

## What's Wrong with Litigation-Driven Science? An Essay in Legal Epistemology

Susan Haack\*

If Science, for a consideration, can be induced to prove anything which a party litigant needs in order to sustain his side of the issue, then Science is fairly open to the charge of venality and perjury, rendered the more base by the disguise of natural truth in which she robes herself.

John Ordronaux<sup>1</sup>

### I. INTRODUCTION

Jeremy Bentham's memorable description of "Injustice and her handmaid, Falsehood"<sup>2</sup> should remind us, if we need reminding, that factual truth is an important element of justice,<sup>3</sup> that it really matters whether this witness's recovered memory of an alleged crime is genuine, whether this is the person who committed the crime, whether this plaintiff's injury was caused by a defect in this manufacturer's tire or seat-belt buckle or lawn-chair, whether this was the

---

\* Distinguished Professor in the Humanities, Cooper Senior Scholar in Arts and Sciences, Professor of Philosophy and Professor of Law, University of Miami. This paper was written for and first presented at the March 2006 Coronado III conference on Litigation-Driven Science, organized by the project on Scientific Knowledge and Public Policy (SKAPP). © 2008 Susan Haack & *Seton Hall Law Review*.

My thanks to Mark Migotti for detailed comments on more than one draft; to Robert Lane, and the participants at the Coronado III conference, for their comments on the penultimate version; and to Barbara Brandon, Michelle Cosby, David Hollander, and Helen Wohl, of the University of Miami Law Library, for research assistance.

<sup>1</sup> John Ordronaux, *On Expert Testimony in Judicial Proceedings*, 30 AM. J. INSANITY 312, 312 (1874).

<sup>2</sup> 1 JEREMY BENTHAM, *RATIONALE OF JUDICIAL EVIDENCE* 22 (Garland Publishing Inc. 1978) (1827).

<sup>3</sup> An important element of substantive justice, that is—with which considerations of procedural justice in some instances compete: *e.g.*, when a person who committed a crime goes unpunished, or a person who suffered an injury goes uncompensated, for lack of admissible evidence making the case to the required degree of proof.

chemical exposure that caused or promoted the plaintiff's cancer, and so on.<sup>4</sup>

Because the factual truths at issue in a case often go beyond what the average juror can be expected to know, courts have come increasingly to rely on expert witnesses, among them scientists testifying on just about every subject imaginable: experts on blood, bullets, bite-marks, battered wives; on PCBs, paternity, poisons, post-traumatic stress; on radon, recovered memories, rape trauma syndrome, random-match probabilities; on psychosis, asbestosis, silicosis (and for all I know, on psittacosis!). But as long as courts have relied significantly on scientific witnesses, there have been complaints: about the scientific ignorance and gullibility of attorneys, judges, and jurors; about "witness-shopping"; and—as my opening quotation illustrates—about the irresponsibility and venality of professional scientific experts willing to say whatever is needed to advance the cause of the party that hires them.

As reliance on expert witnesses has grown, so has the felt need for courts to ensure that the expert testimony admitted is not just flimsy or interested speculation, but reliable enough to be more helpful than misleading; and one factor that courts have sometimes taken as indicating that proffered scientific testimony may not be reliable is that it is based on "litigation-driven" science. As it happens, the context in which I first encountered criticisms of proffered scientific testimony as "litigation-driven" was Judge Kozinski's 1995 ruling in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, on remand from the Supreme Court.<sup>5</sup> As I subsequently discovered, however, that expert testimony is based on litigation-driven research has been construed not

---

<sup>4</sup> This principle has, however, sometimes been deliberately sidestepped in civil cases; see, e.g., *Summers v. Tice*, 199 P.2d 1, 3 (Cal. 1948), where the court held that: where a group of persons are on a hunting party, or otherwise engaged in the use of firearms, and two of them are negligent in firing in the direction of a third person who is injured thereby, both of those so firing are liable for the injury . . . although the negligence of only one of them could have caused [it].

*Id.*; *Sindell v. Abbott Labs*, 607 P.2d 924, 937 (Cal. 1980) (holding that "[e]ach defendant will be held liable for the proportion of the judgment represented by its share of that market unless it demonstrates that it could not have made the product which caused plaintiff's injuries"); and, most strikingly, *Hymowitz v. Eli Lilly & Co.*, 539 N.E.2d 1069, 1081 (N.Y. 1989) (holding that "there should be no exculpation for a defendant who, although a member of the market producing DES [diethylstilbestrol] for pregnancy use, appears not to have caused a particular plaintiff's injury").

<sup>5</sup> *Daubert v. Merrell Dow Pharms., Inc. (Daubert IV)*, 43 F.3d 1311, 1317 (9th Cir. 1995).

only as bearing on its admissibility, but also as lowering its weight;<sup>6</sup> and has been construed as bearing on the admissibility of scientific testimony under *Frye v. United States*<sup>7</sup> as well as under *Daubert*.<sup>8</sup>

This prompts a host of questions, legal and epistemological, theoretical and practical. What role has this factor played in courts' handling of scientific testimony? What exactly does it mean to describe research as "litigation-driven"? What reasons have courts given for regarding litigation-driven science with suspicion? Are these reasons sound? And if they are, does this suffice to show that Judge Kozinski's new "*Daubert* factor"—whether the science on which testimony is based is litigation-driven—is a useful indicator of the (un)reliability of proffered expert testimony?

Part II of this Article will look in some detail at two Bendectin cases: *Daubert* itself, which I'm sure I don't need to tell you was a federal case tried first in 1989 under *Frye*, but reheard by the U.S. Court of Appeals of the Ninth Circuit on remand from the Supreme Court of the United States under the new standards of admissibility the Su-

---

<sup>6</sup> See, e.g., *Perry v. United States*, 755 F.2d 888, 892 (1985) ("[Plaintiff's expert witness] Dr. Goldfield had reached a conclusion as to the connection between encephalitis and the [swine flu] vaccine before commencing his research."). The Court cited similar criticism of Dr. Goldfield's testimony in *O'Gara v. United States*, 560 F. Supp. 786, 790 (E.D. Pa. 1983), and in *Robinson v. United States*, 533 F. Supp. 320, 328 (E.D. Mich. 1982).

<sup>7</sup> *Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923); see, e.g., *Lofgren v. Motorola*, No. CV93-05521, 1988 WL 2999925, at \*32 (Ariz. Super. June 1, 1988) (excluding the testimony of plaintiff's expert witness Dr. Kilburn that the injuries were caused by a single exposure to Rubiflex, in part on the grounds that "the conclusion appeared to be more litigation-driven than science oriented").

<sup>8</sup> *Daubert v. Merrell Dow Pharms., Inc. (Daubert III)*, 509 U.S. 579, 594 (1993); see, e.g., *Burleson v. Tex. Dep't of Criminal Justice*, 393 F.3d 577, 584 (5th Cir. 2004) ("Dr. Carlson's 'radiation hot-spot' theory is nothing more than litigation-driven speculation, not science."); *Prohaska v. Sofamor*, 138 F. Supp. 2d 422, 437 (W.D.N.Y. 2001) (Dr. Austin's testimony that pedicle screws manufactured by the defendant were defective excluded because "litigation-driven expertise has been found to be a negative factor in admissibility"); *Downs v. Prestorp Components, Inc.*, 126 F. Supp. 2d 1090, 1094, 1129 (E.D. Tenn. 1999) (excluding Dr. Kilburn's testimony because it "appeared to be more litigation-driven than science oriented" and is "based upon nothing more than conjecture, speculation, and litigation animus") (internal citations omitted); *Mancuso v. Consol. Edison*, 967 F. Supp. 1437, 1445 (S.D.N.Y. 1997) (Dr. Schwartz's testimony that PCB exposure caused Mr. Mancuso's ailments inadmissible because he "rel[ied] upon plaintiff's attorney to provide him with the scientific literature"); *Celotex Corp. v. AIU Ins. Co.*, 196 B.R. 973, 984–85 (Bankr. M.D. Pa. 1996) ("[T]he 'scientific' evidence regarding asbestos . . . in buildings . . . [seems] more litigation driven than science driven."); *Nelson v. Tenn. Gas Pipeline Co.*, No 95-1112, 1998 WL 1297690, at \*8, \*13 (W.D. Tenn. Aug. 31, 1998) (excluding Dr. Kilburn's testimony partly on the grounds that his study "was performed in connection with litigation and funded by plaintiffs' counsel," and Dr. Hirsch's partly on the grounds of his "failure to have . . . conducted prelitigation research").

preme Court had set in 1993;<sup>9</sup> and a less famous case, *Blum v. Merrell Dow*, from a *Frye* state (Pennsylvania) which began, before *Daubert*, in 1982, but didn't come to a final resolution until 2000.<sup>10</sup> In both cases we find expert opinion criticized as based on "litigation-driven science"—though in *Daubert* this criticism was directed at the reliability of the plaintiff's experts' testimony, while in *Blum* it was directed at the legitimacy of the "scientific consensus" to which the defendants' experts appealed; and both Judge Kozinski (in *Daubert*) and Judge Bernstein (in *Blum*) tried to articulate why litigation-driven science is apt to be less dependable than independently-conducted research.

Part III will explain the distinction between inquiry and advocacy and explore the differences between investigation, plain and simple, and advocacy research; and then clear up an ambiguity in "litigation-driven" and some uncertainties in "reliable." This analysis will reveal that research that is litigation-driven in the stronger of the two senses distinguished is inherently in danger of bias; and in consequence is inherently less likely to be—at least in one understanding of that somewhat elusive concept—evidentially reliable.

This, in turn, will suggest some conclusions, articulated in Part IV. There is some truth, as both Judge Kozinski and Judge Bernstein argue, in the idea that the fact that science is litigation-driven indicates that it is more likely to be unreliable. But there is something not quite right about Judge Kozinski's arguments for this conclusion; and the flaws in his arguments reveal that his new *Daubert* factor is not, after all, as helpful as he hopes, or as it might initially seem. This diagnosis leads to some disturbing thoughts about how scientific work can be distorted and impeded when it gets entangled with litigation, and some hard questions about these interactions of science with the law.

## II. A LEGAL THICKET: THE TANGLED TALE OF *DAUBERT* AND *BLUM*

In *Frye v. United States*, in excluding the results of a then-new blood-pressure lie-detector test, the D.C. Court of Appeals had ruled that novel scientific testimony is admissible only if the "scientific

---

<sup>9</sup> *Daubert III*, 509 U.S. 579; see also *Daubert IV*, 43 F.3d at 1311; *Daubert v. Merrell Dow Pharms., Inc. (Daubert II)*, 951 F.2d 1128 (9th Cir. 1991); *Daubert v. Merrell Dow Pharms., Inc. (Daubert I)*, 727 F.Supp. 570 (S.D. Cal. 1989).

<sup>10</sup> *Blum v. Merrell Dow Pharms., Inc.*, 1 Pa. D. & C.4th 634 (Ct. Comm. Pleas Pa. 1988), *rev'd*, 560 A.2d 212 (Pa. Super. Ct. 1989), *aff'd*, 626 A.2d 537 (Pa. 1993), *remanded to* 33 Phila. Co. Rptr. 193 (Ct. Comm. Pleas Pa. 1996), *rev'd*, 705 A.2d 1314 (Pa. Super. Ct. 1997), *aff'd, sub nom. Blum ex rel. Blum v. Merrell Dow Pharms., Inc.*, 747 A.2d 877 (2000).

principle or discovery” from which it is deduced is “sufficiently established to have gained general acceptance in the particular field in which it belongs.”<sup>11</sup> For a decade, *Frye* wasn’t cited even once, and in its first quarter-century it was cited only a dozen or so times; but by the early 1980s it was being cited over and over, and was “probably the ‘majority rule’” in the country.<sup>12</sup>

The Federal Rules of Evidence (FRE), enacted in 1975, provided in Rule 702 that expert testimony is admissible if it is relevant and not legally excluded under Rule 403 on grounds of prejudice, waste of time, or confusing or misleading the jury.<sup>13</sup> Because Rule 702 said nothing about general acceptance, it provoked debate among legal scholars about whether the Federal Rules had or hadn’t superseded *Frye*. By the late 1980s, there was a burgeoning concern with reliability. By 1991, the publication of Peter Huber’s *Galileo’s Revenge: Junk Science in the Courtroom* was fueling fears that flimsy, interested, wildly speculative science was flooding the courts; the same year, some judges on the Federal Rules Advisory Committee sought to change the FRE to include a reliability requirement. By 1992, the first Bush administration was urging similar changes.<sup>14</sup> These initiatives were preempted, however, by the Supreme Court’s ruling in *Daubert*.

\* \* \*

While pregnant, Mrs. Daubert had taken Bendectin for morning sickness (which, though often just a nuisance, may be serious enough to require hospitalization—and can be fatal).<sup>15</sup> Her baby, Jason, was born with severe birth defects. Coming to suspect that Bendectin was the cause, in 1989 the Dauberts brought suit against Merrell Dow

---

<sup>11</sup> *Frye*, 293 F. at 1014.

<sup>12</sup> See Paul C. Giannelli, *Frye v. United States: Background Paper Prepared for the National Conference of Lawyers and Scientists*, 99 F.R.D. 189, 196 (1983).

<sup>13</sup> FED. R. EVID. 702.

<sup>14</sup> See Michael H. Gottesman, *From Barefoot to Daubert to Joiner: Triple Play or Double Error?*, 40 ARIZ. L. REV. 753, 757–59 (1998) (citing Preliminary Draft of Proposed Amendments to the Federal Rules of Civil Procedure and the Federal Rules of Evidence, 137 F.R.D. 53, 73, 156 (1991) (proposal of the Civil Rules Advisory Committee); Dan Quayle, *Agenda for Civil Justice Reform in America*, 60 U. CIN. L. REV. 997, 999 (1992) (proposal of the President’s Competitiveness Committee)).

<sup>15</sup> See, e.g., American Pregnancy Association, *Morning Sickness*, <http://www.americanpregnancy.org/pregnancyhealth/morningsickness.html> (last visited Aug. 9, 2007); American College of Obstetricians and Gynecologists, *Morning Sickness*, [http://www.acog.org/publications/patient\\_education/bp126.cfm](http://www.acog.org/publications/patient_education/bp126.cfm) (last visited Aug. 9, 2007).

Pharmaceuticals, the manufacturer of the drug.<sup>16</sup> The company had taken Bendectin off the market in 1983, shortly after the first reports appeared of children with limb defects being born to women who had taken the drug<sup>17</sup>—though the company maintained that the withdrawal was not prompted by the alleged dangers, but by the potential costs of litigation.<sup>18</sup> (The chemically identical drug is still on sale, by a different company and under a different name, in Canada.<sup>19</sup>)

The Dauberts proffered experts to testify that their re-analyses of the existing data showed a statistical link between Bendectin and limb-reduction birth defects; that Bendectin causes birth defects in laboratory animals, and so probably causes them in humans too; and that Bendectin is chemically similar to other drugs suspected of causing such defects.<sup>20</sup> Merrell Dow's attorneys presented evidence that no clinical trial had ever been published that showed Bendectin to be teratogenic;<sup>21</sup> that despite a wave of Bendectin litigation the Food and Drug Administration (FDA) had continued to approve the drug for use by pregnant women, because "available data do not demonstrate an association between birth defects and Bendectin;"<sup>22</sup> and that the consensus among medical scientists was that the drug was safe.<sup>23</sup>

---

<sup>16</sup> *Daubert v. Merrell Dow Pharms., Inc. (Daubert IV)*, 43 F.3d 1311, 1313 (9th Cir. 1995).

<sup>17</sup> Gottesman, *supra* note 14, at 767.

<sup>18</sup> Astara March, *Drug Revived to Fight Morning Sickness*, NURSE WEEK, Oct. 11, 2000, <http://www.nurseweek.com/news/00-10/1011morn.asp>.

<sup>19</sup> Duchesnay Inc. sells the drug under the name "Diclectin" in Canada, and is seeking FDA approval to sell it in the U.S. Diclectin, <http://www.diclectin.com/index.html> (last visited Jul. 26, 2007). At an FDA/NIH conference held on Dec. 4, 2000, Dr. Gideon Koren of the University of Toronto asked "How safe is safe?" and answered that while in the first meta-analysis, conducted in Toronto, there were 130,000 case controls, and an odds ratio of 1.0, "there was a confidence interval going to 155, which means we cannot say for sure that there isn't a 55 percent increased risk." *Interface of Clinical Pharmacology and Drug Safety* at FDA/NIH Conference, (Dec. 4, 2000), <http://www.fda.gov/cder/present/clinpharm2000/1204preg.txt> (last visited Jul. 30, 2007). So far as I have been able to determine, as of August 2007, Diclectin had not been approved in the U.S.

<sup>20</sup> *Daubert IV*, 43 F.3d at 1314.

<sup>21</sup> A teratogen (from the Greek word, *tera*, meaning "monster") is a substance that causes birth defects.

<sup>22</sup> U.S. Dep't of Health & Human Servs. News, No. P80-45 (Oct. 7, 1980). The ruling in *Daubert IV* ends the quotation here; however, it continues "[Bendectin] should be used only when conservative treatment fails." Joseph Sanders, *From Science to Evidence: The Testimony on Causation in the Bendectin Cases*, 46 STAN. L. REV. 1, 7 (1993). According to the FDA Orange Book Detail Record, "[Bendectin] was not discontinued or withdrawn for safety or efficacy reasons." Drugs@FDA, [http://www.accessdata.fda.gov/scripts/cder/ob/docs/obdetail.cfm?Appl\\_No=010598&TAB](http://www.accessdata.fda.gov/scripts/cder/ob/docs/obdetail.cfm?Appl_No=010598&TAB) (last visited Jul. 31, 2007).

<sup>23</sup> *Daubert IV*, 43 F.3d at 1314.

Before *Daubert*, the *Frye* Rule had been used almost exclusively in criminal cases.<sup>24</sup> Unusually, however, in *Daubert* the trial court (citing *U.S. v. Kilgus* and *Barrel of Fun v. State Farm Fire and Casualty Co.*)<sup>25</sup> had relied significantly on *Frye* in ruling the plaintiffs' expert evidence inadmissible. The Dauberts' proffered scientific testimony was not, as the *Frye* standard requires, generally accepted in the field to which it belongs; for, the court continues, this would require that there be statistically significant epidemiological evidence of causation,<sup>26</sup> but "none of the *published* studies show a statistically significant association between the use of Bendectin and birth defects."<sup>27</sup> So the trial court granted Merrell Dow summary judgment; and in 1991 (citing *U.S. v. Solomon*), the U.S. Court of Appeals for the Ninth Circuit affirmed.<sup>28</sup>

Because of the trial court's almost unprecedented reliance on *Frye*, the key issue as *Daubert* came before the Supreme Court in 1993 was whether the Federal Rules had or hadn't superseded the older rule. Holding that they had, the Supreme Court reversed and remanded. At the same time, however, the Court (re)interpreted FRE 702 as requiring courts to screen proffered scientific testimony not only for relevance, as the Rule explicitly required, but also for reliability; and provided a "flexible list" of indicia that might be considered in assessing whether such testimony was reliable enough to be

---

<sup>24</sup> So when Peter Huber argued that *Frye* had helped keep junk science out of tort cases before the Federal Rules of Evidence relaxed the standards of admissibility, he misrepresented the relevant legal history. See PETER HUBER, GALILEO'S REVENGE: JUNK SCIENCE IN THE COURTROOM, 14–17, 41, 150, 176–77, 199–201, 204 (1993); Peter Huber, *Junk Science in the Courtroom*, 26 VAL. L. REV. 723 (1992); see also Kenneth J. Cheseboro, *Galileo's Retort: Peter Huber's Junk Scholarship*, 42 AM. U. L. REV. 1637, 1687–96 (1994).

<sup>25</sup> *Kilgus* was a criminal case in which testimony identifying the defendant's aircraft using a "forward looking infrared system" had been excluded under *Frye*. *United States v. Kilgus*, 571 F.2d 508, 510 (9th Cir. 1978) ("[A] necessary predicate to the admission of scientific evidence is that the principle on which it is based 'must be sufficiently established to have general acceptance in the field to which it belongs.'" (quoting *United States v. Brown*, 557 F.2d 541, 546 (6th Cir. 1977))). *Barrel of Fun* was a fire-insurance fraud case in which polygraph testimony had been excluded under *Frye*; so far as I know, it was the only civil case before *Daubert I* which had relied on *Frye*. *Barrel of Fun v. State Farm Fire & Cas. Co.*, 739 F.2d 1028 (5th Cir. 1984).

<sup>26</sup> *Daubert v. Merrell Dow Pharms., Inc. (Daubert I)*, 727 F.Supp. 570, 573 (S.D. Cal. 1989).

<sup>27</sup> *Id.* at 575.

<sup>28</sup> *Daubert v. Merrell Dow Pharms., Inc. (Daubert II)*, 951 F.2d 1128, 1129–30 (9th Cir. 1991) (citing *United States v. Solomon*, 753 F.2d 1522, 1526 (9th Cir. 1985)). *Solomon* was a murder case in which the higher court affirmed the trial court's exclusion, under *Frye*, of evidence concerning narcoanalysis.

admissible: whether the proffered testimony “can be (and has been) tested; the known or potential error rate;” whether the evidence has been subject to scrutiny by way of “peer review and publication”; and (in a nod to *Frye*) whether it is generally accepted in the field to which it belongs.<sup>29</sup>

Rehearing the case on remand, in a memorable passage that would soon be much cited by judges and legal commentators,<sup>30</sup> Judge Kozinski wrote of the formidable task the Supreme Court had set for him and his colleagues on the federal bench:

Federal judges ruling on the admissibility of expert scientific testimony face a far more complex and daunting task in a post-Daubert world than before. . . . [T]hough we are largely untrained in science and certainly no match for any of the witnesses whose testimony we are reviewing, it is our responsibility to determine whether those experts’ proposed testimony amounts to “scientific knowledge,” constitutes “good science,” and was derived by the “scientific method.” . . . [W]e take a deep breath and proceed with this heady task.<sup>31</sup>

The *Daubert* Court, he noted, had not supplied a “definitive checklist” of indicia of reliability, only an illustrative list of the factors to which courts might look;<sup>32</sup> but this list raised some tricky questions:

[H]ow do we determine whether the rate of error is acceptable, and by what standard? . . . [W]hat should we infer from the fact that the methodology has been tested, but only by the party’s own expert . . . ? Do we ask whether the methodology they employ to test their methodology is itself methodologically sound? . . . [T]he

<sup>29</sup> *Daubert v. Merrell Dow Pharms., Inc. (Daubert III)*, 509 U.S. 579, 593–94 (1993).

<sup>30</sup> See, e.g., Sofia Adroque, *The Post-Daubert Court: “Amateur Scientist” Gatekeeper or Executioner?*, 35 HOUS. L. REV., 10, 10 (1998); Mark S. Brodin, *Behavioral Science Evidence in the Age of Daubert: Reflections of a Skeptic*, 73 U. CIN. L. REV. 862, 867 (2005); Judge Harvey Brown, *Eight Gates for Expert Witnesses*, 36 HOUS. L. REV. 743, 784 (1999); Lee Epstein & Gary King, *Empirical Research and the Goals of Legal Scholarship: The Rules of Inference*, 69 U. CHI. L. REV. 1, 133 (2002); David L. Faigman, *The Law’s Scientific Revolution: Reflections and Ruminations on the Law’s Use of Experts in Year Seven of the Revolution*, 57 WASH. & LEE L. REV. 661, 684 (2000); G. Michael Fenner, *The Daubert Handbook: The Case, Its Essential Dilemma, and Its Progeny*, 29 CREIGHTON L. REV. 939, 1066–67 (1996); Robert J. Goodwin, *The Hidden Significance of Kumho Tire v. Carmichael: A Compass for Problems of Definition and Procedure Created by Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 52 BAYLOR L. REV. 603, 646 n.60 (2000).

<sup>31</sup> *Daubert IV*, 43 F.3d at 1315–16.

<sup>32</sup> As the Court confirmed in *Kumho Tire v. Carmichael*, 526 U.S. 137, 138 (1999): “A trial judge determining the admissibility of an engineer’s testimony may consider one or more of the specific *Daubert* factors. The emphasis on the word ‘may’ reflects *Daubert*’s description of the Rule 702 inquiry as ‘a flexible one.’” (quoting *Daubert III*, 509 U.S. at 594).



basic problem . . . is that we must devise standards for acceptability where respected scientists disagree on what's acceptable.<sup>33</sup>

Reviewing the Supreme Court's flexible list of indicia of reliability, Judge Kozinski proposes a new "*Daubert* factor" of his own: whether the proffered expert testimony is based on work undertaken in the normal course of scientific business, or on work conducted specifically for the purposes of litigation. He stresses the likely flaws and failings of litigation-driven science:

One very significant fact to be considered is whether the experts are proposing to testify about matters growing naturally and directly out of research they have conducted independent of the litigation, or whether they have developed their opinions expressly for the purpose of testifying. That an expert testifies for money does not necessarily cast doubt on the reliability of his testimony, as few experts appear in court merely as an eleemosynary gesture. But in determining whether proposed expert testimony amounts to good science, we may not ignore the fact that a scientist's normal workplace is the lab or the field, not the courtroom or the lawyer's office.<sup>34</sup>

Referring to Huber's *Galileo's Revenge*, Judge Kozinski suggests that the fact that an expert testifies on the basis of work he has conducted independent of litigation "provides important, objective proof that the research comports with the dictates of good science";<sup>35</sup> and that the fact that research is litigation-driven is an indication that it may not comport with those dictates. In this context he cites Judge Johnson's ruling in *Perry*: "the examination of a scientific study by a cadre of lawyers is not the same as its examination by others trained in the field of science or medicine."<sup>36</sup>

Judge Kozinski gives two main reasons why science conducted independently of the needs of litigation is more likely to be reliable than litigation-driven science:

[a] [E]xperts whose findings flow from existing research are less likely to have been biased toward a particular conclusion by the promise of remuneration . . . .

---

<sup>33</sup> *Daubert IV*, 43 F.3d at 1316–17 n.3.

<sup>34</sup> *Id.* at 1317 (footnote omitted).

<sup>35</sup> *Id.*

<sup>36</sup> *Id.* at 1318 n.8 (citing *Perry v. United States*, 755 F.2d 888, 892 (11th Cir. 1985)). Now, of course, this stress on the important differences between in-court "testing" by cross-examination and testing in the sciences will bring to mind Judge Pollak's comments about fingerprint identification. *United States v. Llera-Plaza*, No. 98-362-10, 2002 WL 27305, at \*10 (E.D. Pa. Jan. 7, 2002).

[b] [I]ndependent research carries its own indicia of reliability, as it is conducted . . . in the usual course of business and must normally satisfy a variety of standards to attract funding and institutional support.<sup>37</sup>

Referring again to Huber's book, Judge Kozinski suggests that proffered scientific testimony that is not based on research independent of litigation requires some other indication of reliability;<sup>38</sup> specifically, he suggests, had their work been subjected to peer review and publication, this would provide some assurance that the plaintiffs' experts' research was in accordance with the scientific method, as understood by at least a minority of the relevant scientific community. But not only had the plaintiffs' proffered experts conducted their work for the purposes of litigation; not one of them had published his Bendectin research in peer-reviewed journals, as they had their other scientific work. Given that their findings would surely be of interest to the scientific community, Judge Kozinski continues, the fact that they had been unable or unwilling to publish them undermines the idea that these results are, as the *Daubert* standards required, "grounded in the methods and procedures of science."<sup>39</sup>

In a startling but tantalizingly brief footnote to which we shall have to return in due course, he adds that "[t]here are, of course, exceptions"—kinds of litigation-driven science of which, he believes, there is no reason to be skeptical.<sup>40</sup> Some forensic sciences, such as fingerprinting or DNA identification techniques, "have the courtroom as a principal theater of operations"; but here the fact that an expert has developed an expertise primarily for purposes of litigation "will obviously [*sic*] not be a substantial consideration."<sup>41</sup> (As we shall see later, however, really this is far from obvious.)

Early in his ruling, Judge Kozinski had observed that "apart from a small but determined group of scientists testifying on behalf of the Bendectin plaintiffs in this and many other cases, there doesn't appear to be a scientist who has concluded that Bendectin causes limb-reduction defects";<sup>42</sup> under *Frye*, which had been the law of the circuit at the time when the Dauberts' experts submitted their affidavits, their testimony would certainly have to be excluded. However, given

---

<sup>37</sup> *Daubert IV*, 43 F.3d at 1317.

<sup>38</sup> *Id.* at 1316–17.

<sup>39</sup> *Id.* at 1318 (citing *Daubert v. Merrell Dow Pharms., Inc. (Daubert III)*, 509 U.S. 579, 589–90 (1993)).

<sup>40</sup> *Id.* at 1317 n.5.

<sup>41</sup> *Id.*

<sup>42</sup> *Id.* at 1314.

that the law had changed in the meantime, they might have been given an opportunity to submit additional proof that their proffered evidence was, as required by the Supreme Court's ruling in *Daubert*, "derived by the scientific method"—but for the fact, Judge Kozinski argues, that it was already clear this wouldn't change the outcome: the Dauberts' proffered expert testimony would clearly have to be excluded under the new standards, as it was under the old.<sup>43</sup>

Surprisingly, however, Judge Kozinski's reasoning to this conclusion makes little use of the idea that litigation-driven science is especially suspect. In fact—despite his mock-modest announcement at the outset that he will “take a deep breath and proceed with [the] heady task” of assessing the reliability of the proffered science<sup>44</sup>—it leaves scientific issues essentially untouched. Moreover, it calls on the reliability prong of *Daubert* with respect to only one of the plaintiffs' experts, Dr. Palmer—the only proffered expert who would testify that Bendectin actually *did* cause Jason Daubert's birth defects, rather than that it *could possibly* have caused them; and the fact that Dr. Palmer's research was litigation-driven plays no specific role in Judge Kozinski's argument why it would have to be excluded, which is simply that “Dr. Palmer offers no tested or testable theory to explain how . . . he was able to eliminate all other potential causes of birth defects . . . .”<sup>45</sup> The other proffered experts, who would speak in terms of probabilities, would have to be excluded under the relevance prong; for none of them even claimed to show, as required, that Bendectin more than doubles the risk of such defects.<sup>46</sup>

\* \* \*

Like Mrs. Daubert, Mrs. Blum took Bendectin for morning-sickness; like Jason Daubert, Jeffrey Blum was born with severe defects—in his case, clubbed feet; like the Dauberts, the Blums believed Bendectin was the cause. In 1982, seven years before the Dauberts' suit, Jeffrey Blum's parents brought suit against Merrell Dow; and the case slowly wound its way through the Pennsylvania courts for eighteen years before being finally resolved, years after *Daubert*, in 2000.

---

<sup>43</sup> *Daubert IV*, 43 F.3d at 1319–20.

<sup>44</sup> *Id.* at 1316.

<sup>45</sup> *Id.* at 1319.

<sup>46</sup> *Id.* at 1320–21 (“California tort law requires that . . . plaintiffs must establish not just that their mothers' ingestion of Bendectin increased somewhat the likelihood of birth defects, but that it more than doubled it. . . . None of the plaintiffs' epidemiological experts claims that the ingestion of Bendectin during pregnancy more than doubles the risk of birth defects.”); *see also*, Sanders, *supra* note 22, at 16 n.63, on this standard for proof of specific causation.

The first trial ended in 1988 with a jury verdict for the plaintiffs.<sup>47</sup> Merrell Dow appealed, on the grounds that the verdict had been reached by only eleven jurors (the twelfth had fallen ill part-way through the trial); and in 1993 was granted a new trial.<sup>48</sup> On remand, in 1996 the Court of Common Pleas again entered judgment on jury verdict for the plaintiffs (this time with the full complement of jurors).<sup>49</sup> Merrell Dow appealed again, this time on the grounds that the plaintiffs' scientific testimony should have been excluded by the court; the jury—on the vital importance of which they had earlier insisted—should never have been allowed to hear it. In 1997, the Superior Court held that plaintiffs' expert testimony regarding the causal link between Bendectin and birth defects was not admissible under *Frye*, and remanded the case "with instructions to the trial court to enter" judgment n.o.v. in favor of Merrell Dow.<sup>50</sup> In 1999, the Supreme Court of Pennsylvania granted allocatur<sup>51</sup> to consider whether the *Frye* rule still governed the admissibility of expert scientific testimony in Pennsylvania, or had been superseded by *Daubert*. In 2000, declining to replace *Frye* by what it took to be<sup>52</sup> the more relaxed standards of *Daubert*, but arguing that the Blums' expert testimony was inadmissible under either standard, the Pennsylvania Supreme Court affirmed the decision of the Superior Court in favor of the defendant manufacturer.<sup>53</sup>

While Merrell Dow had maintained that the plaintiffs' proffered expert scientific testimony should have been excluded because it wasn't generally accepted in the scientific community, the Blums' at-

<sup>47</sup> *Blum v. Merrell Dow Pharms., Inc.*, 1 Pa. D. & C.4th 634, 635 (Ct. Comm. Pleas Pa. 1988).

<sup>48</sup> *Blum ex. rel. Blum v. Merrell Dow Pharms., Inc.*, 626 A.2d 537, 549 (Pa. 1993).

<sup>49</sup> *Blum ex. rel. Blum v. Merrell Dow Pharms., Inc.*, 33 Phila. Co. Rptr. 193, 243 (Ct. Comm. Pleas Pa. 1996), *rev'd*, 705 A.2d 1314 (Pa. Super. Ct. 1997), *aff'd*, 764 A.2d 1 (Pa. 2000).

<sup>50</sup> *Blum ex. rel. Blum v. Merrell Dow Pharms., Inc.*, 705 A.2d 1314, 1325 (Pa. Super. Ct. 1997), *aff'd*, 764 A.2d 1 (Pa. 2000).

<sup>51</sup> 10 TIMOTHY P. WILE & MARC A. WERLINSKY, WEST'S PENNSYLVANIA PRACTICE DRIVING UNDER THE INFLUENCE § 30:53 (2006–07 ed.) (explaining that "[a]nother name for allowance of appeal is allocatur").

<sup>52</sup> The Supreme Court's rhetoric had described the *Frye* Rule as an "austere standard" which the Federal Rules had relaxed. *Daubert v. Merrell Dow Pharms., Inc. Daubert III*, 509 U.S. 579, 589 (1993). Other states besides Pennsylvania—Florida, for one—also took this rhetoric at face value. *Brim v. State*, 695 So. 2d 268, 271–72 (Fla. 1997) ("[d]espite the federal adoption of a more lenient standard . . . we [Florida] have maintained the higher standard of reliability as dictated by *Frye*"). In practice, however, *Daubert* standards have arguably proven significantly more restrictive than the old *Frye* rule.

<sup>53</sup> *Blum ex. rel. Blum*, 764 A.2d at 4–5.

torneys had argued that Merrell Dow's expert testimony should have been excluded because the supposed "scientific consensus" on this matter was completely artificial; it had been created by the defendant manufacturer's support of favorable research and of questionable peer-reviewed journals that would publish results helpful to the company in defending itself against Bendectin litigation. Dissenting in part from the Pennsylvania Supreme Court's final disposition of the case in favor of Merrell Dow, Justice Castille summed up the issue:

[I]n the litigation-driven Bendectin "scientific community" described to the court in this case, the notion of "general acceptance" or scientific "orthodoxy" . . . on the question of causation was a questionable proposition to begin with . . . because the trial court had heard extensive evidence concerning Merrell Dow's active and deliberate role, motivated by its litigation interests . . . , in actually creating and influencing the scientific orthodoxy that would then operate to suppress any contrary opinion that might harm its Bendectin litigation . . . .<sup>54</sup>

Justice Castille refers us to Judge Bernstein's ruling at the second trial.

This ruling—the tone of which is, to say the least, unusually impassioned—opens with a remarkable excerpt from the testimony of James Newberne, Merrell Dow's Vice-President for Drug Safety:

Q: Sir, it has been the pattern and practice and custom of the Merrell Company, in reporting to the FDA, to pick and choose selective information over the past thirty years, relating to the drug Bendectin, correct?

A: Yes, that's correct.<sup>55</sup>

Judge Bernstein first summarizes the testimony of the Blums' expert witnesses (including some who had been unsuccessfully proffered by the Dauberts),<sup>56</sup> and then subjects the testimony given by Merrell Dow's experts to devastating scrutiny.

Plaintiffs' expert Dr. Gross testified that a review of Merrell Dow's animal testing revealed that there were significant numbers of abnormalities, including club limbs, that had not been reported to

<sup>54</sup> *Id.* at 7–8 (Castille, J., dissenting). Justice Castille later returned to the issue, citing his own dissenting opinion in *Blum* in his concurring opinion in *Grady v. Frito-Lay, Inc.*, 839 A.2d 1038, 1048 (Pa. 2003) (Castille, J., concurring).

<sup>55</sup> *Blum ex rel. Blum v. Merrell Dow Pharms., Inc.*, 33 Phila. Co. Rptr. 193, 194 (Ct. Comm. Pleas Pa. 1996), *rev'd*, 705 A.2d 1314 (Pa. Super. Ct. 1997), *aff'd*, 764 A.2d 1 (Pa. 2000).

<sup>56</sup> Among the Dauberts' proffered experts were Dr. Gross, Dr. Newman, and Dr. Done. *Daubert v. Merrell Dow Pharms., Inc. (Daubert IV)*, 43 F.3d 1311, 1321 (9th Cir. 1995).

the FDA.<sup>57</sup> Dr. Done testified to the chemical similarity of Doxylamine, one of the active ingredients in Bendectin, to other known teratogens, to in vitro studies showing its detrimental effect on limb bud cells, and to his re-analyses of two epidemiological studies which, in his opinion, showed an increased risk of clubfeet in the infants of women who took Bendectin in “the first four months of pregnancy.”<sup>58</sup> Dr. Newman testified that Doxylamine “can pass through the placental barrier” and affect the embryo.<sup>59</sup> The testimony given by Dr. Stolley at the previous trial, that “there was three times the risk of malformations” in babies whose mothers “had filled more than one prescription for Bendectin,” was read into evidence.<sup>60</sup>

Most important here, however, is Judge Bernstein’s summary and scathing commentary on Merrell Dow’s experts’ testimony:

Defense expert Dr. Bracken, a professor of epidemiology at Yale, testified that his study (based on interviews with 1427 mothers, of whom only 122 had taken Bendectin) concluded that Bendectin carried no significant risk of birth defects except for pyloric stenosis; however, he acknowledged that it showed there was a more than two-and-a half times greater risk of birth defects in infants born to women who took Bendectin and also smoked.<sup>61</sup> On cross-examination, he agreed not only that articles that are “less than good” can pass peer review, but also that his own published study of Bendectin and birth defects was itself less than good.<sup>62</sup>

Defense expert Dr. Klebanoff, who began his work on Bendectin long after the drug was taken off the market, testified that Bendectin does not cause birth defects, but acknowledged that his own article showed a statistically significant association with congenital cataracts, underdevelopment of the lungs, and microcephaly.<sup>63</sup> Under cross-examination, he agreed that Bendectin is positively associated with clubbed feet.<sup>64</sup>

Defense expert Dr. Tyl, a developmental toxicologist, was hired by the federal government, again long after Bendectin had been withdrawn, to perform animal studies on the drug.<sup>65</sup> She testified

---

<sup>57</sup> *Blum ex. rel. Blum*, 33 Phila. Co. Rptr. at 202.

<sup>58</sup> *Id.* at 203.

<sup>59</sup> *Id.* at 206.

<sup>60</sup> *Id.*

<sup>61</sup> *Id.* at 207–08.

<sup>62</sup> *Id.* at 207.

<sup>63</sup> *Blum ex. rel. Blum*, 33 Phila. Co. Rptr. at 208–09.

<sup>64</sup> *Id.* at 209.

<sup>65</sup> *Id.* at 209–15.

that “Bendectin is not a teratogen, but it is a ‘developmental toxicant,’” and that as a result of her work the drug had been placed on the “List of Developmental Toxicants” maintained by the U.S. Government. A developmental toxicant, she explained, is defined as an indicator of such defects as reduced “body weight, reduced survival, increased number of variations, reduced ossification; . . . and certain morphological changes.”<sup>66</sup>

Defense expert Dr. Shapiro (whose formal training in epidemiology amounted only to eleven credits toward a Master’s degree) was head of the Slone Center for Epidemiology at Boston University in a period when the unit received over one and a half million dollars in research-support funds from Merrell Dow.<sup>67</sup> He testified that Bendectin could not cause birth defects. However, the data on which he based his opinion lumped together women who took Bendectin during the period when limbs were forming, and those who took the drug only after the baby’s limbs had formed.<sup>68</sup> He agreed that this resulted in an underestimate of the incidence of clubfeet in the group exposed to Bendectin, but refused to attribute any significance to this.<sup>69</sup> If Bendectin did cause birth defects, he explained, his study might have underestimated the risk; but since Bendectin does not cause birth defects, his study could not have done so.<sup>70</sup>

Defense expert Dr. Newberne admitted that Merrell Dow had engaged in “a consistent pattern of underreporting” of adverse effects of Bendectin to the FDA. He acknowledged that during the period when a study by Dr. Smithells supposedly showing the safety of Bendectin had been rejected by the *British Medical Journal*, *The Lancet*, and the *New England Journal of Medicine*, and eventually was accepted by the much less prestigious journal *Teratology*, the author was actively seeking funds from the company, writing that “[m]uch clearly depends upon the value of this publication to Merrell Dow . . . . If it may save the company large sums of money . . . in the California court (which is rather what I thought when we undertook this study), they may feel magnanimous.”<sup>71</sup> Dr. Newberne also testified that

---

<sup>66</sup> *Blum ex rel. Blum*, 33 Phila. Co. Rptr. at 211 n.54, *rev’d*, 705 A.2d 1314 (Pa. Super. Ct. 1997), *aff’d*, 764 A.2d 1 (Pa. 2000).

<sup>67</sup> *Blum ex rel. Blum*, 33 Phila. Co. Rptr. at 214 & n.69.

<sup>68</sup> *Id.* at 215–17.

<sup>69</sup> *Id.* at 217.

<sup>70</sup> *Id.* (the record says that Dr. Shapiro was head of the Department of Epidemiology at Boston University, but Dr. Richard Clapp of the Boston University School of Public Health tells me this is incorrect).

<sup>71</sup> *Id.* at 219.

Merrell Dow had supported Dr. Shapiro's research at Boston University and Dr. Hendricks's in California out of its legal defense funds.<sup>72</sup>

Defense expert Dr. Brent, the editor of *Teratology*, who had been retained as an expert by Merrell Dow for eighteen years, testified that his only formal education in epidemiology was one course in statistics, but considered himself the world authority in "secular trend data"—a scientific field in which, Judge Bernstein adds, there is apparently only one practitioner, Dr. Brent himself.<sup>73</sup> Using his editorial prerogative to sidestep peer review, he had published in his own journal an article entitled *Litigation-Produced Pain, Disease, and Suffering: An Experience with Congenital Malformation Lawsuits*, which concluded, based on his review of deposition and trial transcripts, that seventeen out of seventeen plaintiffs lied.<sup>74</sup> He also testified that he had submitted a draft article entitled *Bendectin: The Most Comprehensively Studied Human Non-Teratogen, and the Foremost Tortogen-Litigen* to Merrell Dow's attorneys for editing, hoping to publish it in *The New England Journal of Medicine*, the *Journal of the American Medical Association*, or *The Lancet*.<sup>75</sup>

Dr. Newberne's testimony, Judge Bernstein comments, revealed "[t]he interaction of 'scientific studies' and litigation defense";<sup>76</sup> Dr. Brent's testimony clearly "revealed a sycophantic relationship be-

---

<sup>72</sup> *Id.* at 221–22.

<sup>73</sup> *Blum ex rel. Blum*, 33 Phila. Co. Rptr. at 224.

<sup>74</sup> Robert L. Brent, *Litigation-Produced Pain, Disease, and Suffering: An Experience with Congenital Malformation Lawsuits*, 16 *TERATOLOGY* 1, 5 (1977).

<sup>75</sup> *Blum ex rel. Blum*, 33 Phila. Co. Rptr. at 225, *rev'd*, 705 A.2d 1314 (Pa. Super. Ct. 1997), *aff'd*, 764 A.2d 1 (Pa. 2000). The article in question appears to be Robert L. Brent, *Bendectin: Review of the Medical Literature of a Comprehensively Studied Human Non-teratogen and the Most Prevalent Tortogen-Litigen*, 9 *REPROD. TOXICOLOGY* 337 (1995). This paper prompted a lawsuit for defamation by Dr. Stuart A. Newman, whom Dr. Brent had misquoted, against Dr. Brent and the editor of *Reproductive Toxicology*. At the suggestion of the presiding judge, the parties were invited to air their differences in a scientific forum. See Stuart A. Newman, *Dr. Brent and Scientific Debate*, 13 *REPROD. TOXICOLOGY* 241 (1999) (complaining of the "partisan" nature of Dr. Brent's work, which "should have raised questions about the objectivity of the peer review and editorial process," and noting his association with the law firm of Dinsmore & Shohl, which represented Merrell Dow in many of its Bendectin cases); Robert L. Brent, *Response to Dr. Stuart Newman's Commentary on an Article Entitled "Bendectin: Review of the Medical Literature of a Comprehensively Studied Human Nonteratogen and the Most Prevalent Tortogen-litigen,"* 13 *REPROD. TOXICOLOGY* 245 (1999) (pointing out that Dr. Newman's testimony had been excluded in several Bendectin cases); Stuart A. Newman, *A Response to Dr. Brent's Commentary on "Dr. Brent and Scientific Debate,"* 13 *REPROD. TOXICOLOGY* 255, 256 (1999) (noting that much of Dr. Brent's response relies on judges' opinions regarding scientific issues); see also Robert L. Brent, *Bendectin and Birth Defects: Hopefully, the Final Chapter*, 67 *BIRTH DEFECTS RES.* 79 (2003) (urging the reintroduction of Bendectin as effective and harmless).

<sup>76</sup> *Blum ex rel. Blum*, 33 Phila. Co. Rptr. at 222.



tween Dr. Brent and the attorneys representing Merrell Dow.”<sup>77</sup> Moreover, testimony presented “clearly demonstrated that not all ‘peer review’ journals are created equal,” that “not all the articles contained in ‘peer review’ journals were even reviewed,” and that “articles were intentionally inserted in peer review journals for use in court.”<sup>78</sup> But most immediately to the present purpose is Judge Bernstein’s exasperated commentary on Dr. Shapiro’s testimony: when asked by the court whether his study underestimated the risk of Bendectin, Dr. Shapiro replied, “yes”; but immediately went on to add that what he meant was only that, if there *were* a causal relationship, it *would have been* underestimated, but “[i]f there were no causal relationship, which is what I believe . . . there could not have been any underestimates.”<sup>79</sup> “The circularity of this reasoning,” Judge Bernstein argues, makes it unmistakably clear that Dr. Shapiro was engaged in “justification science not inquisitive science”; and, he continues, “[c]learly revealed in this testimony is the unalterable preconception from which Dr. Shapiro’s ‘scientific conclusion’ was derived.”<sup>80</sup> Dr. Shapiro’s conviction that Bendectin is not teratogenic was so firm from the outset that he was virtually impervious to any evidence that might suggest otherwise.

\* \* \*

Is Bendectin teratogenic? After reading only *Daubert*, an intelligent, fair-minded layperson would be inclined to say: “almost certainly not.” After reading *Blum*, however, he might well say, as I would: “I’m not so sure as I was; it’s all very confusing.” Maybe Merrell Dow overstepped ethical boundaries in protecting its interests in that self-defeating way to which defendant manufacturers seem prone,<sup>81</sup> but maybe they really had something to hide. For someone

---

<sup>77</sup> *Id.* at 225 (responding to Dr Brent’s claim that there was “a sycophantic alliance between the expert witness and the plaintiff’s attorney”).

<sup>78</sup> *Id.* at 246–47.

<sup>79</sup> *Id.* at 217.

<sup>80</sup> *Id.*

<sup>81</sup> I am thinking here, for example, of the instructions to salespeople uncovered by Dan Bolton, attorney for Maria Stern in her 1984 case against Dow Corning alleging injuries caused by her silicone breast-implants. The incriminating memo reads, in part:

[I]t has been observed that the new mammaries with responsive gel have a tendency to appear oily after being manipulated. This could prove to be a problem with your daily detailing activity. . . . You should make plans to change demonstration samples often. Also, be sure that samples are clean and dry before customer detailing.

outside the relevant fields, it's almost impossible to know.<sup>82</sup> But for someone in my field—epistemology—the tangled tale of *Daubert* and *Blum* is just the kind of tangle to make the fingers itch.

### III. AN EPISTEMOLOGICAL SWAMP: THE SINKING SANDS OF “LITIGATION-DRIVEN” AND “EVIDENTIARY RELIABILITY”

Though, as Jonathan Rauch once observed, a good way “to clear the room at a cocktail party”<sup>83</sup> is to use the word “epistemology,” I hope that in this more academic forum I may be permitted to observe that the law is up to its neck in epistemological concepts and questions.<sup>84</sup> In the present context we need to understand, at a minimum, the difference between inquiry and advocacy; the nature of advocacy research; the contrast between disinterestedness and bias; and the relation of all these to issues about truth and reliability.

Inquiry, investigation—the professional business of scientists, historians, legal and literary scholars, investigative journalists, and so forth—is a matter of trying to discover the answer to some question: who committed the crime, what caused the cancer or made it advance so quickly, where did the money go, etc.? Advocacy, by contrast—the professional business of lobbyists, attorneys, and so on—is a matter of trying to persuade an audience of the truth of some proposition: that my client didn't do it, that it was work-related PCB exposure that promoted the tumor, that the stolen money has been hidden in a numbered account in the Cayman Islands, etc.

---

MARCIA ANGELL, SCIENCE ON TRIAL: THE CLASH OF MEDICAL EVIDENCE AND THE LAW IN THE BREAST IMPLANT CASE 59 (1996).

<sup>82</sup> Sanders, *supra* note 27. Prof. Sanders' conclusion is that the weight of the scientific work indicated that Bendectin is probably not teratogenic, but that the evidence presented to juries in Bendectin cases did not accurately represent the true state of the science. *Id.* at 3. He acknowledges, however, that some *in vivo* studies have shown teratogenic effects; that six epidemiological studies had found a statistically significant correlation between Bendectin and certain types of defect; that many studies failed to pinpoint the time in pregnancy during which mothers took Bendectin; and that the presence of the suspect ingredient, Doxylamine Succinate, in two over-the-counter drugs—Unisom and Nyquil—that some subjects may have taken could have skewed study results. *Id.* at 25–26. Prof. Sanders's description of some of the supposedly reassuring animal-testing work undertaken by Merrell Dow in 1966–1967, in the wake of the Thalidomide disaster, also leaves one a little uneasy: “[a]lthough their test animals suffered several defects, Newberne and Gibson did not attribute the defects to Bendectin.” *Id.* at 21.

<sup>83</sup> JONATHAN RAUCH, KINDLY INQUISITORS: THE NEW ATTACKS ON FREE THOUGHT 35 (1993).

<sup>84</sup> Susan Haack, *Epistemology Legalized: Or, Truth, Justice, and the American Way*, 49 AM. J. JURIS. 43, 44 (2004).

Magistrate Judge Breen observes in *Nelson* that we want expert opinions to be “about science, . . . not advocacy.”<sup>85</sup> That distinction is clear enough; but the most relevant distinction here is between inquiry, investigation, real research, i.e., really trying to find the true answer to some question, whatever that truth may be, and advocacy research, i.e., trying to find the strongest possible evidence for the truth of some proposition determined in advance. This, I take it, was the distinction Judge Bernstein had in mind when he contrasted Dr. Shapiro’s “justification science” (i.e., advocacy research) with “inquisitive science” (i.e., real research, inquiry plain and simple).

Distinguishing genuine inquiry, the real thing, from pseudo-inquiry or “sham reasoning,” C.S. Peirce—a working scientist as well as the greatest of American philosophers—wrote that “[t]he spirit . . . is the most essential thing—the motive”; that genuine inquiry consists in “actually drawing the bow upon truth with intentness in the eye, with energy in the arm.”<sup>86</sup> For the same reason, I am tempted to write of advocacy “research” (in scare quotes); for it is something of a stretch to call advocacy research “research” at all. Advocacy “research” is like inquiry insofar as it involves seeking out evidence. But it is part of an advocacy project insofar as it involves seeking out evidence favoring a predetermined conclusion; and it is undertaken in the spirit, from the motive, of an advocate. In short, it is a kind of pseudo-inquiry.

There’s nothing wrong with advocacy, as such. There’s nothing wrong, even, with a scientist taking on the role of advocate—even on matters related to his own field; indeed, it might be argued that if a medical or environmental scientist, for example, discovers a hitherto unsuspected health risk or benefit, he has a moral obligation to bring it to the public attention as effectively as possible. But there *is* something wrong with advocacy research. Investigating the risks and benefits of taking this dietary supplement or damning that river is a quite different enterprise from advocating that the supplement be banned or that the dam be built; and while it is highly desirable that advocacy be based on the results of well-conducted investigation, it is highly *undesirable* that advocacy be allowed to slant investigation.

Obviously enough, someone straightforwardly investigating a question and someone engaged in advocacy research on behalf of a

<sup>85</sup> *Nelson v. Tenn. Gas Pipeline Co.*, No 95-1112, 1998 WL 1297690, at \*9 (W.D. Tenn. Aug. 31, 1998).

<sup>86</sup> CHARLES SANDERS PEIRCE, *COLLECTED PAPERS OF CHARLES SANDERS PEIRCE* 1.34 (Charles Hartshorne, et. al. eds., 1931–58) (1903); *id.* 1.235 (1902) (references to the *COLLECTED PAPERS* are by volume and paragraph number).

particular answer take different attitudes to the evidence. The plain-and-simple inquirer wants to find the answer (though the upshot may be a realization that his question was in some way misconceived, and when he does find an answer, he will often find himself faced with a slew of new questions). He is motivated to seek out all the evidence he can lay hands on, to weigh it as judiciously as possible, to assess where it leads as carefully as he can, and to suspend judgment unless and until his evidence warrants drawing a conclusion. An advocacy researcher, by contrast, is motivated to seek out all the evidence that favors his predetermined conclusion, but to ignore, play down, or explain away any evidence contrary to that conclusion.

So, being motivated to seek out all the evidence, the plain-and-simple inquirer will be *more thorough* than the advocacy researcher looking only for favorable evidence; being concerned to find the answer whatever the answer may be, he will be *less partial* than the advocacy researcher trying to minimize the importance of unfavorable evidence he can neither ignore nor explain away; and, being ready to acknowledge evidence either way, he will be *more honest* than the advocacy researcher trying to disguise what doesn't suit his purpose. This is why he is likelier than an advocacy researcher—other things (his ability, energy, resources, etc.) being equal—to discover the truth; the more so, the longer he inquires.

Connections with the concepts of interestedness and bias now begin to come into focus. In one sense, to describe an inquirer as “interested” means that he takes an interest in the question he is investigating (he isn't bored by it or uninterested in it, nor is he just dutifully but unenthusiastically doing what is required by his job or demanded by his Ph.D. supervisor). In another and potentially more problematic sense, it means that he has an interest in the answer to the question coming out this way rather than that, i.e., he stands to gain in some way from reaching this conclusion rather than a different one. And in a third sense, the most problematic, an interested investigator is really only an “investigator”; for the way he proceeds is distorted by his desire that the answer come out in the way by which he stands to gain. Often, but not always or inevitably, someone who is interested in the second sense is also interested in the third. It is the third sense that chiefly concerns us here; for an “inquirer” who is interested in this sense is bound to be biased: that is, to lean in one direction, to play up the evidence on one side of his question and play down anything negative. (This reveals the connection between the two senses of “partial”: an investigator who is partial, in the sense of

“biased toward one side of an issue,” will concentrate selectively on evidence which is partial, in the sense of “incomplete.”)

Peirce’s prime example of sham reasoning was the “seminary philosophy” dominant in his day. Theologians, he argued, being professionally committed to the truth of certain propositions, are professionally obliged to adjust their philosophical arguments so as to preserve and support those propositions.<sup>87</sup> So perhaps it is no wonder that a prime contemporary example that comes to my mind is the “research” offered by its proponents in favor of Intelligent Design Theory. So far as I can see, this amounts only to efforts, often botched, and sometimes apparently outright dishonest,<sup>88</sup> to identify “gaps and problems” in the theory of evolution, and to cover up the much more formidable gaps and problems in Intelligent Design Theory.<sup>89</sup> Judge Jones’s unremittingly commonsense analysis in *Kitzmiller*—noting that even some of the expert witnesses for the defendant school district acknowledge that there is no real scientific research supporting Intelligent Design Theory—does a pretty good job of unmasking this sham reasoning.<sup>90</sup>

Of course, the real world is always much messier than philosophers would like. Rather than a simple division into genuine and pseudo-inquiry, honest and dishonest inquirers, we find more and less plain-and-simple investigation, and just about every degree and shade of intellectual honesty and dishonesty.<sup>91</sup> The categorical dis-

<sup>87</sup> *Id.* at 1.620 (1898).

<sup>88</sup> For example, the Intelligent Design biology text, PERCIVAL DAVIS & DEAN KENYON, *OF PANDAS AND PEOPLE* 104 (3d ed. 1993), stresses the absence of transitional fossils of creatures between fishes and amphibians, and the large differences between the two. But when in 2006 scientists discovered the fossil remains of the 375 million-year-old crocodile-headed giant fish, the *tiktaalik*, which appears to have been precisely such a transitional creature, a spokesperson for the Discovery Institute, which has been aggressively promoting Intelligent Design Theory, professed unconcern: “few leading [Intelligent Design] researchers have argued against the existence of transitional forms.” John Noble Wilford, *Fossil Called Missing Link from Sea to Land Animals*, N.Y. TIMES, Apr. 6, 2006, at A1; *If It Walks Like a Fish . . .*, NEWSWEEK, Apr. 27, 2007, at 8.

<sup>89</sup> See SUSAN HAACK, *DEFENDING SCIENCE—WITHIN REASON: BETWEEN SCIENTISM AND CYNICISM*, at X–XIII, 272–82 (paperback ed. 2007); Susan Haack, *Fallibilism and Faith, Naturalism and the Supernatural, Science and Religion*, in *PUTTING PHILOSOPHY TO WORK: INQUIRY AND ITS PLACE IN CULTURE* 183 (2008).

<sup>90</sup> *Kitzmiller v. Dover Area Sch. Dist.*, 400 F. Supp. 2d 707 (M.D. Pa. 2005); see also EDWARD HUMES, *MONKEY GIRL: EVOLUTION, EDUCATION, AND THE BATTLE FOR AMERICA’S SOUL* (2007) (telling the story of the *Kitzmiller* trial, including Eric Rothschild’s devastating cross-examination of Michael Behe, expert witness for the defendant school district).

<sup>91</sup> In Arthur Hailey’s novel, *Strong Medicine* (1984)—clearly based on the Bendectin saga, but telling the story of a fictional drug company, Felding-Roth, and its fic-

inction between genuine inquiry and advocacy research with which I have been working thus far, while agreeably neat and tidy conceptually, isn't adequate to the complexities of real life; it needs to be reconstrued as identifying the two extremes of a continuum. No investigator can approach his question free of any preconceptions whatever; most investigators have some preconception of the expected upshot from the beginning—though those who really want the truth will change their minds should the evidence demand it; and even the most honest and single-minded investigator is vulnerable to that very natural tendency to duck, resist, or conveniently forget evidence that pulls against the view he has previously defended in print, or against his fond hope that this, finally, will be the key to finding a vaccine, and so on.<sup>92</sup> Figuring things out can be really hard, and the temptation to cut corners is ever-present.

So Intelligent Design “research” is only one example among many, for the sad fact is that inquiry that is not quite plain-and-simple, less than perfectly honest, tainted, if not by outright dishonesty, by convenient self-deception, is ubiquitous. We are all only too familiar with the phenomenon of the “Public Inquiry” the purpose of which is to reassure the public that there is no real danger, or that the corruption is all the fault of one junior official; with the “Customer Survey” the purpose of which is to fish for favorable material the publicity department can use; with the “departmental review” the purpose of which is to get friends from outside to endorse the faculty's grandiose hopes for expansion. We are all aware, also, that in many disciplines—economics, public health, the environmental sciences, to mention just a few—the pressures to nudge inquiry in the direction of advocacy are subtle, and the boundary easily transgressed. And we all know that, even in the disciplines furthest removed from policy or practice, academics often succumb to the temptation to divert energy from finding out what they can, or from

---

tional morning-sickness drug, Montayne—one fictional scientist, Martin Peat-Smith, is a paradigm of the honest inquirer, and another, Vincent Lord, of the self-deceived advocacy researcher. See also Susan Haack, *The Ideal of Intellectual Integrity*, in *Life and Literature*, 36 NEW LITERARY HIST., 359 (2005), reprinted in Haack, *supra* note 89, at 195–208.

<sup>92</sup> Sinclair Lewis's novel, *Arrowsmith* (1924), conveys the point: Martin Arrowsmith destroys the integrity of his test of a vaccine by giving it, out of sympathy with their suffering and hope of curing them, to all those who have been exposed. John Berry's historical study, *The Great Influenza* (2004), illustrates it: scientists desperate to find a vaccine ignored evidence that influenza is not bacterial; only Oswald Avery patiently held out. Dr. Brent, whom we encountered in *Blum*, seems to have been motivated in part by the fear that, with Bendectin off the market, physicians would have no effective treatment for a potentially serious condition.

seriously thinking things through, into efforts to promote their area, their line, or their clique.

There are many kinds of advocacy research, and many sources of bias: some advocacy researchers are too concerned to arrive at a result favorable to a sponsor; some are over-anxious to find a cure quickly; some are too protective of a pet approach or theory, or too deferential to an idea endorsed by a hero of their profession; some get careless out of concern over global warming or pollution, or etc.; some want to reach politically-correct conclusions potentially beneficial to their careers, or to avoid reaching politically-incorrect conclusions potentially damaging to their careers; and many are simply too certain they are right—and so feel entirely justified in suppressing apparently unfavorable evidence which, as they see it, can only be misleading.<sup>93</sup>

\* \* \*

To describe research as “litigation-driven” may mean either (a) that the need for this work arises out of litigation, or (b) that the work is undertaken for the purpose of finding evidence favoring one side in litigation, and explaining away or otherwise playing down evidence favoring the other side. Research which is litigation-driven in sense (a) may, but need not, also be litigation-driven in sense (b). Research which is litigation-driven in the first sense is not peculiarly susceptible to bias merely by virtue of being, in this sense, litigation-driven. But research which is litigation-driven in the second sense is (one kind of) advocacy research; and so, if my analysis is correct, is inherently in danger of bias.

This danger is mitigated somewhat if advocacy research rests on science which has non-judicial as well as judicial uses, but it is not completely averted. Think of DNA identifications: the underlying

---

<sup>93</sup> William McBride, the Australian physician who first drew attention to the teratogenic effects of Thalidomide, was apparently so distressed at the delay before his warnings about Thalidomide were heeded that when, subsequently, he began to suspect Bendectin (sold in Australia under the name “Debendox”) of causing birth defects, he resorted to fraud in his study of pregnant rabbits given the related anticholinergic Scopolamine. Before the fraud was revealed, Dr. McBride had testified for the plaintiffs in seventeen Bendectin cases. Sanders, *supra* note 23, at 36. See A. Skolnick, *Key Witness Against Morning Sickness Drug Faces Scientific Fraud Charges*, 263 J. AM. MED. ASS’N 1468 (1990); G.F. Humphrey, *Scientific Fraud: The McBride Case*, 32 MED. SCI. L. 199 (1992); G.F. Humphrey, *Scientific Fraud: The McBride Case—Judgment*, 34 MED. SCI. L. 299 (1994). Scopolamine is now marketed in the form of a patch as an anti-nausea drug, under the name “Transderm Scop.” RXList.com, Clinical Pharmacology, [http://www.rxlist.com/cgi/generic2/transscop\\_cp.htm](http://www.rxlist.com/cgi/generic2/transscop_cp.htm) (last visited Feb. 20, 2006).

theoretical principles are deeply interconnected with a whole range of other areas of well-established science, and these techniques are used, for example, to identify disaster victims as well as to identify the perpetrators of crimes. The theory is about as solid as scientific theory gets. But it's not the underlying principles that are disputed at trial; courts are not (by now, anyway) trying to determine whether these principles are sound, but whether they have been reliably applied in this instance. There is plenty of room for bias to creep into the application of even the soundest science.

Research may be prompted by the needs of a particular case, or class of cases; or it may be prompted, not by cases already ongoing, but by the fear that there will, or may, be litigation. Moreover, there is very often more than one motive for conducting research, which may, for example, be intended to make the case for FDA approval, to be useful for marketing purposes, and to provide protection against possible litigation. Obviously enough, besides the hope of prevailing in litigation, some of these other motives—the marketing-oriented, for example—are also likely to introduce bias.

Since *Daubert* gives this concept a crucial role, we also need to give some thought to what it means to describe scientific testimony as “reliable.” *Merriam-Webster's* definition is: “suitable or fit to be relied on[;] . . . giving the same results in successive trials”;<sup>94</sup> the *Oxford English Dictionary's* is: “may be relied upon, of sound & consistent character or quality.”<sup>95</sup> Unless it is intended to be read disjunctively, *Webster's* definition seems a little odd, for the second clause seems to allow that a procedure or technique may be reliable even though it usually gives false results, provided it does so consistently—which hardly seems compatible with fitness to be relied on. (A weighing machine that consistently takes fifty pounds off a person's real weight, or a clock that runs perfectly but was set to the wrong time to begin with, are not, in the ordinary sense of the term, reliable; though I suppose you might describe them as, though “off,” at least reliably off.) The *Oxford English Dictionary's* definition, by virtue of its reference to the “soundness” of the results, is closer to my understanding of the word.

Not unexpectedly, however, the legal concept of reliability articulated in *Daubert* diverges somewhat from the ordinary sense; as Justice Blackmun's phrase “evidentiary reliability” signals, it is a spe-

---

<sup>94</sup> MERRIAM-WEBSTER'S COLLEGIATE DICTIONARY 1051 (11th ed. 2003).

<sup>95</sup> THE CONCISE OXFORD DICTIONARY OF CURRENT ENGLISH 1028 (4th ed. 1959).



cialized legal concept. It is also far from transparent. Justice Blackmun writes:

We note that scientists typically distinguish between “validity” (does the principle support what it purports to show?) and “reliability” (does application of the principle produce consistent results?). Although “the difference between accuracy, validity, and reliability may be such that each is distinct from the other by no more than a hen’s kick,” our reference here is to evidentiary reliability—that is, trustworthiness. In a case involving scientific evidence, *evidentiary reliability* will be based upon *scientific validity*.<sup>96</sup>

This tells us that the legal or “evidentiary” concept of reliability is to be tied to scientific “validity,” not to scientific “reliability”; which seems to mean, in part, that yielding consistent results (which is Justice Blackmun’s understanding of “scientific reliability”) is not enough. The reference to “trustworthiness” points in the same direction: “evidentiary reliability” requires scientific testimony to be based on methods and processes that yield “sound,” and not merely consistent, results. But Justice Blackmun’s understanding of “sound” is apparently quite modest; it does not require that the principle on which expert testimony is based yield true or even probably true results, but only that “the principle support[s] what it purports to show.”<sup>97</sup>

The fact that research is litigation-driven in the stronger sense, I have argued, makes it likely to be biased. Biased research doesn’t necessarily produce false results; nor does it necessarily produce false results more often than true. After all, the proposition(s) toward which it is slanted may be true; and when there is biased research on both sides of a legal case, if the propositions on each side genuinely contradict each other, the proposition(s) toward which one side’s research is slanted must be true. But biased research tends toward the predetermined conclusion irrespective of where the evidence points; the results it produces don’t depend on where the evidence really leads. So if this is, as it seems to be, a reasonable interpretation of the *Daubert* Court’s “evidentiary reliability,” then, indeed, biased research is unreliable in the relevant sense.

---

<sup>96</sup> *Daubert v. Merrell Dow Pharms., Inc. (Daubert III)*, 509 U.S. 579, 590 n.9 (1993) (citations omitted). The internal quotation is from James Starrs, *Frye v. United States Restructured and Revitalized: A Proposal to Amend Federal Rule of Evidence 703*, 26 JURIMETRICS J. 249, 256 (1986).

<sup>97</sup> *Daubert III*, 509 U.S. at 590 n.9.

IV. THROUGH THE THICKET, OUT OF THE  
SWAMP, AND ONTO THE HIGH ROAD? NOT YET!

So there is some foundation for Judge Bernstein's strictures against "justification science"; indeed, his observation that Dr. Shapiro's work seems to have been based on an "unalterable preconception" that Bendectin was harmless closely parallels the argument here, that science that is litigation-driven in the stronger sense fails to meet Justice Blackmun's standard of evidentiary reliability because the conclusions drawn are not sensitive to the evidence in the way they ought to be. And, again provided that "litigation-driven" is understood in the stronger sense, there is some foundation, also, for Judge Kozinski's conclusion that the fact that testimony is based on litigation-driven research speaks negatively to its (evidentiary) reliability.

However, there is something amiss with Judge Kozinski's arguments for that conclusion. His first argument, remember, is that science flowing from existing research is less likely to be biased toward a particular conclusion by the promise of remuneration; this is true, but it proves much more than he intends. Many studies confirm that company-sponsored research into drugs or medical devices is significantly more likely than independent research to be favorable to the sponsor's product;<sup>98</sup> but this suggests, not just that litigation-driven science may be below par, but also that marketing-oriented science should also be regarded with suspicion.

Moreover, this first argument also undermines the exception Judge Kozinski makes with regard to evidence from the forensic sciences. It is true, as he says, that the fact that forensic scientists acquire their expertise for the purposes of the justice system isn't in itself grounds for doubting the reliability of their testimony; but this is not enough to establish his point. Perhaps the thought implicit here is that forensic science is litigation-driven only in the weaker, less troubling sense: that while it is needed only because there are crimes to be solved and prosecuted, it is not inherently motivated by the desire to make one side of a case; but this is Pollyannish to say the least. After all, such work is undertaken almost exclusively for the police or

---

<sup>98</sup> See e.g., Richard A. Davidson, *Source of Funding and Outcome of Clinical Trials*, 1 J. GEN. INTERNAL MED. 155 (1986); Paula Rochon et al., *A Study of Manufacturer-Supported Trials of Non-Steroidal Anti-Inflammatory Drugs in the Treatment of Arthritis*, 154 ANNALS INTERNAL MED. 157 (1994); Lee S. Friedman & Elihu D. Richter, *Relationship Between Conflict of Interest and Research Results*, 19 J. GEN. INTERNAL MED. 51 (2004).

prosecution;<sup>99</sup> and it seems likely that forensic scientists' and technicians' understandable but inappropriate desire to be helpful, to find something to make a case against a suspect, sometimes biases their judgment. Or perhaps the thought is that forensic experts will curb their biases because they know they will be called on to testify on numerous other occasions;<sup>100</sup> but this seems no less doubtful. After all, the fact that expert witnesses in tort cases are "repeat testers" or "professional expert witnesses," as we say pejoratively, is often seen, not without reason, as grounds for distrusting them.<sup>101</sup> Judge Bernstein's worry that "general acceptance in the field to which it belongs"<sup>102</sup> is a poor indicator of reliability if the consensus is an artificial one is also relevant here; for in some areas of forensic science there is a real danger that a supposed "scientific consensus" has been

---

<sup>99</sup> See e.g. William C. Thompson, *A Sociological Perspective on the Science of Forensic DNA Testing*, 30 U.C. DAVIS L. REV. 1113, 1114 (1997).

The primary clients of the vast majority of forensic scientists are law enforcement agencies. Most forensic scientists are employed directly by law enforcement agencies. Their role in litigation is typically, and often exclusively, to provide evidence in support of criminal prosecutions. Forensic scientists who work in private laboratories may occasionally be employed by criminal defense lawyers. However, the bulk of their work is for law enforcement as well. The major market for commercial laboratories that develop new technology for forensic testing also consists of law enforcement personnel.

*Id.*

<sup>100</sup> Dr. Thompson (the author of the article in note 99, *supra*) tells me that this is the reason Judge Kozinski gave him.

<sup>101</sup> In the first *Blum* trial, the court prevented the plaintiff's attorneys from referring to the fact that Merrell Dow's experts had testified in other Bendectin trials. Appendix 6, Order and Opinion of Philadelphia Court of Common Pleas, May 12, 1988, at 28 (D'Alessandro, J.) (on file with author). Judge Kozinski himself suggests that the fact that the Dauberts' proffered experts have been testifying in Bendectin cases all over the country is reason to be suspicious of them. *Daubert v. Merrell Dow Pharms., Inc. (Daubert IV)*, 43 F. 3d 1311, 1314 (9th Cir. 1995). In fact, there were numerous repeat testers on both sides throughout the Bendectin litigation. Sanders, *supra* note 45, at 36. Cf. *Chaulk v. Volkswagen of Am. Inc.*, 808 F.2d 639, 644 (7th Cir. 1986) (Posner J., dissenting) (writing that an expert's testimony was either the work of

a crank, or, what is more likely, of a man who is making a career out of testifying for plaintiffs in automobile accident cases in which a door may have opened; at the time of trial he was involved in 10 such cases. His testimony illustrates the age-old problem of expert witnesses who are "often the mere paid advocates or partisans of those who employ or pay them."

(quoting *Keegan v. Minneapolis & St. Louis R.R.*, 78 N.W. 965, 966 (Minn. 1899)).

<sup>102</sup> *Blum ex rel. Blum v. Merrell Dow Pharms., Inc.*, 33 Phila. Co. Rptr. 193, 256 (Ct. Comm. Pleas Pa. 1996).

generated by a kind of guild or trade union of mutually supportive practitioners with an interest in protecting their livelihoods.<sup>103</sup>

Judge Kozinski's second argument, that litigation-driven science is not, like university science, kept up to the mark by the need to attract funding and institutional support, rests on a false premise. For by now a significant proportion of the medical research in universities is not truly independent, but is sponsored by drug companies and such<sup>104</sup> (and a significant proportion of research in the social sciences is in one way or another politically motivated). And in combination with the first argument, this suggests that there may be reason to doubt the reliability of such university science, as well as science specifically undertaken to support one side or another in litigation, or to provide data that can be used in marketing.

Moreover, the peer-review process for funding and publication, on which Judge Kozinski puts quite a lot of weight, is a frail safeguard at best. Even if all the work published in peer-review journals were peer-reviewed—which it isn't—this would be only very weak assurance of its reliability. As the *Daubert* Court's comments on "peer-review and publication" obliquely acknowledge, it is not peer-reviewed publication as such that indicates reliability, but the long-run survival of published results on which other scientists find they can build successfully.<sup>105</sup>

Still, given that, as I have argued, there is merit in the idea that the fact that science is litigation-driven in the stronger sense indicates that it is likely to be unreliable, in something like the sense Justice Blackmun explained in *Daubert*, might this not be a helpful factor to be added to his list of indicia of (un)reliability? Unfortunately, matters are not so simple; for the sad fact—obvious once you think about it—is that there *can be no* simple, mechanically applicable test that would accurately discriminate strong science from weak. The *Daubert* Court observes that its list of indicia of reliability is "flexible," and can't simply be applied mechanically. And a mechanical application of Judge Kozinski's new *Daubert* factor would certainly be as ill-advised

---

<sup>103</sup> See e.g., D. Michael Risinger, Mark P. Denbeaux & Michael J. Saks, *Brave New "Post-Daubert World"—A Reply to Professor Moenssens*, 29 SETON HALL L. REV. 405 (1998). The knife-mark examiners in *Ramirez v. State*, 542 So. 2d 352 (Fla. 1989), *Ramirez v. State*, 651 So. 2d 1164 (Fla. 1995), and *Ramirez v. State*, 810 So. 2d 836 (Fla. 2001) illustrate the problem.

<sup>104</sup> See e.g., SHELDON KRIMSKY, *SCIENCE IN THE PRIVATE INTEREST* (2003); Marcia Angell, *Is Scientific Medicine for Sale?*, 342 NEW ENG. J. MED. 1516 (2000).

<sup>105</sup> Brief for Chubin et al. as Amici Curiae Supporting Petitioners, *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579 (1993) (No. 92-102). See also Susan Haack, *Peer Review and Publication: Lessons for Lawyers*, 36 STETSON L. REV. 789 (2007).

as a mechanical application of a requirement that testimony be based on research that has been published in peer-reviewed journals—and for structurally similar reasons: “peer-reviewed” and “litigation-driven” both have (a) a readily-applicable sense that has little to do with reliability, and (b) a subtler sense which bears more closely on reliability but isn’t readily applicable. “Published after peer review” is easily applied, but it is a frail indicator of reliability; “has been out there long enough, has been read by enough others knowledgeable enough in the field, links up in an explanatory way with enough other bits of scientific theorizing, and has proven robust enough when new experiments or theoretical work assume its reliability” is a much better indicator of reliability, but it is much more difficult to apply.<sup>106</sup> Similarly, “undertaken in the course of or in anticipation of litigation” is easily applied, but a frail indicator of reliability; “skewed by the desire to advance one side in litigation” indicates unreliability, all right, but is much more difficult to apply.

\* \* \*

I haven’t forgotten that the epistemological rationale for the adversarial system is that having rival advocates each present the evidence favoring their side of a case is a good way to ensure, so far as possible, that the truth comes out. As I have argued elsewhere, the best argument that could be made for the epistemological efficacy of such a system would run something like this:

Since for good reasons the legal process, unlike the process of scientific inquiry, has to be concluded within a relatively short time frame, we need a way of ensuring that the search for and scrutiny of evidence is as thorough as that time frame allows. An adversarial system is one way to do this. If everyone involved knows that eventually, at the trial stage, the determination will be made by an impartial jury weighing the evidence developed and presented by the parties, each subject to cross-examination by the other, this should encourage precisely the kind of thoroughness we are aiming to achieve. For an advocate’s goal is to win; so counsel for each party is motivated to seek out evidence favoring his side of the case, and to bring out the flaws in evidence pointing the other way. To be sure, the process isn’t perfect; but it is a reasonable substitute for the ideal.<sup>107</sup>

---

<sup>106</sup> See Haack, *supra* note 105.

<sup>107</sup> Haack, *supra* note 84, at 51.

This is a good argument—in principle; however, it is a serious question how well it applies to our adversarial system, as that functions in practice.

A quite general problem is that there is often a vast asymmetry between the resources available to one side in litigation and those available to the other. And where scientific testimony is concerned there are further problems as well. In 1901, Judge Learned Hand complained that the expert-witness system “set[s] juries to decide, where doctors disagree”;<sup>108</sup> more than a century later, Justice Rehnquist and Judge Kozinski complain, in effect, that now it “set[s] judges to decide, where doctors disagree.”<sup>109</sup> The fact is that judges, jurors, or attorneys, however conscientious and thorough, probably don’t fully understand scientific testimony; to make matters worse, the more an area of science gets entangled with litigation, the more scientists in that area seem (like Dr. Brent and Dr. Newman) to fall into advocacy mode. And this makes the difficult business of getting at the truth of the questions at issue even harder than it would otherwise be—which is presumably what Justice Castille had in mind when he expressed concern about “the litigation-driven Benedictin ‘scientific community’ described to the court” in *Blum*.<sup>110</sup>

In the criminal justice system, besides a troubling asymmetry between the scientific resources ordinarily available to the defense and those available to the prosecution, there seem to be grounds for concern both that, in some areas of forensic science, a self-serving guild mentality may predominate over the scientific attitude, and that courts are reluctant to reconsider their long-standing reliance on identification techniques such as fingerprinting (about the reliability of which much is claimed, but little seems to be known)<sup>111</sup> or psychi-

---

<sup>108</sup> Learned Hand, *Historical and Practical Considerations Regarding Expert Testimony*, 15 HARV. L. REV. 40, 54 (1901).

<sup>109</sup> “I do not doubt that Rule 702 confides to the judge some gatekeeping responsibility in deciding questions of the admissibility of proffered expert testimony. But I do not think it imposes on them either the obligation or the authority to become amateur scientists in order to perform that role.” *Daubert v. Merrell Dow Pharms., Inc. (Daubert III)*, 509 U.S. 579, 600–01 (1993) (Rehnquist, J., dissenting). “[W]e are largely untrained in science and certainly no match for any of the witnesses whose testimony we are reviewing.” *Daubert v. Merrell Dow Pharms., Inc. (Daubert IV)*, 43 F.3d 1311, 1316 (9th Cir. 1995).

<sup>110</sup> *Blum ex rel. Blum v. Merrell Dow Pharms., Inc.*, 764 A.2d 1, 13–14 (Pa. 2000).

<sup>111</sup> See, e.g., Simon Cole, *What Counts for Identity? The Historical Origins of the Methodology of Latent Fingerprint Identification*, 12 SCI. CONTEXT 139 (1999); Robert Epstein, *Fingerprints Meet Daubert: The Myth of Fingerprint “Science” is Revealed*, 75 S. CAL. L. REV. 605 (2002); Jennifer Mnookin, *Fingerprints: Not a Gold Standard*, ISSUES SCI. & TECH., Fall 2003, at 47; Simon Cole, *Grandfathering Evidence: Fingerprint Admissibility from Jennings to Llera Plaza and Back Again*, 41 AM. CRIM. L. REV. 1189 (2004); Sharon Be-

atric techniques such as predictions of future dangerousness.<sup>112</sup> And in the civil arena, toxic-tort and product-liability litigation seems like a kind of lottery, where it is hard to feel confident either that all and only those plaintiffs who really were injured by defendants' products are compensated, or that the system provides effective incentives to manufacturers to investigate their products as thoroughly as possible.

In any case, while compensating the victims of dangerous products after the damage has been done, insofar as such compensation is possible, is better than nothing, it is hardly the ideal. It would be better, surely, to ensure so far as humanly possible that safe and beneficial drugs, devices, chemicals, etc., are available, but dangerous or damaging drugs, etc., are kept off the market, or taken off the market as soon as the dangers are known, and that manufacturers are discouraged from hiding or disguising risks posed by their products. Policy proposals are not exactly a philosopher's forte, but I will venture to ask some of the tough questions that the story of *Daubert* and *Blum* prompts in my mind. Do we rely too much on what Justice Breyer describes as "the powerful engine of tort litigation,"<sup>113</sup> ideally the last resort? Are other technologically advanced countries where the engine of tort litigation is less powerful, invariably less successful, also, in keeping beneficial products on, and dangerous products off the market? Are other countries' regulatory agencies more effective, and if so, why? Might some of the energy now devoted to discussions of how best to fine-tune the rules of admissibility of expert testimony be more profitably diverted to thinking about other and possibly better ways to approximate the ideal more closely? And (perhaps you will think this a naïve question, but I'll ask it anyway): what if the time, energy, intelligence, and resources spent on cases like *Daubert* and *Blum* had been spent instead on independent, honest, solid scientific investigation of the factual issues?

---

gley, *Despite Its Reputation, Fingerprint Evidence Isn't Really Infallible*, WALL ST. J., June 4, 2004, at B1.

<sup>112</sup> See, e.g., Erica Beecher-Monas & Edgar Garcia-Hill, *Genetic Predictions of Future Dangerousness: Is There a Blueprint for Violence?*, 69 LAW & CONTEMP. PROBS. 201 (2006); Thomas Regnier, *Barefoot in Quicksand: The Future of Future Dangerousness Predictions in Death Penalty Sentencing in the World of Daubert and Kumho*, 37 AKRON L. REV. 467 (2004).

<sup>113</sup> Justice Breyer writes that courts' gatekeeping can help ensure that "the powerful engine of tort liability, which can generate strong financial incentives to reduce, or eliminate production, points towards the right substances and does not destroy the wrong ones." *G.E. v. Joiner*, 522 U.S. 136, 148–49 (1997) (Breyer, J., concurring). Of course, it spoils the effect somewhat that the substance in question in the case was PCBs (polychlorinated biphenals), so dangerous that they have been banned since 1978!