases that had been completed, including the consolidated trial. At that point the emerging pattern could have become a *de facto* trial-court rule, if juries had no role in the remaining cases. In any new Bendectin case, everybody—including the trier of fact—would know in advance that Merrell had already won repeatedly on the same issue and the same evidence. To rule against Merrell would seem both unfair and inefficient, and (since appellate courts could be expected to feel the same way) would court reversal. As it happened Merrell won both of the Bendectin bench trials that went to verdict after 1986 (the only bench trials in the entire set), but only 11 of the 17 jury trials. 107

(Merrell's attorneys did make this point, as best they could. For example, in *DeLuca v. Merrell Dow Pharmaceuticals*, ¹⁰⁸ the Third Circuit comments that the defendant's "apparent litigation strategy . . . was to emphasize that 'in all material respects, the instant case is identical to the cases where summary judgment has been granted in Merrell Dow's favor.' "109 This was no doubt an excellent plan even, as in *DeLuca*, in the context of a motion for summary judgment; it would have succeeded more readily if the argument could have been addressed to a judge with plenary power to weigh the evidence.)

The only way to achieve consistency in a set of cases of this sort is to take them away from juries. The most effective method is to bar the underlying claim, procedurally (by a statute of limitations, by some form of immunity, or whatever) or substantively (by holding that there is no cause of action for these injuries against this defendant)—if there is some legal basis for doing so. In the case of Benedectin, there appears to have been none. If that's not possible, the courts will be forced to

dismiss any suits that are filed on one of the two available evidentiary grounds: by holding that the evidence is *insufficient* as a matter of law, or by holding that it is *inadmissible*. As the Bendectin cases illustrate, courts do not like to admit evidence and say it's insufficient. Instead, they will go to unfortunate lengths to find that essential parts are inadmissable, and then say that there's not enough left to go to the jury.

The basic problem seems to be that judges do not want to look as though they are abrogating the role of the jury as trier of fact. The legal sufficiency of evidence is, technically, a question of law, but it looks and sounds like a judgment on the weight of the evidence—it is a judgment on the weight of the evidence, only an extreme one. This is particularly true for expert evidence, since traditionally courts have held that the testimony of any qualified expert is sufficient to sustain a verdict on any issue on which she testified.

Expert testimony is admitted in evidence as an exception to the usual rules limiting opinion evidence and restricting witnesses to facts within their personal knowledge. In effect, such evidence is an exception to a basic division of functions in common-law trials: witnesses *present* information, jurors *evaluate* it. *Expert* witnesses *evaluate* facts as well as present them; indeed, that is often their main function. The frequently they testify to evaluations—expert conclusions—that are identical to those the jury must make, and yet we consider this *evidence*. It is hard to say that the testimony of a witness who speaks in such terms—"the plaintiff is diabetic" or "Bendectin causes birth defects"—is admissible but insufficient to support a finding of fact that simply restates the testimony. That would amount to saying: "The jurors may hear and consider this witness—in fact they are required to do so—but they are also required to disbelieve her."

Evidence is often said to be either circumstantial or direct. Circumstantial evidence of a fact at issue is evidence which (if believed) supports an inference that the fact exists. Testimony that there were bootprints in the snow (to choose a traditional cliche) is circumstantial evidence that a person has walked by, since the conclusion depends on the (extremely plausible) inference that the prints were made by a walking person. Direct evidence is evidence which, if believed, directly proves the fact at issue, with no intervening inference—for example, testimony from a witness who claims to have seen a person crossing the snowy field. McCormick explains how the issue of sufficiency differs in the context of these two types of evidence: "Generally no difficulty occurs where the evidence is direct. Except in rare cases, it is sufficient, though given by one witness only, however negligible a human being she may be. But if the evidence is circumstantial, forensic disputes often arise as to its sufficiency. . . . "112"

Strictly speaking, expert opinion testimony is neither direct nor circumstantial evidence in the usual sense of these terms. It is a judgment

about facts that the expert has considered. However, an expert opinion on causation in a Bendectin case *sounds* very much like direct evidence: the opinion, if believed, resolves the issue at stake with no additional steps. As a consequence, the testimony of a single expert—"however negligible"—is often given the same dignity (for the purpose of rulings on the sufficiency of evidence) as the testimony of a lone eyewitness. A case from the California court of appeals illustrates this link:

The testimony of a medical witness in answer to hypothetical questions based on the facts in the record is sufficient to support a finding contrary to the testimony of other medical experts who have seen and examined the patient. The testimony of one credible witness, if believed, is sufficient to support a finding of the court. 113

In the context of toxic tort litigation, this classic position was restated by the District of Columbia Circuit in Ferebee:

[I]f experts are willing to testify that such a link exists [between long term exposure to paraquat and pulmonary fibrosis] it is for the jury to decide whether to credit their testimony.¹¹⁴

But, of course, if the court finds the expert evidence *inadmissible*, the jury never faces this task.¹¹⁵

In short, the nature of expert testimony invites substantive judicial review. Since expert witnesses are allowed to pronounce judgment on the central issues in dispute—a prerogative we otherwise reserve to juries and judges—it is natural that the courts sometimes intervene to evaluate the quality of their evidence. Our rules of practice, however, drive judges who do so to pretend that they are acting on a question of procedure, namely, admissibility.

Unfortunately attempts to use admissibility to achieve consistency in toxic tort litigation do not work, at least not if the attempt resembles anything like a true rule of evidence. On the one hand, a court can get specific and say, as the District of Columbia Circuit has said, that "an expert opinion that Bendectin is a human teratogen . . . is inadmissible."116 But it is unprincipled to pretend that this is really an evidentiary ruling. Rulings on objections to evidence are supposed to be made on the basis of categorical criteria-testimony is hearsay because it repeats an outof-court statement and is offered to prove the truth of that statement; a document is not privileged because it was not written as a communication to an attorney; and so forth. To be sure, many evidentiary rulings require the court to consider the item in question in the context of other evidence. (This is particularly true for objections that evidence is irrelevant in the absence of other evidence of a necessary precondition, 117 or that the probative value of evidence is substantially outweighed by the danger of confusion and prejudice. 118) But it takes judicial chutzpah to decide the

central factual issue of a case, grant summary judgment, and call it a ruling on admissibility.

A court could try to achieve the same result by devising a genuine rule of exclusion. But the tool does not fit the task. Admissibility, like liability, is an either/or proposition. Not surprisingly, courts that do this try to define evidentiary requirements that have the same dichotomous quality: general acceptance, statistical significance, peer review. At first blush, this seems appropriate: a technique or a finding either is or is not generally accepted, or statistically significant, or whatever, so a court can easily say whether it is or is not admissible. In practice, these rules break down, one after the other.

The worst problem with such rules is the misuse of scientific concepts. Scientists do not use general acceptance or peer review as preconditions for *considering* information, and most no longer use statistical significance for that purpose, if they ever did. Peer review¹¹⁹ and tests of statistical significance¹²⁰ were designed for other purposes, and general acceptance is not a scientific criterion for anything. Sometimes the most informative findings on an issue will be not statistically significant (perhaps because of a limited number of observations), or not published, or not generally recognized.

In addition, the major virtue of these rules turns out to be illusory: In practice, they are not easy to apply after all. Statistical significance looks like a readily determinable attribute, but in fact it can become quite tricky. For example, in many situations an expert (or a lawyer) can select (or manipulate) the data that are subjected to this statistical test – as the use of meta analysis in the Bendectin cases illustrates. Peer review looks like a straightforward issue-did it take place or didn't it?-but that may be true only so long as peer review does not have determinative forensic significance. If peer review became a general requirement for admissibility, it would also become distorted and problematic, and courts might be forced to settle disputes over whose peers and which journals are good enough to count. And general acceptance, of course, is a notoriously slippery concept, even as applied to theories and techniques. You can't see it; you have to infer it from whatever evidence you can gather about the field assuming that the field is somehow well defined. An attempt to determine the general acceptance of a particular scientific conclusion quickly becomes an inquiry into its truth-which, again, cannot be a method of determining the admissibility of evidence of the truth of that very same conclusion.

Finally, creating a rule of admissibility in order to resolve a substantive problem breeds new problems in future cases. It may satisfy the needs of the moment to say that the expert's opinion is inadmissible because there is no epidemiological research, no statistically significant finding, or no peer reviewed study. But will this test work in the next case up? Should we have to go through a plenary trial in a Bendectin case (or

five of them) just because a plaintiff produces evidence of a statistically significant finding in a single epidemiological study of uncertain quality that was published in a "peer reviewed" journal of extremely modest distinction? Should we dismiss outright a case like *Ferebee*, or a stronger case, in which the available evidence consistently suggests that the substance in question did cause the plaintiff's pathology, because no epidemiological studies have yet been completed?

IV. What to Do?

A. Easy Cases

Daubert was an easy case in more senses than one. Not only was the legal issue before the Court a clear call, but so was the underlying factual dispute as well. By 1993 just about everybody agreed that Bendectin is not a human teratogen, or at least that there is no substantial evidence that it is. This consensus is reflected in the decisions of most juries that tried Bendectin cases, and in the actions (if not always words) of almost all judges who presided over and reviewed those trials. The final problem was, and is, merely procedural: how to dispose of the remaining cases efficiently and consistently.

The simple straightforward way to do so would be to redefine the question of causation—does Bendectin cause birth defects?—as one of "law" (to be decided by the judge) rather "fact" (to be decided by the jury). At first blush, this sounds like cheating. How can this be anything other than a question of fact? But courts constantly resolve factual questions in the process of deciding "legal" issues, from "was the statement made under a belief in impending death?" ¹²¹ to "is this the sort of evidence on which experts in the field reasonably rely in reaching conclusions on such matters?" ¹²² The issue is whether this factual question is for juries to decide or for judges; the category in which it is placed, "factual" or "legal," is simply a label we attach to that assignment. In this case, there is a strong argument that it should be given to judges.

To be sure, the question of causation in a tort case is one we normally submit to juries. But not always. A plaintiff who claimed that the defendant caused his ulcer by witchcraft would not get far, nor would a defendant who admitted exposing the plaintiff to staphylococcus but denied that bacteria cause illness. Either case would present a traditional context for judicial notice that the claims are wrong, since contrary propositions are widely regarded as indisputable. Unfortunately, the case for Bendectin is less clearcut. Even given the lopsided pattern of epidemiological findings, the claim that Bendectin is a human teratogen is

hardly "indisputably" false—although one could argue that it is indisputable that the available evidence is inadequate to prove that claim by any standard.

In other contexts, courts frequently make similar factual decisions without requiring anything like indisputability. When a court decides that five-member juries function less well than six-member juries, ¹²³ or (applying *Daubert*) that DNA profiling is a reliable technique for determining the identity of a suspect in a criminal case, ¹²⁴ it is making a factual determination on a debatable question as a step in the process of announcing a "rule of law." In 1942, Professor Kenneth C. Davis named this age-old process "legislative fact finding," ¹²⁵ and the name has stuck. Legislative fact finding is often described as a species of judicial notice, ¹²⁶ which makes some sense, since under both procedures the court makes factual decisions and is not restricted to the record in doing so. But, as Professor Davis wrote, the role of the common law judge requires more latitude than the traditional restrictions on judicial notice permit:

[J]udge made law would stop growing if judges, in thinking about questions of law and policy, were forbidden to take into account the facts they believe, as distinguished from facts that which are "clearly . . . within the domain of the indisputable." 127

Judging from published opinions, legislative fact finding must be something like sex: judges do it all the time, but rarely talk about it in public. Most if not all decisions that announce significant innovations in legal doctrine include factual premises that are open to dispute, and typically these premises rest at least in part on extra-judicial information. But very few opinions actually mention this embarrassing activity by name, and those that do generally keep the references to the margins—in a footnote, ¹²⁸ in a concurring or a dissenting opinion, ¹²⁹ describing what some other court has done in some other case. ¹³⁰ The reason, I suppose, is that the name smacks of judicial legislation, which is a political taboo. In any event, this shyness makes it possible to overlook the fact that the practice (like sex) is essential, widespread, and accepted.

The real issue is not whether legislative fact-finding by courts is legitimate, but when and where. That question requires a comparison with the competing procedure, adjudicative fact finding on the records by the trier of fact:

When a court or agency finds facts concerning the immediate parties—who did what, where, when, how, and with what motive or intent—the court or agency is performing an adjudicative function, and the facts are conveniently called adjudicative facts. . . .

Stated in other terms, adjudicative facts are those to which the law applies in the process of adjudication. They are the facts that normally go to the jury in the process of adjudication. They relate to the parties, their activities, their properties, their businesses.¹³¹

By contrast:

Legislative facts are those which help the tribunal to determine the content of a law and policy and to exercise its discretion in determining what course of action to take. Legislative facts are ordinarily general and do not concern the immediate parties. 132

Many issues can be neatly classified by this dichotomy. How fast was the bus going?—Adjudicative. Are buses subject to the speed limit for automobiles or that for trucks?—Legislative. Did the defendant know that it was wrong to kill the deceased?—Adjudicative. What is the test for legal insanity?—Legislative. Unfortunately, the main issue at stake in the Bendectin cases—is this drug a teratogen?—falls between these two stools. The factual issue is not one that concerns the immediate parties—"who did what, where, when, how, and with what motive or intent"—but it does look like the type of issue that normally goes to the jury. Is it the sort of fact that helps judges "determine the content of law"? Maybe, but that merely begs the question. The issue at stake—causation in a tort case—is one we ordinarily submit to juries; but that doesn't mean that we are required to do so in this context.

Courts are not reluctant to base findings of legislative fact on scientific evidence. For example, several courts have held that the results of a blood-grouping test that purports to exclude paternity are conclusive. The only issues left for the jury are whether the test was properly conducted, and whether that was the finding:

There should be no occasion for expert testimony in every case to prove the scientific validity of blood-grouping tests resulting in exclusion of paternity. The scientific opinion on that point is so general that courts may take judicial notice of it in filiation proceedings. 133

And, of course, courts are perfectly comfortable making legislative findings of scientific fact when they decide that a scientific technique has gained general acceptance or is sufficiently reliable to warrant admission in evidence. For example, in *United States v. Jakobetz*, ¹³⁴ after a detailed discussion of the scientific bases and techniques of DNA profiling, the Second Circuit held that "in future cases with a similar evidentiary issue, a court could properly take judicial notice of the general acceptability of the general theory and the use of these specific techniques." ¹³⁵

The evidence on the issue of general causation in the Bendectin cases is roughly the same type as the evidence in the DNA profiling or the blood-group paternity cases. The conclusion the courts might reach, however, is not so similar. In the cases on DNA profiling, the courts have found that the technique is sufficiently reliable to be admissible. This is a traditional, old-fashioned evidentiary ruling that opens the door to a category of evidence. In the absence of other information, the outcome of a particular case cannot be predicted from such a ruling; the

DNA evidence could help prove a defendant's identity as the criminal, or it could help prove his innocence, or it could be indeterminant. The decisions on blood tests in paternity cases are closer to a legislative finding that Bendectin does not cause birth defects. The statement "a drug cleared by 30 epidemiological studies did not cause the plaintiff's birth defects" does sound quite a bit like the statement "a man cleared by blood tests is not the father of the infant in this case." In fact, the former statement is more sweeping. It is one thing to decide that a blood test (or any other item of scientific evidence) has determinative weight on an issue; the jury still has to find that this particular man was cleared by a properly conducted test. It is another matter to conclude that the universe of available evidence will not support a claim. To do that is to abolish a cause of action in an entire set of cases, without regard to the position of any individual party.

In an article on the use of science in the courts published nearly thirty years ago, Professor Harold Korn wrote:

The effect of characterizing a determination as "legal" is to place it within the court's sphere of influence and to acknowledge that the determination has the capacity for authoritative assimilation into the legal system as transmissible doctrine. ¹³⁶

Judges, according to Korn, engage in a "ceaseless search for consistency and predictability in the legal order." They achieve this consistency by a form of forensic manifest destiny; to quote Thayer, a jury question "is likely to be absorbed by the judge, 'whenever a rule about it can be laid down." ¹³⁸

The Bendectin cases cry out for absorption into transmissible legal doctrine. And in fact, that has taken place. The unmistakable meaning of the last decade of Bendectin litigation is this: No claim that Bendectin caused a birth defect will be considered. This is a rule of law. The rule could be changed in the unlikely event that some surprising new research undermines the factual premises on which it is based, but that is not unusual. What is slightly more unusual is that the courts are reluctant to own up to what they have done.

The two Circuit Courts that held the evidence against Merrell insufficient as a matter of law were fairly explicit about what they were doing. The normal standard for decision will not support their decisions—the plaintiffs' expert evidence, considered in isolation, is legally sufficient to go the a jury—but these are not normal cases. The essence of the problem is the generality of the main issue. The Sixth Circuit:

The cases are variations on a theme, somewhat like an orchestra which travels to different music halls, substituting musicians from time to time but playing essentially the same repertoire. 139

[I]n mass torts the same issue is often presented over and over to juries in different cases, and juries often split both ways on the issue. . . . Appellate courts, if they take the lead in resolving those questions . . . can reduce some of the uncertainty which can tend to produce a sub-optimal amount of new drug development. 140

Both Circuits were also careful to limit their rulings to Bendectin cases. The basis for the decisions is the plaintiffs' failure to present epidemiological evidence that Bendectin is a teratogen; however, "we do not hold that epidemiological proof is a necessary element in all toxic tort cases," and "[w]e do not mean to intimate that animal studies lack merit or power when it comes to predicting outcomes in humans." It is only in the context of this record, with all the negative epidemiological studies, that such a failure is fatal. These courts did not quite say "we have weighed the evidence on both sides, and find as a matter of law that Bendectin is not a teratogen," but they came close, and their meaning is clear.

Even this level of explicitness goes against the grain. Summary judgment is only supposed to be granted when the losing party has failed to present sufficient evidence to sustain a verdict, not when it has presented substantial evidence but the opposition has presented an overwhelming rebuttal. At least, that's the rule when summary judgment is ordered on the "facts," because of insufficient evidence, rather than on the "law"—and this one looks like a decision on the "facts." So courts retreat to the safer ground of ruling on admissibility.

There was a strong push in that direction in the Fifth Circuit. In *Brock*, a suggestion for a rehearing en banc was rejected by a vote to 8 to 6, and the dissenters made clear that they were more concerned about the rationale for the panel's opinion than the outcome: "It strikes me that the issue in this case revolves around the admissibility of the expert testimony. . . . the panel [however] chooses to accept the admissibility of the testimony and to quarrel with its effect." Admissibility, of course, was the basis for decision in the First, Ninth, and District of Columbia Circuits, the other three federal appellate courts that have explicitly closed the door on Bendectin suits. As far as this set of cases is concerned, the effect is same. The route would be shorter, however, and the law of torts and of evidence would be clearer, if judges recognized what they were doing and said it.

B. Hard Cases

Litigation favors hard cases. Truly easy issues are not likely to be disputed at all, and if a disputed question *becomes* easy—e.g., Bendectin—the courts tend to weed out the cases that depend on it, by one means or another. They are replaced by a steady stream of new cases, raising

questions that are not so easy-at least not yet. It's for these hard cases that the evidentiary and substantive rules governing scientific evidence matter most.

Hard cases are not all equally hard. Some depend on issues that have been the subject of a great deal of scientific research-enough to make legislative fact finding possible, even though the conclusion is not foregone. A prime example is Judge Weinstein's famous opinion in In re "Agent Orange" Product Liability Litigation. 147 The cases decided by this opinion were claims by Vietnam veterans against various chemical companies, alleging that they had been injured by exposure to dioxin, a toxic contaminant of the herbicide "Agent Orange" which had been manufactured by those companies and widely used in the Vietnam war. The vast majority of the Agent Orange claims were included in a class-action settlement, engineered by Judge Weinstein. 148 However, 281 plaintiffs elected to opt out of this class action. These remaining "opt-out" cases were dismissed by Judge Weinstein, who granted the defendants' motion for summary judgment on the ground that the plaintiffs had presented insufficient evidence that their injuries had been caused by exposure to Agent Orange. 149 A lot has been written about Judge Weinstein's opinion in the opt-out cases. 150 I will add very little to that body of writing—only my own general conclusion that parts of the opinion are hard to defend, but that Judge Weinstein's basic judgment was correct. The plaintiffs could not prove that exposure to Agent Orange had harmed them, although the research was not as one-sided as the Bendectin studies. The Agent Orange opt-out opinion encouraged other courts to take similar action in other toxic tort cases, not least in the Bendectin cases. In Brock, for example, the Fifth Circuit saw itself as following Judge Weinstein's lead:

We are not without precedent in our approach to this problem. The case before us parallel's in many respects the recently conducted Agent Orange Litigation.¹⁵¹

A more recent toxic tort case, however, provides a better example of a comparatively hard case in which legislative fact finding may still be appropriate.

In In re Joint Eastern and Southern District Asbestos Litigation (the "Joint Asbestos" case), 152 the central issue was whether a deceased sheet metal worker's colon cancer had been caused by workplace exposure to asbestos. After a jury returned a \$4,510,000 verdict against several defendants, Judge Sweet of the Southern District of New York granted the defendants' motion for j.n.o.v. on the ground that the plaintiff had presented insufficient evidence of causation. Judge Sweet (who cites the Agent Orange opt-out opinion and various Bendectin cases profusely) appears to have excluded no scientific evidence at trial. Instead, he claims to apply the usual standard for deciding a motion for a judgment

as a matter of law: "[T]he sufficiency of the Plaintiff's epidemiological evidence must be determined by analyzing that evidence on its face and without weighing the evidence against it." His published opinion belies that claim. In a twelve-page discussion of the epidemiological and clinical evidence, the judge goes back and forth from the plaintiff's case, to the defendants' case, to the published literature, comparing and criticizing. A fair reading of the opinion leaves no doubt that his conclusion—"the Plaintiff's epidemiological evidence fails to support the claim that exposure to asbestos causes colon cancer" based on a detailed analysis of the entire record, on both sides.

The record on which Judge Sweet based his opinion is extensive. "During the course of the trial, approximately 45 epidemiological studies and surveys of studies were discussed by the parties' experts. . . ."156 Judging from the opinion, most of these studies show no association between asbestos and colon cancer, but a few do seem to show a comparatively weak correlation. Given these studies, Judge Sweet's conclusion sounds right: exposure to asbestos may increase the likelihood of developing cancer of the colon, but the effect (if there is one) is too small to justify a conclusion that any individual's cancer was more likely than not caused by asbestos.

Exposure to asbestos is a common problem in the United States, and colon cancer is a common cause of death. Inevitably, the possible causal relationship between these two events has become a recurring issue in litigation. 157 In that context-and given a sufficiently well developed scientific basis-legislative fact finding is a sensible way to handle the issue. That does not mean it's easy. Judge Sweet's opinion, like Judge Weinstein's Agent Orange opt-out opinion, clearly reflects long study and careful attention to abstruse technical issues. Most judges would probably never try to reach reasoned scientific judgments in cases where many studies find no causal relationship, but (unlike Bendectin) a few seem to go the other way. And just as well. It should only be done with great care. 158 Judge Sweet seem to have done well, but without a great deal of expert knowledge, there's no sure way to tell. The real test, of course, is time. 159 In any event, if judges do write opinions like these it would be better if they would say what they mean. If it is appropriate for the courts to make substantive scientific judgments on liability as a matter of law, they should not muddy the waters by pretending merely to apply procedural and evidentiary rules.

C. Harder Cases

The great majority of scientific issues that are presented in litigation are not even remotely suitable for legislative fact finding. This is true, of course, for legions of cases where the critical issue is the evaluation of one person or a unique occurrence. It is also true of the issue of general

causation in many toxic tort cases. Vann v. City of Woodhaven¹⁶⁰—which is described in detail by Professor Troyen A. Brennan¹⁶¹—is a good example. The plaintiffs in Vann were the parents of a 12 year old boy who died of aplastic anemia; they claimed that the disease was caused by exposure to a pesticide, Pratt 505K, which the defendant had sprayed to kill mosquitos. Various chemical companies that manufactured the ingredients of Pratt 505K were joined as third-party defendants. Since there were no epidemiological studies of Pratt 505K or its constituents, the plaintiffs' experts relied on less telling evidence: case reports, structural similarities between the chemicals involved and others that were known to be toxic, in vitro studies and animal studies. The defense experts denied that these were valid bases for reaching a conclusion on causation. The jury, ultimately, found for the defense.

There was no evidentiary basis for excluding the plaintiffs' evidence in *Vann*. It was, in fact "standard toxicological evidence." At the same time, there was no scientific basis for deciding whether or not Pratt 505K causes aplastic anemia: the available evidence was sufficient to pose the question, but not to answer it. Worse yet (from a scientific point of view) the problem is too uncommon to ever justify the time and expense it would take to find out. Courts could resolve this type of difficulty by requiring epidemiological evidence in every toxic tort case. But that requirement would virtually foreclose an entire class of cases—individual claims of uncommon types of injuries from exposure to allegedly toxic substances. "There are over 100,000 synthetic chemical used in the United States. Only a very small fraction have been subject to epidemiological investigation." Short of such a drastic step, the only solution is to let the trier of fact—the jury—decide the merits of the competing scientific claims.

In other words, there is no procedural short cut. In *Vann*, and in many thousands of other law suits, courts must resolve scientific or technical disputes, the significance of which is limited to the case at hand or to a small set of related incidents. The issue may be the toxicity of a pesticide, or the safety of a highway exit ramp, or the cause of an individual's brain tumor, or the interpretation of fiber-match evidence. Ultimately such questions are resolved by that notorious embarrassment of adversarial fact finding, the battle of the experts.

Criticisms of this method are legion. They have been stated, forcefully, for well over a hundred years. I will not even attempt to summarize them in this paper. For present purposes, Professor Brennan's comment on the evidence in *Vann* will suffice:

Plaintiffs had every incentive to emphasize the importance of case studies. Defense witnesses in turn had the same incentives to deny the importance of animals studies and short-term tests. Thus, the court never heard an even-handed account of toxicology. 166

The common solution proposed for this problem is to call non-partisan court-appointed experts to testify in addition to the parties' witnesses. ¹⁶⁷ I have written about this issue in detail elsewhere; ¹⁶⁸ I will merely state my main conclusions: Court appointment of expert witnesses is a widely praised widely available procedure that is almost never used. The best hope for improving our use of scientific evidence, and of expert evidence in general, is to devise procedures that succeed in encouraging judges to make wider use of this option.

Non-partisan experts might be as useful in cases in which courts engage in legislative fact finding as they would be in trials. The ultimate judgment on Bendectin was clear, but it might have been reached more easily and quickly if some judge along the line—or several judges—had appointed a qualified epidemiologist to provide a non-partisan appraisal of the evidence. Apparently that never happened. In harder cases, the value of non-partisan expertise is greater. While Judge Weinstein may have done an excellent job in the Agent Orange opt-out case, his method, as Professor Schuck has pointed out, was chancy. Despite all his hard work, "he remained self-taught and incompletely informed, lacking the intuition and finely honed technical judgment of the experienced scientist." Wouldn't his judgment have been safer if he had appointed an experienced scientist to advise him? 171

The same applies to a determination of the "general acceptance" or the "validity" of a scientific theory or technique as the basis for a ruling on the admissibility under Frye or Daubert. Christopherson v. Allied Signal Corp. 172—yet another toxic tort case—is a good example. The plaintiff in Christopherson claimed that her deceased husband contracted small-cell colon cancer—a rare disease—as a result of work-place exposure to nickel and cadmium. In the absence of epidemiological evidence, the plaintiff's expert—Dr. Miller—based his conclusion on the following reasoning: Nickel and cadmium are known to cause small-cell carcinomas in other sites in the human body, especially the lungs. Small-cell carcinomas have the same histology—that is, they appear to be made up of the same cells—regardless of their location. Therefore, it is likely that these metals also cause small-cell carcinomas in the colon, and—given the evidence on exposure and on other possible causes—it is likely that they did so in this individual.

The trial court granted the defendant's motion for summary judgment, concluding, among other things:

Dr. Miller's conclusion that a small cell carcinoma of the lung is likely to be associated with a small cell carcinoma located elsewhere in the body is "without precedent in cancer epidemiology and is not scientifically correct." ¹⁷³

The Fifth Circuit, en banc, held that "[t]his finding of what is a scientifically correct conclusion is not for the district court." Nonetheless,

the court affirmed (over a bitter dissent¹⁷⁵) on the ground that this impermissible judgment amounted to an implicit finding that Dr. Miller's methodology failed *Frye's* general acceptance test. The same result could just as easily have been reached under *Daubert* if the district court, or the Fifth Circuit, had concluded that Dr. Miler's reasoning was not scientifically valid.

Was the district court judge right? The only evidence before him were affidavits from partisan experts; faced with a conflict he "simply chose sides in this battle of the experts" and accepted the defendants' theory of science. On what basis? Was it because there were several defense experts but only one for the plaintiff? Or because the defense experts had flashier credentials, or wrote better prose? Or was it because the judge understood the scientific issues and accurately chose the correct position?

As a matter of science, the judge in *Christopherson* may have made the right choice, ¹⁷⁷ but the process by which he reached his decision hardly inspires confidence. Judges, no less than jurors, do get such questions wrong. ¹⁷⁸ If a judge is going to "usurp[] the role of the jury in evaluating the evidence" on a difficult scientific issue, he ought to have some basis for his conclusion beyond the statements of experts hired by the winning side. Of course, if non-partisan expert evidence *were* available, the temptation to take the issue from the jury in the first place would have been greatly reduced. ¹⁸⁰

One last note. Critics of the use of court-appointed experts - almost invariably, trial lawyers - argue that this procedure increases the power of judges alarmingly: "Trial by jury . . . becomes no more than an empty illusion, a shibboleth, to which lip service is paid while its destruction is endorsed."181 I doubt if this argument impresses many judges, but for one reason or another they do go along with the conclusion: they almost never appoint expert witnesses. But this hardly represents a conservative position on the power of judges. They won't seek independent expert advice, but they will exclude the core of a party's case on the basis of their lay assessment of its scientific merit. They won't appoint non-partisan experts to provide information to their jurors, but they will dismiss an entire class of claims as unsound. At least in the context of toxic torts, we have come to accept judicial responses to scientific problems that are both more drastic than the appointment of a witness, and less effective. Whatever the issue-admissibility or sufficiency or liability-the solution to the problem of troublesome scientific evidence is more information, not less, but from a trustworthy source.

Appendix

In the text I predict that if *Daubert* has any systematic effect on the treatment of scientific evidence, it will encourage courts to exclude more

often than they had done under *Frye*. It's too early to reach any solid conclusion, but the first year-and-a-half's worth of federal cases applying *Daubert* are consistent with that prediction.

For the most part, these post-Daubert decisions are unsurprising. There has been a shift, conspicuously in the types of the cases that raise questions concerning scientific evidence. Cases applying Frye were usually criminal—until the last several years almost exclusively so (see supra note 46 and accompanying text)—but most of the cases applying Daubert are civil. Moreover, there are signs of instability among the civil cases, especially on appeal. In this respect, Daubert seems to have accelerated a trend that was already underway.

Criminal Cases

Among reported decisions in federal criminal cases, Daubert has changed almost nothing. The best defined subgroup consists of cases on the admissibility of DNA identification evidence. Like all appellate opinions on point, they let it in. United States v. Martinez, 3 F.3d 1191, 1198 (8th Cir. 1993); United States v. Bonds, 12 F.3d 540, 554 (6th Cir. 1993); United States v. Chischilly, 30 F.3d 1144, 1156 (9th Cir. 1994); United States v. Davis, 40 F.3d 1069, 1075 (10th Cir. 1994); Gov't of the Virgin Islands v. Penn, 838 F.Supp. 1054, 1073 (D. Virg. Is. 1993). Another set of non-path breaking cases deal with a variety of chemical tests in drug prosecutions. As before Daubert, their admission, too, is upheld on appeal. United States v. Bynum, 3 F.3d 769, 773 (4th Cir. 1993)(chromatographic analysis is admissible to show that three samples of crack cocaine came from the same batch); United States v. McCaskey, 9 F.3d 368, 380 (5th Cir. 1993)(evidence of Fourier Transform Infrared Spectrophotometer test for cocaine base admissible at sentencing hearing); United States v. Muldrow, 19 F.3d 1332, 1337 (10th Cir. 1994)(testimony by forensic chemist on tests to identify cocaine admissible); United States v. Harris, No. 93-5943, 1994 U.S. App. LEXIS 20145 at *5 (4th Cir. 1994)(evidence of field test for cocaine admissible). Cf. United States v. Lee, 25 F.3d 997, 999 (11th Cir. 1994)(remand to determine the admissibility under Daubert of trace evidence of narcotics collected via the Sentor and Ionscan machines).

Although *Daubert* speaks directly only to the admissibility of expert scientific evidence, several cases cite the standards announced in *Daubert* as the basis for approving admission of other types of expert testimony: United States v. Markum, 4 F.3d 891, 896 (10th Cir. 1993)(expert testimony by fire chief that fire was caused by arson); United States v. Locascio, 6 F.3d 924, 938–939 (2nd Cir. 1993)(expert testimony by a police officer on the workings of the underworld); United States v. Muldrow, supra 19 F.3d at 1337–1338 (police officer's testimony on whether possession of one kilo of cocaine would be for personal use or distribution); United States v. Johnson, 28 F.3d 1487, 1497 (8th Cir. 1994)(expert testimony by

unindicted coconspirator on gang structure and drug trafficking). See also *United States v. Sepulveda*, 15 F.3d 1161, 1183 (1st Cir. 1993)(trial court's exclusion of *improper* expert opinion testimony by a police commander on defendant's roles in cocaine trafficking organization did not necessitate a mistrial). None of these decisions would likely have been different two years earlier.

Various types of expert scientific evidence that have traditionally been considered unreliable were still excluded after *Daubert*, at least when offered by criminal defendants: polygraph evidence, *United States v. Black*, 831 F. Supp. 120, 122 (E.D.N.Y. 1993); forensic anthropology in a robbery case that turned on identification, *United States v. Dorsey*, 45 F.3d 809, 814–816 (4th Cir. 1995); evidence that the defendant in an attempted extortion case had had "dependant personality disorder," *United States v. Marsh*, 26 F.3d 1496, 1502–1503 (9th Cir. 1994); and expert voice identification, *United States v. Jones*, 24 F.3d 1177, 1180 (9th Cir. 1994)(although the witness was allowed to give *lay* opinion evidence on the same issue). By contrast, the admission of testimony by a prosecution expert on a photogrammetry process for calculating the height of an individual from surveillance photographs was affirmed, on the basis of the trial court's finding that the calculations were neither novel nor controversial. *United States v. Quinn*, 18 F.3d 1461, 1465 (9th Cir. 1994).

The only post-Daubert criminal cases that were remanded on appeal dealt with expert evidence on the reliability of eyewitness identification testimony. In United States v. Amador-Galvan, 9 F.3d 1414, 1418 (9th Cir. 1993), the Ninth Circuit remanded for a trial-court determination on whether the evidence qualified for admission under Daubert. And in United States v. Minnis, No. 93-50330, 1994 U.S. App. LEXIS 14727 at *4 (9th Cir. 1994), the same court followed Amador-Galvan and ordered a similar evidentiary hearing. Between these two decisions, in United States v. Rincon, 28 F.3d 921, 923 (9th Cir. 1994), the Ninth Circuit affirmed the decision of a different district court that such evidence was inadmissible under the same standard. It's hard to see any trend here-in Rincon the court emphasized that it was not setting a general rule, and that other litigants may meet the criteria for admission for similar evidence. Daubert seems to have been an occasion for the Ninth Circuit to make the undramatic statement that at least in some types of cases, trial judges are required to consider admitting expert evidence on evewitness identifications, but having done so in the approved manner they can admit or exclude the evidence as they please.

Civil Cases

A sizeable minority of the civil post-Daubert cases concern medical evidence. Two are Bendectin cases: Elkins v. Richardson-Merrell, Inc., 8 F.3d 1068, 1071 (6th Cir. 1993) in which the Sixth Circuit followed Turpin,

supra note 20, and held the plaintiff's evidence admissible but legally insufficient; and Daubert itself on remand, 43 F.3d 1311 (9th Cir. 1995), in which the Ninth circuit held the same evidence inadmissible under the new standard. See supra note 60. In both Elkins and Daubert the plaintiff claimed that the defendants' affidavits were inadmissible under Daubert; in both the circuit court responded that this argument misconstrued the defendant's burden on summary judgment.

Most of the rest of the medical-evidence cases involve other toxic tort claims, and in most of these the scientific evidence was excluded or its exclusion was upheld: Porter v. Whitehall Laboratories, 9 F.3d 607, 616 (7th Cir. 1993)(summary judgment to defendant ibuprofen manufacturers affirmed); O'Conner v. Commonwealth Edison Co., 13 F.3d 1090, 1107 (7th Cir. 1994)(summary judgment affirmed in radiation exposure case); Claar v. Burlington N. R.R. Co., 29 F.3d 499, 503 (9th Cir. 1994)(summary judgment for defendant affirmed on complaint alleging various injuries from exposure to various chemicals); Sorensen v. Shaklee Corp., 31 F.3d 638, 650 (8th Cir. 1994)(summary judgment for defendant in case alleging that plaintiffs' birth defects were caused by their parents' consumption of alfalfa health food tablets affirmed); Bradley v. Brown, 42 F.3d 434, 438 (7th Cir. 1994)(exclusion at bench trial of evidence on Multiple Chemical Sensitivity); Hayes v. Raytheon Co., No. 92-4004, 1994 U.S. App. LEXIS 8415 at *17, (7th Cir. 1994)(summary judgment for defendant upheld in case alleging that video display terminals caused cervical cancer); Chikovsky v. Ortho Pharmaceutical Corp., 832 F.Supp. 341, 346 (S.D. Fla. 1993)(summary judgment to manufacturer of acne medication Retin-A in birth defect case). See also Wheat v. Pfizer, Inc., 31 F.3d 340, 343 (5th Cir. 1994)(dictum that testimony linking defendant's drug to plaintiff's hepatitis would be inadmissible under Daubert; summary judgment for defendant affirmed on other grounds). On the other hand, in a few post-Daubert federal toxic tort cases, the scientific evidence was admitted at trial or on appeal: Cantrell v. GAF Corp., 999 F.2d 1007, 1014 (6th Cir. 1993)(admission of evidence linking laryngeal cancer to asbestos affirmed); Glaser v. Thompson Medical Co., Inc., 32 F.3d 969, 975 (6th Cir. 1994)(exclusion of evidence that Dexatrim can cause hypertension and summary judgment for defendant reversed); Hopkins v. Dow Corning Corp., 33 F.3d 1116, 1125 (9th Cir. 1994)(admission of plaintiff's expert evidence affirmed in breast implant case).

As a group, the post-Daubert toxic tort cases suggest that (at least in this context) federal judges are, if anything, more willing than before to scrutinize and exclude scientific evidence that they see as weak—e.g., evidence that is based on "common sense" with no published support, Chikovsky, supra., 823 F.Supp. at 345–46, or evidence from a witness who represents a "lone voice" in the scientific community, O'Conner, supra., 13 F.3d at 1105–1106. The two non-toxic tort medical-evidence cases add little. Both are medical malpractice cases, and in both the admission

of the disputed testimony was affirmed: Cella v. United States, 998 F.2d 418, 425 (7th Cir. 1993)(plaintiff's medical expert allowed to testify to cause of polymyositis); Carroll v. Morgan, 17 F.3d 787, 790 (5th Cir.), reh'g denied 26 F.3d 1117 (1994)(defendant's expert cardiologist allowed to testify to cause of death).

The non-medical post-Daubert civil cases are an extremely assorted group. In most the court excluded or approved the exclusion of testimony by experts whose qualifications or data were questionable: Joy v. Bell Helicopter Textron, Inc., 999 F.2d 549, 569 (D.C. Cir. 1993)(economist's testimony on future earning capacity of deceased plaintiff if he had survived); Frymire-Brinati v. KPMG Peat Marwick, 2 F.3d 183, 186 (7th Cir. 1993)(accountant's valuation of investments as worthless); Rosado v. Deters, 5 F.3d 119, 124 (5th Cir. 1993)(accident reconstructionist's tesimony about motorcycle accident); Wilson v. City of Chicago, 6 F.3d 1233, 1238 (7th Cir. 1993)(pathologist's testimony that plaintiff's story of electroshock torture was credible); Marcel v. Placid Oil Co., 11 F.3d 563, 567 (5th Cir. 1994)(economist's testimony on life expectancy of oil rig workers); Berry v. City of Detroit, 25 F.3d 1324, 1354 (6th Cir. 1994)(expert testimony that failure to discipline police officers caused defendant officer to shoot plaintiff); Habecker v. Clark Equip. Co., 36 F.3d 278, 290 (3d Cir. 1994)(expert testimony on simulation on forklift accident); American and Foreign Ins. Co. v. General Elec. Co., 45 F.3d 135, 139 (6th Cir. 1995)(affidavit of engineer regarding design of circuit breaker); Stanczyk v. Black and Decker, Inc., 836 F. Supp. 565, 568 (N.D. Ill. 1993) (feasibility of safer rotary saw design); McLendon v. Georgia Kaolin Co., 841 F.Supp. 415, 418-419 (M.D. Ga. 1994)(geologist's testimony on Kaolin deposits); Israel Travel Advisory Serv., Inc. v. Israel Identity Tours, Inc., No. 92 C 2379, 1993 U.S. Dist. LEXIS 13749 at *5, 1993 WL 387346 at *6 (N.D. III. 1993)(economic testimony on damages to a business); Liu v. Korean Air Lines Co., No. 84 Civ. 0690 (PLN), U.S. Dist. LEXIS 16233, at *1-11, 1993 WL 478343 at *1-10 (S.D.N.Y. 1993)(economist's testimony on future earnings of deceased plaintiff, expert restricted but not excluded); Doe v. Tag, Inc., No. 92 C 7661, 1993 U.S. Dist. LEXIS 16356, at *3-7, 1993 WL 484212 at *3-7 (N.D. Ill. 1993)(expert testimony by social workers and economist excluded, but evidence from MD's admitted, in civil child abuse case).

On the other side, however, there are several cases in which evidence on similar issues was admitted, or its exclusion was reversed on appeal: Petruzzi's IGA Supermarkets, Inc. v. Darling-Delaware Co., 998 F.2d 1224, 1240–1241 (3d. Cir. 1993)(affidavits by economists in antitrust case); United States v. Deccarett, 6 F.3d 37, 58 (2d Cir. 1993)(expert testimony by DEA agent on money laundering, in civil forfeiture case); Iacobelli Constr., Inc. v. County of Monroe, 32 F.3d 19, 24–25 (2nd Cir. 1994)(affidavits by experts on underground construction in contract case); Pioneer Hi-Bred Int'l v. Holden Found. Seeds, Inc., 35 F.3d 1226, 1229 (8th Cir.

1994)(electrophoresis, liquid chromatography, and growout testing in suit alleging misappropriation of the genetic make-up of seed corn); *Auvil v. CBS "60 Minutes"*, 836 F. Supp. 741, 741 (E.D. Wash. 1993)(evidence on toxicity of apples treated with daminozide).

There is a clear pattern across this entire set of civil post-Daubert cases. Almost all the action involves plaintiffs' witnesses. For example, seven of these cases report trial-court decisions: Chikovsky, McLendon, Liu, Stanczyk, Israel Travel Advisory Service, Doe and Auvil. As one might expect, all but one (Auvil) are cases in which purportedly scientific evidence was excluded (although in two of those, Liu and Doe, some such evidence was admitted as well). What's less predictable and more striking is that in every case in which evidence was excluded it was offered by a plaintiff, while in the only reported trial court opinion in which the scientific evidence was admitted it was offered by the defendant.

The largest well-defined cluster of appellate opinions consists of six toxic tort cases in which the plaintiffs' scientific evidence was excluded in the trial court, and the resulting summary judgment for the defendant was affirmed: Porter, O'Conner, Hayes, Sorenson, Claer, and Daubert on remand. If the original Ninth Circuit opinion in Daubert was written in part to encourage trial courts to clamp down on plaintiffs' experts in toxic tort cases, its reversal by the Supreme Court opinion has not stopped federal judges from doing just that. There are no post-Daubert cases of any sort that involve summary judgment for the plaintiff.

There are eleven opinions in this set that affirm judgments after trial: six for plaintiffs and five for defendants. With two exceptions, Marcel and Carroll, these cases are similar in a telling way: they all concern plaintiffs' witnesses. (Indeed, Marcel, Carroll and Auvil are the only three civil post-Daubert cases that involve defense witnesses, regardless of the issue or the outcome.) When a plaintiff's judgment is affirmed, the circuit court opinion upholds the district court's decision to admit the plaintiff's scientific evidence (Pioneer Hi-Bred Int'l, Daccarett, Cantrell, Hopkins, Cella); when a defense judgment is affirmed, the appellate court affirms the decision to exclude the plaintiff's evidence (Habecker, Rosado, American and Foreign Ins. Co., Chicovsky).

The most interesting pattern of all shows up in the seven cases in which trial courts were reversed. They fall into two groups: four cases in which a jury verdict for the plaintiff was reversed because of improper admission of scientific evidence (Frymire-Brinati, Wilson, Berry, Joy), and three in which summary judgment for the defendant was reversed because of improper exclusion of plaintiff's expert affidavits (Glasser, Iacobelli Construction, Petruzzi's IGA Supermarkets). In other words, sometimes the trial courts are too lax and let plaintiffs introduce scientific evidence that should be kept out; when that happens, judgments may have to be reversed. Other times they go too far in the other direction and exclude evidence too quickly, on summary judgment; then cases have to be sent

back to do things the right way. One way or another, judging from these early reversals under *Daubert*, the issue of the moment is what to do about scientific experts who testify for plaintiffs.

Notes

- 1. 113 S. Ct. 2786 (1993).
- 2. The name and the corporate identity of the manufacturer of Bendectin both changed a couple of times between 1956 and 1983, and a couple more times since. Joseph Sanders, The Bendectin Litigation: A Case Study in the Life Cycle of Mass Torts, 43 Hastings L.J. 301, 311-12 (1992). Following Sanders's lead, I will refer to the company throughout this article simply as "Merrell."
- Michael D. Green, Expert Witnesses and Sufficiency of Evidence in Toxic Substances Litigation: The Legacy of Agent Orange and Bendectin Litigation, 86 Nw. U.L. Rev. 643, 661 (1992).
- 4. Mekdeci v. Merrell Nat'l Lab., 711 F.2d 1510 (11th Cir. 1983).
- Sanders, supra note 2, at 396–401, 410–418; Sanders, From Science to Evidence: The Testimony on Causation in the Bendectin Cases, 46 Stan. L. Rev. 1, 4–8 (1993).
- See Michael D. Green, The Paradox of Statutes of Limitations in Toxic Substances Litigation, 76 Calif.L.Rev. 965, 973 (1988) (asbestos); Sindell v. Abbot Labs., 26 Cal.3d 588, 594, 607 P.2d 924, 925 (1980) (DES).
- 7. Sanders, supra note 2, at 331-348.
- 8. Id. at 347-48.
- 9. Sanders, supra note 2. See also Sanders, supra note 5; Green, supra note 3.
- See 506 A.2d 1100,1113 (D.C. App., 1986), appeal after remand 563 A.2d
 (D.C. App. 1989), cert. denied, 493 U.S. 1074 (1990), appeal after remand
 A.2d 1023 (D.C. App. 1991).
- 11. See In re Richardson-Merrell, Inc. "Bendectin" Products Liability Litigation, 624 F.Supp 1212, 1216 (S.D. Ohio, 1985).
- 12. Id.
- 13. In re Bendectin Litigation, 857 F.2d 290 (6th Cir. 1988).
- Sanders, supra note 5, at 6. I have excluded three European trials listed by Sanders and included one more recent trial: Blum v. Merrell Dow Pharmaceuticals, Inc., Pa. Ct. Common Pleas, Phila. County No. 1C27 (discussed in Products Liability Daily, June 21, 1994, p.3).
- Richardson v. Richardson-Merrell, Inc., 857 F.2d 823 (D.C. Cir. 1987), cert. denied, 493 U.S. 882 (1989); Brock v. Merrell Dow Pharmaceuticals, 874 F.2d 307 (5th Cir.), modified 884 F.2d 166 (5th Cir. 1989), cert. denied, 494 U.S. 1046 (1990); Ealy v. Richardson-Merrell, Inc., 987 F.2d 1159 (D.C. Cir.), cert. denied, 498 U.S. 950 (1990).
- Raynor v. Richardson-Merrell, Inc., No. 93–7109, 1993 U.S. App. LEXIS 35818 (D.C. Cir. 1993).
- Lynch v. Merrell Dow Laboratories, 838 F.2d 1190 (1st Cir. 1987); Richardson, supra note 15.

- 18. Brock, supra note 16.
- 19. Turpin v. Merrell Dow Pharmaceuticals, 959 F.2d 1349 (6th Cir. 1992).
- 20. E.g., Hull v. Merrell Dow Pharmaceuticals, 700 F.Supp. 28 (S.D. Fla. 1988); Daubert v. Merrell Dow Pharmaceuticals, 727 F.Supp. 570 (S.D. Cal. 1989), aff'd, 951 F.2d 1128 (9th Cir. 1991); Turpin v. Merrell Dow Pharmaceuticals, 736 F.Supp. 737 (E.D. Ken. 1990), aff'd, 959 F.2d 1349 (6th Cir. 1992); DeLuca v. Merrell Dow Pharmaceuticals, 131 F.R.D. 71 (D.N.J. 1990), rev'd, 911 F.2d 942 (3d Cir. 1990), summ. judgment granted on remand, 791 F.Supp. 1042 (D.N.J. 1992), aff'd without opinion, 6 F.3d 778 (3rd Cir. 1993); Lee v. Richardson Merrell, 772 F.Supp. 1027 (W.D. Tenn. 1991), aff'd, 1992 U.S.App. LEXIS 8478 (6th Cir. 1992).
- 21. Sanders, supra note 5, at 6.
- 22. Daubert, supra note 20.
- 23. Sanders, supra note 2, at 321-331; Green, supra note 3, at 644-658.
- 24. Sanders, supra note 2, at 395.
- 25. Id. at 430.
- 26. Daubert, 727 F.Supp. at 575.
- 27. Id. at 576.
- 28. Id. at 575.
- 29. Judge Gilliam also disposed peremptorily of another expert's "unsupported allegation" that the data of a single published study show a statistically significant teratogenic link. Id. at 576.
- 30. 951 F.2d 1128 (9th Cir. 1991).
- 31. 293 F. 1013, 1014 (D.C. Cir. 1923).
- 32. Id. at 1014.
- 33. 911 F.2d 941, 955 (3rd Cir. 1990).
- 34. 951 F.2d at 1130.
- 35. As applied by the Ninth Circuit in *United States v. Solomon*, 753 F.2d 1522, 1626 (9th Cir. 1985) (cited in *Daubert*, 951 F.2d at 1129–30).
- 36. 951 F.2d at 1131.
- 37. 951 F.2d at 1131 (quoting Peter Huber, Galileo's Revenge: Junk Science in the Courtroom 228 (1991)).
- 38. The issue is discussed in detail in several briefs amicus curiae that were filed in Daubert by organizations and individuals with deep experience in the process of scientific peer review: Brief Amici Curiae of Physicians, Scientists and Historians of Science in Support of Petitioners; Brief Amici Curiae of American Society of Law, Medicine and Ethics, et al., in Support of Petitioners; Brief Amici Curiae of Daryl E. Chubin, et al., in Support of Petitioners; Brief of the Carnegie Commission on Science, Technology, and Government as Amici Curiae in Support of Neither Party; Brief of the New England Journal of Medicine, et al., as Amici Curiae in Support of Respondent; Brief of the American Medical Association, et al., as Amici Curiae in Support of Respondents; Brief Amicus Curiae of the American College of Legal Medicine in Support of the Respondent; Brief of Amici Curiae of Nicolaas Bloembergen, et al., in Support of Respondents; Brief Amicus Curiae of Professor Alvan Feinstein in Support of Respondent; Brief for the American Association for the Advancement of Science, et al., as Amici Curiae in Support of Respondent. Significantly, while several amici

- wrote in favor of peer review in general, none endorsed it as an all-purpose test for the admissibility of scientific evidence.
- 39. 113 S. Ct. 2786 (1993).
- 40. The Court also sidestepped another argument by the Petitioners: that Frye, as a federal judge-made rule could not, under Erie R. v. Tompkins, 304 U.S. 64 (1938), be applied in a federal diversity law suit (such as Daubert). 113 S. Ct. at 2794 n.6. This issue had not surfaced before Daubert, no doubt because Frye had been applied almost exclusively in criminal cases. See infra note 46 and accompanying text.
- 41. See Edward J. Imwinkelreid, The Standard for Admitting Scientific Evidence: A Critique From the Perspective of Juror Psychology, 28 Vill. L. Rev. 554, 556–57 (1982–83): "Frye was not only the majority view among American courts; it was the almost universal view, with the overwhelming majority of federal and state courts following it. Indeed, at one point in the mid-1970's, Frye seemed to be the controlling test in at least forty-five states." See also Note, Changing the Standards for the Admissibility of Novel Scientific Evidence: State v. Williams, 40 Ohio St. L.J. 757, 769 (1979).
- See, e.g., Daubert v. Merrell Dow Pharmaceuticals, Inc., 951 F.2d 1128 (9th Cir. 1991); Christophersen v. Allied-Signal Corp., 939 F.2d 1106, 1110–1111, 1115–1116 (5th Cir. 1991) (en banc), cert. denied, 112 S.Ct. 1280 (1992); United States v. Smith, 869 F.2d 348 351 (7th Cir. 1989); United States v. Shorter, 809 F.2d 54, 59–60 (D.C. Cir.), cert. denied, 484 U.S. 817 (1987); United States v. Metzger, 778 F.2d 1195, 1203 (6th Cir. 1985), cert. denied, 477 U.S. 906 (1986); United States v. Smith, 776 F.2d 892, 898 (10th Cir. 1985); United States v. Alexander, 526 F.2d 161, 163–164 (8th Cir. 1975).
- See, e.g., United States v. Williams, 583 F.2d 1194 (2d Cir. 1978); United States v. Piccinonna, 885 F.2d 1529, 1536–1537 (11th Cir. 1989); United States v. Downing, 753 F.2d 1224, 1237–1240 (3d Cir. 1985); United States v. Baller, 519 F.2d 463, 465–466 (4th Cir.), cert. denied, 423 U.S. 1019 (1975).
- 44. In deVries v. St. Paul Five and Marine Insurance Company, 716 F.2d 939, 945 (1st Cir. 1983), the First Circuit cited Frye with approval, but on the narrow issue of the unreliability of polygraph tests.
- 45. The status of Frye in state courts was always a tricky issue, since many state followed Frye with modifications or limitations, and some adopted a "general acceptance" test without citing Frye. See Paul C. Giannelli, The Admissibility of Novel Scientific Evidence: Frye v. United States, A Half-Century Later, 80 Colum. L. Rev. 1197, 1128-29 (1980). In Daubert, the States of Texas, Montana, Idaho, and South Dakota stated in a brief amicus curiae that Frye had been rejected at least as an exclusive test in 25 states; for 18 of the 25 states they list, the opinion cited is from 1987 or later. Brief Amici Curiae of the States of Texas, et al., in Support of Petitioners, at 6-7.
- See id.; Paul C. Giannelli, "Junk Science": The Criminal Cases, 84 J. Crim. L. & Criminology 105 111 (1993).
- 47. The two earlier cases were Barrel of Fun, Inc. v. State Farm Fire and Cas. Co., 739 F.2d 1028, 1031 (5th Cir. 1984); and Christophersen v. Allied Signal Corp., 939 F.2d 1106, 1115–16 (5th Cir. 1991) (per curiam) (en banc).
- See, e.g., Giannelli, supra note 45, and authorities cited therein at 1206 n.
 Hanson, James Alphonso Frye is Sixty-Five Years Old; Should He Retire?,
 W. St. U.L.Rev. 357 (1989).

- 49. 113 S. Ct. at 794. Nor does Rule 703 help preserve the *Frye* test. Rule 703 permits an expert to rely on *inadmissible* evidence if it is "of a type reasonably relied upon by experts in the particular field in forming opinions or inferences upon the subject." It does nothing to limit the *admissible* bases for expert opinion evidence. Assuming that "reasonable reliance" is equivalent to "general acceptance," this rule would permit an expert witness to rely on a systolic blood pressure deception test as the basis for an opinion on a witness's credibility (assuming the evidence was otherwise admissible) if she had performed it herself (and could therefore testify from personal knowledge), but not if an associate conducted the test and informed her of the results (which would be hearsay)—unless the test were "of a type reasonably relied upon," etc. Obviously, this is *not* the *Frye* test.
- 50. The procedure that governs these preliminary judicial findings is described in Federal Rule of Evidence 104(a).
- 51. The rejected *Frye* test (as we have noted) was limited to evidence of *novel* scientific principles and techniques. Rule 702 is not, and therefore neither is the Court's interpretation of that rule. However, the Court recognizes that for "well established propositions" it will be easy to satisfy these requirements, and some may be subject to judicial notice. 113 S. Ct. at 2796 n.11.
- 52. Id. at 2796 (footnote omitted) (emphasis added).
- 53. Id. (emphasis added).
- 54. Id. at 2797.
- 55. Id.
- 56. Id.
- 57. Id.
- 58. Id.
- 59. Id.
- 60. Id. On remand from the Supreme Court, the Ninth Circuit reaffirmed its earlier judgment. Daubert v. Merrell Dow Pharmaceuticals, 43 F.3d 1311 (9th Cir. 1995). The court (in another opinion by Judge Kozinski) found the plaintiffs' evidence inadmissible under the Supreme Court opinion in Daubert, in part because there was insufficient evidence to support a finding that it constituted reliable "scientific knowledge," id. at 1316–21, and in part because in the absence of any claim that Bendectin more than doubled the risk of limb-reduction birth defects, the evidence could not show that it "caused" the plaintiffs' injuries by the standard that applies under California law, and therefore the evidence would not "assist the trier of fact," id. at 1321–22.
- 61. See supra note 45 and accompanying text.
- 62. Cleary et al., McCormick on Evidence §206 at 912-14 (4th ed. 1992).
- 63. State v. Dorsey, 88 N.M. 184, 539 P.2d 204 (1975).
- State v. Lindemuth, 56 N.M. 257, 243 P.2d 325 (1952); see also Montoya v. Metropolitan Court, 98 N.M. 616, 651 P.2d 1260 (1982), overruled, State v. Alberico, 116 N.M. 156, 861 P.2d 192 (1993).
- 65. State v. Sims, 52 Ohio Misc. 31, 369 N.E.2d 24 (1977), overruled sub silentio, State v. Souel, 53 Ohio St.2d 123, 372 N.E.2d 1318 (1978).
- 66. State v. Williams, 446 N.E.2d 444, 446-48 (Ohio 1983).
- 67. Anno., Modem Status of Rule Relating to Admission of Results of Lie Detector (Polygraph) Test in Federal Criminal Trials, 43 ALR Fed 68, and supp.

- See Margaret A. Berger, Novel Forensic Evidence: The Need for Court-Appointed Experts after Daubert, 1 Shepard's Expert & Scientific Evid. Q. 487, 488 (1994).
- 69. Committee on Evaluation of Sound Spectrograms, National Research Council, On The Theory and Practice of Voice Identification (1979).
- State v. Williams, 388 A.2d 500 (Me. 1978); United States v. Williams, 583
 F.2d 1194 (2d Cir. 1978); State v. Williams, 446 N.E.2d 444, 446–48 (Ohio 1983).
- 71. Anno., Admissibility and Weight of Voiceprint Evidence, 97 ALR 3d 294, and supp.; Berger, supra note 68, at 488.
- 72. Anno., Admissibility of DNA Identification Evidence, 84 ALR 4th 313, and supp. Courts are divided, however, on the admissibility of inferential evidence based on estimates of the frequency of various DNA types in relevant human populations. Compare, e.g., State v. Vandebogart, 136 N.H. 365, 616 A.2d 484, 494 (1992) (excluding population frequency estimates under Frye), with United States v. Yee, 134 F.R.D. 161, 210 (N.D. Ohio 1991) (admitting such evidence, also under Frye). See generally Berger, supra note 68, at 489–90; William C. Thompson & Susan Ford, DNA Testing: Acceptance are Weight of the New Genetic Identity Tests, 75 Va. L. Rev. 45 (1989).
- 73. 753 F.2d 1224 (3rd Cir. 1985).
- 74. The Downing factors are: (1) the "degree of acceptance" of the technique or theory in the relevant scientific community; (2) the "novelty" of technique or theory, and (3) "the existence of a specialized literature" dealing with it, both which bear on (4) the extent to which it "has been exposed to critical scientific scrutiny"; (5) "the qualifications and professional status" of the witnesses offering the evidence; (6) "the non-judicial uses to which the scientific technique are put"; (7) the "frequency with which the technique leads to erroneous results"; and (8) "the type of error" it produces. Id. at 1238–39. The Court in Daubert gave qualified approval to the Downing test, and to various other tests for admissibility proposed by commentators: "all these versions may have some merit, although we express no opinion regarding any of their particular details." 113 S. Ct. at 2797 n.12.
- 75. See DeLuca, supra note 20.
- 76. United States v. Franks, 511 F.2d 25, 33 (6th Cir. 1975).
- 77. 113 S. Ct. at 2796 n. 10 and accompanying text.
- 78. Id. at 2797-98.
- 79. Id. at 2796-97.
- 80. Id. at 2798-99.
- 81. Id. at 2794.
- 82. It's too early to say for sure, but the first batch of federal cases applying Daubert is consistent with this prediction. See Appendix for a summary of such cases through early 1995.
- 83. 830 F.2d 1190 (1st Cir. 1987).
- 84. Id. at 1194 (citation omitted).
- 85. 857 F.2d 823 (D.C. Cir. 1988).
- 86. 736 F.2d 1529 (D.C. Cir. 1984).
- 87. Id. at 1535.
- 88. Id. at 1535.
- 89. Id. at 1534.

- 90. 857 F.2d at 832.
- 91. 897 F.2d 1159 (D.C. Cir. 1990).
- Id. at 1160 (citation to Richardson v. Richardson-Merrell, 857 F.2d 823 (D.C. Cir. 1989) omitted).
- 93. 113 S.Ct. at 2797.
- 94. 874 F.2d 307 (5th Cir. 1989).
- 95. 959 F.2d 1349 (6th Cir. 1992).
- Kevin Clermont & Theodore Eisenberg, Trial by Jury or Judge: Transcending Empiricism, 77 Cornell L. Rev. 1124, 1127 n.7 (1992).
- 97. In re Richardson-Merrell, 624 F.Supp. at 1268.
- 98. E.g., Federal Rule of Evidence 606(b).
- 99. E.g., In re Japanese Electronic Products Antitrust Litigation, 631 F.2d 1069, 1086 (3rd Cir. 1980); Warren Burger, The Use of Lay Jurors in Complicated Civil Cases, Remarks to the Conference of State Chief Justices 3–5 (Aug. 7, 1979); Goodman, Greene & Loftus, What Confuses Jurors in Complex Cases: Judges and Jurors Outline the Problem, Trial, Nov. 1988 at 65. But see In re United States Fin. Sec. Litig., 609 F.2d 411, 429–30 (9th Cir. 1979). See generally Richard O. Lempert, Civil Juries and Complex Cases: Lets Not Rush to Judgment, 80 Mich. L. Rev. 68 (1981); Joe S. Cecil, Valerie P. Hans and Elizabeth C. Wiggins, Citizen Comprehension of Difficult Issues: Lessons From Civil Jury Trials, 40 Am. U.L. Rev. 727 (1991); Richard Lempert, Civil Juries and Complex Cases: Taking Stock after Twelve Years, in Robert E. Litan, Ed. Verdict, Assessing the Civil Jury System, 181–247 (The Brookings Institution 1993).
- 100. E.g., Reid Hastie, Steven Penrod & Nancy Penington, Inside the Jury (1983), The Effects of Death Qualification on Jurors' Predisposition to Convict and on the Quality of Deliberation, 8 L. & Hum. Beh. 53 (1984).
- Phoebe C. Ellsworth, Are Twelve Heads Better Than One?, 52 L. & Contemp. Probs. 205 (1989).
- 102. Sanders, supra note 5, at 6.
- 103. 615 F.Supp. 262 (N.D. Ga. 1985), aff'd and modified in part, 788 F.2d 741 (11th Cir. 1986).
- 104. Samuel R. Gross, Expert Evidence, 1991 Wisc. L. Rev. 1113, 1121-24.
- 105. Wells, 615 F.Supp. at 298. The judgment was reduced on appeal to \$4.7 million. 788 F.2d at 747.
- Smith v. Ortho Pharmaceutical Corporation, 770 F.Supp. 1561 (N.D. Ga.1991).
- 107. Sanders, supra note 5, at 6 (omitting European cases).
- 108. 911 F.2d 941 (3rd Cir. 1990).
- 109. Id. at 944 (quoting Appellant's Brief at 38).
- 110. See Gross, supra note 104, at 1153-58.
- 111. Federal Rule of Evidence 704(a), for example, specifically permits expert testimony that "embraces an ultimate issue to be decided by the trier of fact." (Except for some statements about the mental state of criminal defendants, which are not permitted by Rule 704(b).)
- 112. McCormick, supra note 62, §338 at 434 (footnote omitted). The "rare case" cited is instructive: Scott v. Hansen, 228 Iowa 37, 289 N.W. 710 (1990), in which a witness testified that he saw a cow hit by a car, after which the cow

- "flew pretty near thirty feet without touching the pavement at all; and then she hit the pavement and bounced better than five feet. . . ." Id. at 712.
- 113. Gimbel v. Laramie, 181 Cal. App.2d 77, 80, 5 Cal. Rptr. 88, 90 (1960) (citations omitted).
- 114. Ferebee, supra note 86, 736 F.2d at 1534.
- 115. In Federal diversity cases, treating the problem of questionable scientific evidence as an issue of admissibility has another consequence that judges may like: it enables them to avoid the requirement of Erie R.R. v. Tompkins, 304 U.S. 64 (1938), that in diversity cases federal courts apply the substantive law of the states in which they sit. In Richardson, the plaintiffs argued that Erie meant that the Federal Courts for the District of Columbia were obliged to follow the holding of the District of Columbia Court of Appeals in Oxendine v. Merrell-Dow Pharmaceuticals, Inc., 563 A.2d 330 (D.C. 1989), reversing a j.n.o.v. in a Bendectin case and reinstating the jury's verdict for the plaintiff. The Circuit Court replied: "[T]he Erie doctrine applies to substantive, not procedural, law. . . . [I]t is clear that Erie does not affect the federal courts' application of the Federal Rules of Evidence. . . . Because the admissibility of testimony . . is the crux of our decision in this case . . . the Richardsons' argument must be rejected." See also Ealy v. Richardson-Merrell, Inc., 897 F.2d 1159, 1163 (D.C. Cir. 1990).
- 116. Ealy, 897 F.2d at 1160.
- 117. See, e.g., Federal Rule of Evidence 104(b).
- 118. See, e.g., Federal Rule of Evidence 104.
- 119. See supra note 38 and accompanying text.
- See, e.g., David H. Kaye, Is Proof of Statistical Significance Relevant?, 61 Wash. L. Rev. 1333 (1986).
- 121. See Federal Rule of Evidence 804(b)(2).
- 122. See Federal Rule of Evidence 703.
- 123. Ballew v. Georgia, 435 U.S. 223 (1978).
- 124. U.S. v. Martinez, 3 F.3d 1191 (8th Cir. 1993).
- 125. Kenneth C. Davis, An Approach to Problems of Evidence in the Administrative Process, 55 Harv. L. Rev. 364 (1942).
- 126. E.g., Advisory Committee's Note to Federal Rule of Evidence 201(a).
- 127. Kenneth C. Davis, A System of Judicial Notice Based on Fairness and Convenience, in Perspectives of Law 69, 82 (1964).
- Lockhart v. McCree, 476 U.S. 162, 166 n.3 (1986); Dunagin v. City of Oxford, Miss., 718 F.2d 738, 748 n.3 (5th Cir. 1983).
- In re Asbestos Litigation, 829 F.2d 1233, 1245–51 (Becker, J., concurring) (3rd Cir. 1987); id. at 1256–60 (Hunter, J., dissenting).
- 130. Id., 829 F.2d at 1233, 1245–51, 1256–60 (interpreting Beshada v. Johns-Manville Products Corp., 90 N.J. 191, 447 A.2d 539 (1982)). But see Democratic Party v. Nat. Conservative P.A. Committee, 578 F.Supp. 797, 830–31 (E.D. Pa. 1983), modified sub. nom., Fed. Election Comm. v. Nat. Conservative P.A. Committee, 470 U.S. 480 (1985).
- 131. Kenneth C. Davis, Administrative Law Treatise 353 (1978).
- 132. Kenneth C. Davis, Judicial Notice, 55 Colum. L. Rev. 945, 952 (1955).
- 133. Commissioner of Welfare ex rel. Tyler v. Costonio, 277 App. Div. 90. 92, 97 N.Y.S.2d 804, 806 (1st Dep't 1950) (Shientag, J., concurring). See also Jordan v. Mace, 144 Me. 351, 69 A.2d 760 (1949).

277

- 134. 955 F.2d 786 (2d Cir. 1992).
- 135. Id. at 799-800.
- Harold L. Korn, Law, Fact and Science in the Courts, 66 Colum. L. Rev. 1080, 1102 (1966).
- 137. Id. at 1105.
- 138. James Bradley Thayer, A Preliminary Treatise on Evidence at the Common Law 249 (1898) (quoting Lord Mansfield in Tindfal v. Brown, 1 T.R. 167, 168, 99 Eng. Rep. 1033, 1034 (K.B. 1786)).
- 139. Turpin, note 20, 959 F.2d at 1351.
- 140. Brock, 874 F.2d at 310 (footnotes omitted).
- 141. Id. at 313.
- 142. Turpin, 959 F.2d at 1360.
- 143. E.g., "In light of the evidence presented, we are convinced that the Brocks did not present sufficient evidence regarding causation to allow a trier of fact to make a reasonable inference that Bendectin caused Rachel Brock's limb reduction defect." Brock, 874 F.2d at 315 (emphasis added). "Dr. Palmer [a plaintiff's expert] does not . . . take issue with his peers and explain the grounds for his differences." Turpin, 959 F.2d at 1360.
- 144. See, e.g., Jack H. Friedenthal, Cases on Summary Judgment: Has There Been a Material Change in Standards?, 63 Notre Dame L. Rev. 770, 781 (1988).
- 145. Brock, 884 F.2d at 169 (Higginbotham, J., concurring in the dissent); see also id. at 168 (Reavley, J., dissenting).
- 146. Thus, for example, some courts have interpreted *Brock* as setting a general requirement of statistically significant epidemiological evidence for any toxic tort case. E.g., *Thomas v. Hoffman-La Roche*, 731 F.Supp. 224, 228 (N.D. Miss. 1989), *aff'd*, 949 F.2d 806 (5th Cir. 1992); Porter v. Whitehall Labs, Inc., 791 F.Supp. 1335, 1347 (S.D. Ind. 1992).
- 147. 611 F. Supp.1223 (E.D.N.Y. 1985), aff'd on other grounds, 818 F.2d 187 (2nd Cir. 1987), cert. denied, 487 U.S. 1234 (1988).
- 148. See Peter H. Schuck, Agent Orange on Trial, Mass Toxic Disasters in the Courts 143-167 (1986).
- 149. 611 F. Supp. 1223. Judge Weinstein also granted summary judgment on the ground that the plaintiffs could not, in any event, prove which defendant might have harmed them, 611 F. Supp. at 1263, and on the "government contract defense." Id. at 1263-64. The Second Circuit affirmed on this last ground. 818 F.2d at 194.
- 150. E.g., Schuck, supra note 148; Green, supra note 3; Charles Nesson, Agent Orange Meets the Blue Bus: Factfinding at the Frontiers of Knowledge, 66 B.U.L. Rev. 521 (1986).
- 151. 874 F.2d at 310 (citation omitted).
- 152. 827 F. Supp. 1014 (S.D.N.Y. 1993).
- 153. Id. at 1050.
- 154. Id. at 1038-50.
- 155. Id. at 1050.
- 156. Id. at 1038.
- 157. See Washington v. Armstrong World Indus., 839 F.2d 1121 (5th Cir. 1988); Landrigan v. Celotex Corp., 127 N.J. 404, 605 A.2d 1079 (1992); Caterinicchio v. Pittsburgh Corning Corp., 127 N.J. 428, 605 A.2d 1092 (1992). The Joint Ashestos Case itself concerns a single claim out of a group of 16 that were

- initially jointly prosecuted. See *In re Joint E. and S. Dist. Asbestos Litig.*, 758 F.Supp. 199.200 (S.D.N.Y. 1991).
- 158. See Schuck, *supra* note 148 at 239. ("Weinstein personally read an enormous amount of the technical literature on toxicology, statistics, and other relevant disciplines. . . . Were all judges as incisive as Weinstein, there might be much to be said for this intervention.")
- 159. A very recent study found that exposure to one of two common types of asbestos—amphibole, but not serpentine—"may be associated with colorectal cancer," but this association appears relatively weak and might be caused by "misdiagnosis of lung cancer or other types of cancer in reported causes of death." David M. Homa, David H. Garabrant, & Brenda W. Gillespie, A Meta-Analysis of Colorectal Cancer and Asbestos Exposure, Am. J. Epidemiology (forthcoming). This is not likely to be the last word on the subject.
- 160. No. 84 425 091 NI (Wayne County Cir. Ct. Mich. June 12, 1988).
- Troyen A. Brennan, Helping Courts with Toxic Torts: Some Proposals Regarding Alternative Methods for Presenting and Assessing Scientific Evidence in Common Law Courts, 51 U. Pitt. L. Rev. 1, 44–48 (1989).
- 162. Id. at 45-46.
- 163. Id. at 46.
- 164. Id. at 47, n. 197.
- 165. E.g., Emory Washburn, Testimony of Experts, 1 Am. L. Rev. 45, 48–49 (1866); Rutherford v. Morris, 77 Ill. 397, 405 (1875); Learned Hand, Historical and Practical Considerations Regarding Expert Testimony, 15 Harv. L. Rev. 40 (1901); Mason Ladd, Expert Testimony, 5 Vand. L. Rev. 414 (1952); Jack B. Weinstein, Improving Expert Testimony, 20 U. Rich. L. Rev. 473 (1986).
- 166. Brennan, supra note 161, at 48.
- 167. Id.; see also Washburn, supra note 165; Hand, supra note 165.
- 168. Gross, supra note 104, at 1187-1208, 1220-30.
- 169. Needless to say, experts would have been useful to the jurors for in those Bendectin cases that went to trial. For example, Joseph Sanders has published a fascinating description of the jury deliberations in one Bendectin trial, Havner v. Merrell Dow Pharmaceuticals, Texas Dist. Ct., 214th Fud. Dist. (1991), based on post-deliberation interviews with several jurors. He found a "strong tendency [for the jurors] to conclude that [scientific] opinion is evenly divided" on whether Bendectin is a teratogen. Joseph A. Sanders, Jury Deliberation in a Complex Case: Havner v. Merrell Dow Pharmaceuticals, 16 Justice Sys. J. 45, 64 (1993). Any court-appointed expert would have disabused the jury of that mistake; and yet the defense, which would clearly have benefited, apparently did not seek the appointment of such an expert. Their failure to do so was no aberration. Trial lawyers hardly ever ask to have experts appointed, whether they are likely to gain or lose from such an appointment. See generally Gross, supra note 104, at 1198–1200.
- 170. Schuck, supra note 148, at 239.
- 171. Various commentators have suggested more radical proposals for dealing with scientific issues in mass-tort cases—e.g., the creation of a "Science Court," James A. Martin, *The Proposed "Science Court,"* 75 Mich. L. Rev. 1058 (1977), or a "Science Board" that, among other functions, would appoint panels to resolve important recurring issues. Brennan, *supra* note 161, at 62–71. These might be worthy suggestions, but they require legislative innovations. Almost all American jurisdictions already permit judges to appoint non-partisan

- expert witnesses in any type of case, on motion of a party or on their own motion. See, e.g., Federal Rule of Evidence 706.
- 172. 902 F.2d 302 (5th Cir. 1990), superseded en banc, 939 F.2d 1106 (1991).
- 173. 939 F.2d at 1116. The quotation is from an affidavit submitted by a defense expert; apparently, it is taken out of context. See 902 F.2d at 366 n.3.
- 174. 939 F.2d at 1116.
- 175. Id. at 1124-34 (Reavely, J., dissenting). See also Id. at 1119-20 (Clark, C.J., concurring); 902 F.2d 362 (panel opinion of Reavely, J., reversing).
- 176. 902 F.2d at 366.
- 177. As it happens, I believe the district court was right—but only because I have checked with a non-partisan expert on my own, an Associate Professor of Occupational Medicine at a major university who is an expert on cancer epidemiology. He confirmed that there is no epidemiological evidence linking nickel or cadmium and colon cancer of any sort, and that small-cell colon cancer is rare, but he dismissed the argument that the fact that a substance produces small-cell carcinomas in one organ means its likely to do the same in other organs: "There's no evidence of that. That's just not the way it works." Conversation between the author and Dr. David H. Garabrant, M.D., M.P.H., Associate Professor of Occupational Medicine, University of Michigan Medical School and School of Public Health, 12/12/93. Needless to say, I do not mean to suggest that this informal consultation is a model for the use non-partisan experts in litigation. And for any purpose, if more depended on the issue, I would want more detailed information and explanation—but from the same type of source.
- 178. Wells is a clear example.
- 179. Christopherson, 902 F.2d at 366.
- 180. Court-appointed experts may promote speed and efficiency, as well as accuracy and consistency, since their presence might encourage the parties to settle. Thus, for example, a study by the Federal Judicial Center found that in most of the rare occasions in which court-appointed experts have been used, their main function is to facilitate settlement rather than testify at trial. Joseph Cecil & Thomas A. Willging, *Defining a Role for Court-Appointed Experts*, 4 Fed. Jud. Ctr., Directions 6 (Aug. 1992).
- 181. Elwood S. Levy, *Impartial Medical Testimony Revisited*, 34 Temp. L.Q. 416, 425 (1961).