

RESEARCH SUBPOENAS AND THE SOCIOLOGY OF KNOWLEDGE

SHEILA JASANOFF*

I

INTRODUCTION

Like Moliere's Monsieur Jourdain on learning that he had been speaking prose for forty years, most U.S. judges would probably be astonished to find that their everyday assessments of scientific evidence recapitulate some of the most profound contemporary debates in the philosophy and sociology of scientific knowledge. In an increasing variety of lawsuits, judges are confronting, and innocently "resolving," epistemological questions that have perplexed academic analysts of science for many decades.¹ What is the nature of "truth" and "validity" in science? What makes one disputed scientific claim "better" than competing interpretations and points of view? When is the science invoked by parties to a lawsuit sufficiently reliable to be heard in court, let alone to justify the imposition of serious financial liability or restraints on personal liberty? And when conflicts occur between alternative approaches to interpreting scientific information, who should decide which approach is better?

Questions like these have arisen in remarkably diverse factual contexts, and they tax judicial understandings of how scientific research is done. For example, the landmark environmental controversies of the 1970s and 1980s invited courts to compare legal and scientific standards of certainty and to give administrative agencies a reasoned basis for acting as they did when there was no consensus in science to guide them.² Toxic tort claims pitted medical and scientific experts against each other, requiring courts to assess their relative credibility as well as their competence to testify on matters relevant to determinations of liability.³ Requests to exclude various types of novel scientific evidence forced courts to act in effect as gatekeepers for scientific methodologies and

Copyright © 1996 by Law and Contemporary Problems

*Professor, Department of Science and Technology Studies, Cornell University.

1. Particularly influential works in this genre include THOMAS KUHN, *THE STRUCTURE OF SCIENTIFIC REVOLUTIONS* (1962); KARL V. POPPER, *OBJECTIVE KNOWLEDGE: AN EVOLUTIONARY PERSPECTIVE* (1972).

2. See, e.g., *Baltimore Gas and Elec. Co. v. Natural Resources Defense Council, Inc.*, 462 U.S. 87 (1983); *Gulf South Insulation v. United States Consumer Prods. Safety Comm'n*, 701 F.2d 1137 (5th Cir. 1983); *Ethyl Corp. v. Env'tl. Protection Agency*, 541 F.2d 1 (D.C. Cir. 1976).

3. Representative cases over more than a decade include *In re Paoli RR Yard PCB Litig.*, 35 F.3d 717 (3d Cir. 1994); *Christopherson v. Allied-Signal Corp.*, 939 F.2d 1106 (5th Cir. 1991); *In re Paoli RR Yard PCB Litig.*, 916 F.2d 829 (3d Cir. 1990), cert. denied sub nom. *General Elec. Co. v. Knight*, 499 U.S. 961 (1991); *Sterling v. Velsicol Chem. Corp.*, 24 ERC 2017 (W.D. Tenn. 1986); *In re Agent Orange Prod. Liab. Litig.*, 611 F. Supp. 1223 (E.D.N.Y. 1985).

practices of uncertain reliability.⁴ In *Daubert v. Merrell Dow Pharmaceuticals, Inc.*,⁵ the Supreme Court attempted to systematize judicial thinking on these issues with a new rule on the admissibility of scientific evidence.

The research subpoena cases discussed in this symposium present similar challenges, but perhaps in even sharper form. In evaluating research subpoenas, courts regularly have dealt with questions concerning the relevance and reliability of the requested materials,⁶ the balance between the requesting party's interest in disclosure and the researcher's interest in confidentiality,⁷ the reasonableness of the burden on the researcher or the researcher's institution,⁸ and the possible right of experts to be protected against legal intrusions.⁹ There is, however, an epistemological dimension to research subpoenas that to this point has received relatively little attention from judges and legal analysts.¹⁰ Once materials are obtained by subpoena, are courts institutionally competent to take the next step beyond gatekeeping? That is, can they evaluate both the probity and scientific validity of these materials and their adequacy as a foundation for challenged evidence?

Advocates contend that subpoenaed research information allows the requesting party, and in due course the legal fact finder, to determine whether the reported scientific results were properly derived from the researcher's underlying acts of observation, recording, and interpretation.¹¹ But this attempt to reassess in a legal forum what scientists know, and how they come to know it, threatens to put both judges' and attorneys' lay conceptions of "the scientific method" on a collision course with scientists' actual practices. Drawing on re-

4. See *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 43 F.3d 1311 (9th Cir. 1995) (on the statistical technique of "meta-analysis"); *People v. Ojeda*, 225 Cal. App. 3d 404 (1990) (on "horizontal gaze nystagmus," a field test for drunk driving); *People v. Castro*, 545 N.Y.S.2d 985 (1989) (on "DNA typing").

5. 509 U.S. 579 (1993).

6. See, e.g., *Dow Chem. Co. v. Allen*, 672 F.2d 1262 (7th Cir. 1982); *Andrews v. Eli Lilly & Co.*, 97 F.R.D. 494 (N.D. Ill. 1983), *vacated sub nom. Deitchman v. E.R. Squibb & Sons, Inc.*, 740 F.2d 556 (7th Cir. 1984).

7. See, e.g., *Deitchman*, 740 F.2d at 559; *Farnsworth v. Procter & Gamble Co.*, 101 F.R.D. 355, 357 (N.D. Ga. 1984), *aff'd*, 758 F.2d 1545 (11th Cir. 1985); see also Elizabeth C. Wiggins & Judith A. McKenna, *Researchers' Reactions to Compelled Disclosure of Scientific Information*, 59 LAW & CONTEMP. PROBS. 67, 88-91 (Summer 1996).

8. See, e.g., *In re Application of Am. Tobacco Co.*, 880 F.2d 1520, 1523 (2d Cir. 1989); *Dow Chemical*, 672 F.2d at 1267; *Anker v. G.D. Searle & Co.*, 126 F.R.D. 515, 518-19 (M.D.N.C. 1989); *In re Application of R.J. Reynolds Tobacco Co.*, 518 N.Y.S.2d 729, 732 (1987).

9. See, e.g., *In re Grand Jury Subpoena* dated Jan. 4, 1984, 583 F. Supp. 991 (E.D.N.Y.), *rev'd and remanded*, 750 F.2d 223 (2d Cir. 1984); *Wright v. Jeep Corp.*, 547 F. Supp. 871 (E.D. Mich. 1982); see also Paul D. Carrington & Traci L. Jones, *Reluctant Experts*, 59 LAW & CONTEMP. PROBS. 51 (Summer 1996); Michael Traynor, *Countering the Excessive Subpoena for Scholarly Research*, 59 LAW & CONTEMP. PROBS. 119, 120 (Summer 1996).

10. On possible conflicts between the judicial interest in open discovery and scientists' interest in controlling who should evaluate science, see SHEILA JASANOFF, *SCIENCE AT THE BAR: LAW, SCIENCE, AND TECHNOLOGY IN AMERICA* 98-99 (1995).

11. See, e.g., Eliot Marshall, *Court Orders "Sharing" of Data*, 261 SCI. 284 (1993). For this purpose, researchers may be asked to produce any or all of the following types of documents: unpublished working papers; research protocols; written commentaries on or criticisms of research designs; documents pertaining to peer review of protocols and findings; questionnaires, letters, and interview forms; responses to interviews; field notes; data on causes of death; and data sheets or computer disks recording raw data. See *In re Application of Am. Tobacco Co.*, 880 F.2d at 1523.

cent research in the sociology of science, this article seeks to identify the points at which disjunctions between judges' and scientists' perceptions of science are most likely to occur and to propose methods by which courts can reduce the risk of judging science by ad hoc, arbitrary, and inappropriate standards.

II

WHY SUBPOENA RESEARCH RECORDS?—SOME COMMON ASSUMPTIONS

Requests for the records on which scientists have based their published findings may be motivated by both legitimate and illegitimate considerations. Litigants may try to use the research subpoena to intimidate and overburden researchers or to introduce irrelevant and confusing issues into the fact-finding process.¹² As discussed elsewhere in this symposium, vigilance by judges and researchers offers the best protection against such unjustified "fishing expeditions."¹³

More serious and conceptually more interesting are the cases in which litigants expect to find genuine discrepancies between the "facts" reported by scientists and the observations on which their findings were based. The expectation that research subpoenas will enable courts to distinguish between valid and invalid scientific claims appears to rest on certain widely held but empirically questionable assumptions about the practice of science:

- that scientific records are kept in accordance with standardized and generally accepted rules;
- that evidence of fraud, error, and poor scientific practice can be detected unambiguously from written records;
- that challenged methodologies (including methods of analysis and interpretation) clearly conform or do not conform to "scientific" standards;
- that such standards preexist and hence can be mechanically applied during legal inquiry; and
- that the adversarial process is conducive to sorting out disputes concerning the validity and reliability of competing research practices.

In fact, as seen below, each of these assumptions is at odds with well-established findings in the sociology of science. They are further contradicted by the experience of legal institutions in evaluating novel scientific evidence and investigating claims of scientific misconduct.

12. Abuses of the discovery process are most likely to occur in high-stakes mass tort cases. See MARCIA ANGELL, *SCIENCE ON TRIAL: THE CLASH OF MEDICAL EVIDENCE AND THE LAW IN THE BREAST IMPLANT CASE 142-46* (1996) (describing a research subpoena by a plaintiff's attorney for documents of a type that did not exist, such as evidence of improper financial dealings between herself and implant manufacturers); Gina Kolata, *Legal System and Science Come to Differing Conclusions on Silicone*, N.Y. TIMES, May 16, 1995, at D6 (describing "staggering" demand for documents served on the principal author of article on breast implants).

13. See, e.g., Barbara B. Crabb, *Judicially Compelled Disclosure Of Researchers' Data: A Judge's View*, 59 LAW & CONTEMP. PROBS. 9, 26-32 (Summer 1996); Michael Traynor, *supra* note 9, at 120.

III

THE CONSTRUCTION AND DECONSTRUCTION OF SCIENCE

Although scientific knowledge and discovery are commonly represented as the product of the genius-scientist's lonely, heroic struggles with nature, recent scholarship accords greater prominence to the mundane and collective aspects of scientific activity. Through detailed descriptions of laboratory work and other forms of scientific practice, researchers have shown that the production of facts and bodies of knowledge is invariably embedded in a matrix of social interactions, agreements, and understandings.¹⁴ Scientific knowledge, according to these studies, must be communally certified to be legitimate; it must be subscribed to by groups of like-minded investigators, not asserted as indisputable truth by isolated individuals. Negotiation within and among scientific communities establishes the conventions for acceptable research and determines whether a given scientist's work falls within or outside the boundaries of good science.¹⁵ These observations are frequently summarized in the statement that science is socially constructed.¹⁶

The model of social construction has several important implications for legal fact-finding. First, as seen from this perspective, the scientific method looks neither singular nor static, but rather is continually renegotiated within particular research communities so as to accommodate new observations and technical capabilities. Comparison across scientific sub-disciplines shows that even the most basic principles of the scientific method, such as the use of experimental controls, may vary to some degree from one area of science to another. For example, though mixing historical and contemporaneous controls may be strictly forbidden in molecular biology, it may be accepted practice when conducting studies of drug safety and efficacy on human or animal populations.¹⁷ Scientific facts are therefore most authoritative when they have been generated in accordance with the established conventions of some well-defined field or discipline. At the same time, there are no abstract, universally applicable standards against which the validity of scientific facts can be tested.

Second, in evolving areas of science it may take time and considerable discussion among researchers to reach a consensus on the validity of new research methodologies and criteria of evaluation. Thus, in his classic analysis of how the Wassermann test for syphilis gained acceptance, the German biologist Ludwik Fleck described the emergence of a scientific "thought collective" that

14. For a helpful review of this literature, see Steven Shapin, *Here and Everywhere: Sociology of Scientific Knowledge*, 21 ANNUAL REV. OF SOC. 289 (1995). See also H.M. COLLINS & TREVOR PINCH, *THE GOLEM: WHAT EVERYONE SHOULD KNOW ABOUT SCIENCE* (1993); BRUNO LATOUR, *SCIENCE IN ACTION: HOW TO FOLLOW SCIENTISTS AND ENGINEERS THROUGH SOCIETY* (1987).

15. SHEILA JASANOFF, *THE FIFTH BRANCH: SCIENCE ADVISERS AS POLICYMAKERS* 61-83 (1990).

16. See Sheila Jasanoff, *What Judges Should Know about the Sociology of Science*, 32 JURIMETRICS 345 (1992).

17. See JASANOFF, *supra* note 15, at 75-76 (illustrating differences among scientific specialties, and individual practitioners, on the use of controls).

eventually ratified the test's efficacy.¹⁸ Until such a consensus is stabilized, there may be genuine doubt and a legitimate plurality of thought about the validity of particular research approaches. In reflecting on the history of the Wassermann test, Fleck concluded that

[t]he following facts are therefore firmly established and can be regarded as a paradigm of many discoveries. *From false assumptions and irreproducible initial experiments an important discovery has resulted after many errors and detours.* The principal actors in the drama cannot tell us how it happened, for they rationalize and idealize the development. Some among the eyewitnesses talk about a lucky accident, and the well-disposed about the intuition of a genius. It is quite clear that the claims of both parties are of no scientific value.¹⁹

Fleck's solution to this paradox was to stress the non-individualistic character of valid scientific knowledge, a solution that has gathered strength from subsequent work in the sociology of science.²⁰ Moreover, the knowledge that a "thought collective" or scientific subculture holds in common may be impossible to reduce to formal principles; rather, it may consist in large part of tacit knowledge, experience, and skill.²¹ If this is the case, only those who belong to the relevant subculture may be truly in a position to evaluate each other's competence. Outsiders to the subculture may not be able to discriminate as readily between skilled and unskilled research practices in that field.

Third, the model of social construction highlights the role of prior negotiation and consensus building in the determination of whether an experiment or a study has been successfully replicated. Agreement is required to establish that one result is "the same" as another, because absolute identity in experimental conditions or in results can never be achieved. Experiments, moreover, invariably are matters of skill; talented experimentalists are sometimes said to have "golden hands." Therefore, failure to obtain "the same" results as those in a prior experiment can always be attributed to the second investigator's lack of competence rather than to the mistakes or misinterpretations of the first. Disputes about replication—and indeed about scientific validity more generally—often degenerate into what H.M. Collins has termed "experimenter's regress," a potentially endless series of questions designed to probe every aspect of the "failed" experiment, from the reliability of the instruments used to the honesty of the researchers.²² Such disputes are most likely to arise in novel or developing, and hence relatively unstandardized, fields of science.

Standardization of research protocols and evaluation criteria is one widely accepted safeguard against "experimenter's regress." If researchers can agree

18. LUDWIK FLECK, *GENESIS AND DEVELOPMENT OF A SCIENTIFIC FACT* (1979) (originally published as *DIE ENTSTEHUNG UND ENTWICKLUNG EINER WISSENSCHAFTLICHEN TATSACHE* (1935)).

19. *Id.* at 76 (emphasis in original). For an ethnographic approach to scientific fact-making, see BRUNO LATOUR & STEVE WOOLGAR, *LABORATORY LIFE: THE CONSTRUCTION OF SCIENTIFIC FACTS* (1979).

20. See SHEILA JASANOFF ET AL., *HANDBOOK OF SCIENCE AND TECHNOLOGY STUDIES* (1995).

21. H.M. COLLINS, *The TEA Set: Tacit Knowledge and Scientific Networks*, 4 *SCI. STUD.* 165 (1974).

22. H.M. COLLINS, *CHANGING ORDER: REPLICATION AND INDUCTION IN SCIENTIFIC PRACTICE* 2 (1985) (describing experimenter's regress); see also H.M. COLLINS, *The Seven Sexes: A Study in the Sociology of a Phenomenon, or the Replication of Experiments in Physics*, 9 *SOC.* 205 (1975).

in advance on what significance they will attach to variations in controls or other background conditions, they may prevent subsequent conflicts over discrepancies in the results obtained by different researchers conducting "the same" experiment. Standardization only works, however, if everyone concerned agrees to play by the designated rules. When scientific results are disputed, it is frequently because dissenters from within a given scientific community, or observers from outside it, have begun to question their colleagues' methodological compromises or tacit assumptions.

Scientific claims are especially prone to being pulled apart, or deconstructed, when they are used to justify significant legal or political decisions. Skepticism is institutionalized in these settings. The familiar courtroom "battle of experts" is made possible, in part, by the fact that consensus in science is usually provisional and hardly ever all-inclusive. An expert witness who happens not to subscribe to the governing consensus can sow reasonable doubt about the reliability of particular scientific practices, the adequacy or completeness of specific interpretive frameworks, the comparative saliency of different kinds of evidence, and the credibility of competing experts.²³ In short, litigation presents not so much a contest between "true" and "false" beliefs²⁴ as a test of the strength and unanimity of the prevailing consensus.

IV

STANDARDS AS A RESPONSE TO SKEPTICISM

Contrary to the expectations of many judges, standards for conducting valid scientific research in a given field may not exist or may not have been firmly codified before a legal controversy begins. In such cases, the legal dispute itself becomes an occasion for debating competing scientific claims and may indirectly promote standardization. Past controversies surrounding the exclusion of DNA typing evidence present one example of such a process.

The lack of standardized methods for interpreting DNA evidence became readily apparent in the 1989 sexual molestation case of *Maine v. McLeod*. Lifecodes Corporation, a commercial testing laboratory, conducted scientific analysis of the evidence in the case and claimed that it had identified a match between separate DNA samples taken from the suspect and the crime scene.²⁵ Such determinations are ordinarily made on the basis of a pattern of bands resembling a supermarket bar code (the so-called "fingerprint") that is derived through the process of gel electrophoresis.²⁶ Although Lifecodes asserted that

23. See generally EXPERT EVIDENCE: INTERPRETING SCIENCE IN THE LAW (Roger Smith & Brian Wynne eds., 1989).

24. This rather reductionist view of scientific controversies in the courtroom has, of course, gained wide currency among both legal practitioners and the public. See, e.g., PETER W. HUBER, GALILEO'S REVENGE: JUNK SCIENCE IN THE COURTROOM (1991).

25. See Colin Norman, *Maine Case Deals Blow to DNA Fingerprinting*, 246 SCI. 1556 (1989).

26. To obtain a DNA "fingerprint," long molecules of DNA are first cut into fragments of varying lengths with the use of special restriction enzymes. The fragments are separated by size through a

it had found a match, the bands from the two samples did not line up perfectly. The distribution of the bands was the same in both prints, but the entire pattern was displaced in a way that suggested the DNA fragments from one sample were slightly larger and heavier than those in the other, and thus had traveled shorter distances through the electrically charged gel. On what basis, then, had Lifecodes declared the two prints identical?

Experts from Lifecodes revealed at trial that they had employed a previously unreviewed and unpublished methodology to reconcile the differences between the two DNA prints.²⁷ A device known as a monomorphic probe had been used to tag a particular fragment of DNA that is the same in every person. Extrapolating from the relative displacement of this tagged fragment, the Lifecodes experts had concluded that all of the bands in the seemingly displaced sample should be corrected by a factor of 3.15%, a calculation that enabled them to treat the bandshifting as immaterial.

Subsequent discussion among scientists both inside and outside the courtroom showed that the technique used by Lifecodes was far from generally accepted.²⁸ Nor was there, as yet, a generally accepted alternative on which most molecular biologists could agree. Once their interest was awakened, researchers not only questioned Lifecodes's methods and tacit assumptions but also began proposing a variety of techniques that they asserted would work better. One defense expert stated, "The whole experiment wasn't done with the kind of rigor you would expect."²⁹ An exchange of letters in *Science* carried the "experimenter's regress" still further, pinpointing specific technical options that Lifecodes should have considered and either should have adopted or explicitly ruled out.³⁰

These criticisms underscored the fact that Lifecodes's declaration of a match was contingent upon a particular testing technique, namely the use of monomorphic probes, whose acceptance within a wider scientific community had never been openly debated and was by no means a foregone conclusion. Leading scientists agreed that controversies such as the one that occurred in *Maine v. McLeod* had the beneficial effect of exposing unresolved and important technical issues.³¹ The "black box" of DNA typing was usefully opened up

process called electrophoresis. This involves placing samples of DNA fragments on an electrically charged gel where smaller fragments move faster than larger ones. When the process is concluded, DNA fragments of different lengths congregate in separate bands along the gel. Radioactive probes are then used to obtain a picture of their distribution. See NATIONAL RESEARCH COUNCIL, DNA TECHNOLOGY IN FORENSIC SCIENCE 36-37 (1992).

27. Norman, *supra* note 25, at 1558.

28. *Id.* The standard for the admissibility of scientific evidence at the time of the *Maine* case was that of "general acceptance" in the scientific community. See *Frye v. United States*, 293 F. 1013, 1014 (D.C. Cir. 1923).

29. Norman, *supra* note 25, at 1557.

30. See *Letters: DNA Fingerprinting*, 247 SCI. 1018-19 (1989).

31. Eric S. Lander & Bruce Budowle, *DNA Fingerprinting Dispute Laid to Rest*, 371 NATURE 735, 735 (1994) ("The initial outcry over DNA typing standards concerned laboratory problems: poorly defined rules for declaring a match; experiments without controls; contaminated probes and samples; and

to a kind of extended peer review as scientists openly evaluated alternative methods of interpreting such evidence. Confrontations between prosecution and defense DNA experts thus prevented the overly rapid acceptance of a technique whose use in criminal trials could influence jury decisions on whether to deny a defendant his life or liberty. Litigation in this instance prompted the scientific and legal communities to cooperate in developing standard approaches to DNA typing that would prove more resistant to criticism.³²

In other instances, however, scientific disagreements arising under the deconstructive pressure of litigation may point less persuasively to correctable flaws in theory or practice. Conflicts between experts may merely highlight areas of ambiguity where multiple approaches and interpretations could peacefully coexist, and may even create a misleading impression of uncertainty where most scientists believe none exists.³³ Locating disputes among experts on a continuum ranging from clearly meritorious to probably specious presents formidable difficulties for non-specialist courts; yet, this is precisely the kind of assessment that judges may be called upon to make in research subpoena cases.

V

AMBIGUOUS DISCREPANCIES

In litigation initiated by victims of toxic shock syndrome, Procter and Gamble sought the names and addresses of women who had been interviewed by the Centers for Disease Control ("CDC") so that the company could reinterview the women and thereby validate the CDC's results.³⁴ Such requests rest on the suspicion that raw data and interview records will disclose evidence of scientific fraud or incompetence.³⁵ On many occasions, however, the line between acceptable and unacceptable scientific practice may be much less clear than a research subpoena presupposes. Conflict about the amount of misconduct that exists in the U.S. scientific community offers *prima facie* support for this point.

Two highly publicized cases of scientific misconduct that shook American biomedical science in the 1980s cast light on the kinds of interpretive ambiguities that may arise in litigation involving subpoenaed research materials. The first case grew out of a dispute between Robert Gallo of the U.S. National Institutes of Health ("NIH") and Luc Montagnier of France's prestigious Pasteur

sloppy interpretation of autoradiograms.").

32. See NATIONAL RESEARCH COUNCIL: THE EVALUATION OF FORENSIC DNA EVIDENCE (1996); FEDERAL JUDICIAL CENTER, REFERENCE MANUAL ON SCIENTIFIC EVIDENCE (1994); NATIONAL RESEARCH COUNCIL, DNA TECHNOLOGY IN FORENSIC SCIENCE (1992).

33. Arguably, this is what happened in a number of trials involving the prescription drug Bendectin. For a review of the legal and scientific issues in these cases, see Joseph Sanders, *From Science to Evidence: The Testimony of Causation in the Bendectin Cases*, 46 STAN. L. REV. 1 (1993).

34. *Farnsworth v. Procter & Gamble*, 101 F.R.D. 355 (N.D. Ga. 1984); Wiggins & McKenna, *supra* note 7, at 81.

35. See, e.g., WILLIAM BROAD & NICHOLAS WADE, BETRAYERS OF THE TRUTH (1985).

Institute over which of the two scientists had first isolated the human immunodeficiency virus (“HIV”) suspected of causing AIDS. The second case focused on the allegation that Thereza Imanishi-Kari, a scientist at the Massachusetts Institute of Technology (“MIT”), had falsified some of the data reported in an article she had co-authored with, among others, the Nobel laureate David Baltimore. Both cases involved an exhaustive inquiry into the relationship between published scientific results and the records (including notebooks, memoranda, slides, and other forms of data) on which they were supposedly based. In each case, investigators from outside the relevant research community played a major role in determining whether misconduct had occurred. More importantly, it proved easier in both cases to deconstruct and undermine claims made by particular individuals than to produce a coherent, compelling account of the research practices that would have been appropriate.

A. The Gallo Case

On March 31, 1987, President Ronald Reagan and French Prime Minister Jacques Chirac issued a joint statement acknowledging that U.S. and French scientists had jointly discovered the cause of AIDS and that both would share the patent rights to the invention of the AIDS blood test.³⁶ This official statement was backed by a twenty-three page “history” composed by the disputing parties. Despite the formal end to the controversy, John Crewdson, a Pulitzer prize winning reporter for the *Chicago Tribune*, conducted a twenty-month investigation into suggestions that Gallo and his American colleagues had wrongfully appropriated the AIDS virus from Montagnier and the Pasteur group.³⁷ In a lengthy report published on November 19, 1989, Crewdson set out in painstaking detail the raft of disparities that he had uncovered between Gallo’s accounts of the chronology of the discovery and the records actually made in Gallo’s laboratory during the period in question. Crewdson’s tenacious inquiry left little doubt, at least in his mind, that the American group’s role in this momentous achievement raised serious questions about research ethics, honesty, and professionalism. Gallo’s publication of the “genetic twin” of a virus isolated a year earlier in France, Crewdson concluded, must have been “either an accident or a theft.”³⁸

Spurred by Crewdson’s article and heightened public concern with scientific fraud and misconduct,³⁹ the NIH launched a tortured and ultimately inconclusive investigation into the behavior of Gallo and his principal colleague and

36. John Crewdson, *The Great AIDS Quest*, CHI. TRIB., Nov. 19, 1989, §§ 5, 15.

37. *Id.* at 11. The degree of genetic identity between the LAV virus isolated by the Pasteur group and the HTLV-3 virus isolated and used to develop the AIDS blood test by the NIH group left little doubt that they were not merely closely related but virtually the same. The rapid mutation rate of the AIDS virus as it is transferred from patient to patient essentially ruled out the possibility of an accidental similarity, possibly occasioned by close contact between the patients carrying the two viruses. The remaining possibilities were accidental or purposeful contamination occurring in Gallo’s lab.

38. *Id.*

39. See generally BROAD & WADE, *supra* note 35 (reviewing many instances of scientific misconduct that aroused public attention in the early 1980s).

coauthor, Czech-born virologist Mikulas Popovic.⁴⁰ Questions centered on whether Gallo and Popovic had actively fabricated, falsified, or misrepresented data or whether they had otherwise deviated from accepted research practices. Inevitably, as in research subpoena cases, the misconduct inquiry necessitated a look behind the claimed scientific facts and a skeptical second-guessing of the means by which they had been produced.

One issue of research practice spotlighted in Crewdson's report and in the subsequent NIH investigation exemplified, in concentrated form, the problems of tacit knowledge, interpretive flexibility, and experimenter's regress that have been noted by sociologists of scientific knowledge. It was discovered that Popovic had pooled viruses taken from up to ten different AIDS patients in a single culture in order to induce the virus to replicate.⁴¹ Was this good science or a recipe for deception? Popovic's experiment proved successful, as the AIDS virus, HTLV-3, was ultimately isolated from his pooled culture. At the same time, however, pooling made it more difficult to trace the exact genealogy of HTLV-3, and thus to resolve the legally and ethically significant mystery of how HTLV-3 came to be nearly identical to the French virus LAV, which was also present in cultured form in Popovic's laboratory.

Crewdson's interviews revealed substantial differences of opinion among scientists concerning the validity of Popovic's method.

Asked why Popovic had pooled patient materials in his quest for the virus, Gallo replied: "To any virologist, it's obvious why. Anybody knows why we would mix. We weren't doing well with singles and you're trying to complement. Mika is a classic virologist. He comes from the old school in Prague, which used to be No. 1 in the world." The reaction among Popovic's colleagues, however, has been mostly one of puzzlement mixed with dismay that the genealogy of the AIDS virus that has become the diagnostic "gold standard" is lost to history.

"As a virologist, you don't throw lots of viruses into one culture," says Jay Levy. "You do it independently." Lee Ratner, the scientist who coordinated the genetic sequencing of the AIDS virus in Gallo's lab agrees: "It was not the cleanest scientific experiment. It confused things endlessly. I don't know why you would ever mix things. All I know is Mika did it."⁴²

Was Popovic an inspired experimentalist trained, as Gallo asserted, by "classic" methods at the "No. 1" school in the world? Or was he a scientific renegade whose unorthodox practices "stole" a seminal discovery from somebody else? The point to note for our purposes is that all the twists and turns of the extensive NIH misconduct investigation failed to yield definitive answers to these questions. Popovic was deemed in turn a hero, a suspect, and a victim of excess investigative zeal. Courts, lacking the flexibility to revisit the same litigation twice, may render judgments that appear more conclusive, but their apparent finality should not obscure the fact that the underlying subject matter is equally open to multiple readings.

40. See, e.g., Joseph Palca, "Verdicts" Are in on the Gallo Probe, 256 SCI. 735 (1992).

41. Crewdson, *supra* note 36, at 12.

42. *Id.*

B. The *Imanishi-Kari Baltimore Case*

Thereza Imanishi-Kari's work at MIT gained notoriety even though it was conducted for lower stakes and under less competitive circumstances than the race to find the AIDS virus. Like Gallo, Imanishi-Kari also was ultimately exonerated,⁴³ but the protracted inquiry into her scientific integrity proved just as inconclusive a device for uncovering "the truth." In this case, suspicion was aroused by Margot O'Toole, a junior researcher in Imanishi-Kari's lab, who had repeatedly failed to replicate an experiment using a reagent characterized by her superior in a co-authored paper published in the journal *Cell*. Imanishi-Kari blamed O'Toole's incompetence for the failed experiments, whereas O'Toole became convinced that some of the data underlying the *Cell* paper had been deliberately falsified. O'Toole's suspicion was intensified when she accidentally came upon seventeen pages of missing lab notes that seemed to show different results from those reported by Imanishi-Kari.

What began as a perfectly commonplace dispute between two scientists snowballed into a public drama of national proportions. Scientific fraud was already on the political agenda, as demonstrated by the Gallo case. O'Toole's charges, moreover, implicated not only Imanishi-Kari and MIT, but also David Baltimore, one of the country's most eminent and powerful scientists and a co-author of the controversial *Cell* paper. Inquiries launched by a congressional subcommittee chaired by Representative John Dingell⁴⁴ and by NIH's Office of Scientific Integrity ("OSI")⁴⁵ kept the case in the limelight as a symbol of what many saw as science's disgraceful lack of accountability to the public. More important for our purposes, the investigations generated two crucially different interpretations of Imanishi-Kari's behavior—that she was guilty of fraud or that she had committed pardonable error—each supported by evidence from her scientific records. The parties, interests, and observations that constituted the "fraud story" and the "error story" have been described by the historian of science, Daniel Kevles.⁴⁶

1. *The "Fraud Story."* Key to this interpretation were extracts from Imanishi-Kari's notebooks, including the seventeen pages discovered by O'Toole, as well as subsequent forensic analyses by the Secret Service claiming that her data had been recorded out of sequence and that original data could not have been produced at the times she claimed.⁴⁷ The chief proponents of this interpretation were, to varying degrees, outsiders to the mainstream culture of bio-

43. On June 21, 1996, the Research Integrity Adjudications Panel of the Department of Health and Human Services, an appeals board, rejected ORI's findings that Imanishi-Kari had fabricated data. *See Noted Finding of Science Fraud Is Overturned by a Federal Panel*, N.Y. TIMES, June 22, 1996, at A1.

44. Dingell at the time chaired the Subcommittee on Oversight and Investigations of the House Committee on Energy and Commerce. In this capacity, he crusaded vigorously against perceived misuses of public funds. Daniel Kevles, *The Assault on David Baltimore*, NEW YORKER, May 27, 1996, 94, 101.

45. This office was later reconstituted as the Office of Research Integrity.

46. Kevles, *supra* note 44, at 94-109. As is evident from the title of Kevles's piece, his sympathies lie with Baltimore and Imanishi-Kari; he is thus himself a partisan of the "error story," which portrays both scientists in a more favorable light.

47. *Id.* at 107.

medical science. O'Toole, the reluctant whistle-blower, was probably the closest to being an insider as she was an active post-doctoral student until her suspicions set her apart from her professional community. Interestingly, in Kevles's account, O'Toole at first suspected error but later changed her interpretation to fraud for personal reasons.⁴⁸ Further out of the mainstream were Walter Stewart, a scientist at NIH, who was reshaping his career as an outspoken critic of scientific misconduct,⁴⁹ Dingell, the legislator and vigilant guardian of public expenditure, and Suzanne Hadley, the committed OSI official whom Kevles accuses of misguided sympathy with O'Toole and a faulty understanding of how science is done.⁵⁰

2. *The "Error Story."* This interpretation accepted, but characterized as sloppiness or poor scientific judgment, some of the evidence offered by proponents of the "fraud story," such as discrepancies between Imanishi-Kari's notebooks and the *Cell* paper. It rejected outright, however, other evidence such as some of the Secret Service's forensic findings.⁵¹ Proponents of this interpretation were mostly members in good standing of the relevant research community. Imanishi-Kari herself was one. She excused her research practices as "messy" but reliable, since she knew how to make sense of her own data.⁵² Baltimore, whose support of Imanishi-Kari was construed by some as an immoderately arrogant defense of science, was another. Also included were several investigative panels consisting of scientists appointed by Tufts, MIT, and NIH to look into the substance of O'Toole's inquiries, none of whom found a case of serious professional misconduct.

C. Conclusion

The characteristics of experimental practice and of experimentally derived knowledge were as much at stake in the Imanishi-Kari case as in the controversy over Gallo and Popovic. In each case, popular opinion split between two idealized—and equally unrealistic—versions of the scientific method: science as cleanly rule-bound and governed by well-defined, communal standards of correct and deviant procedure, and science as the preserve of inspiration, intuition, and risk-taking, where messy means could be condoned and even applauded provided they supplied successful ends. For members of the "science is clean" school, both Popovic and Imanishi-Kari strayed from the paths of scientific righteousness, perhaps to the point of moral transgression; for adherents of the

48. *Id.* at 106.

49. Stewart, along with another NIH colleague, Ned Feder, was the co-author of a controversial and widely discussed article on scientific misconduct in the journal *Nature*. See Walter W. Stewart & Ned Feder, *The Integrity of the Scientific Literature*, 325 *NATURE* 207 (1987).

50. Kevles, *supra* note 44, at 108. "Some of the charges struck at [Imanishi-Kari's] tendency not to practice science in the mechanical way that Suzanne Hadley, for one, seems to think that it ought to be practiced . . ." *Id.*

51. In a nice display of experimenter's regress, the Secret Service's own techniques were declared wanting by a forensic expert working for Imanishi-Kari. This expert, as Kevles notes, "analyzed the Secret Service findings and original material and reported that it was in fact impossible through the techniques employed by the agency to draw any conclusions about when the [data] had been generated." *Id.*

52. *Id.* at 104.

“science is messy” school, Popovic was a hero whose unique approach made a recalcitrant virus grow, and Imanishi-Kari was not a heroine only because she could report no comparable success.

In fact, these cases tell the observant analyst a more complex story than any of the participants told. “In human affairs, as in science,” asserts Kevles, “truth is inseparable from the standards and processes used in determining it.”⁵³ Truth, in other words, is a function of the socially constructed methods by which human beings go about seeking the truth. These methods, as the Gallo and Imanishi-Kari cases bear out, are seldom unanimously accepted. The acceptability of scientific methods depends considerably on the context of practice. Messy and unorthodox methods that may be appropriate for cutting-edge academic research, where replication is likely to weed out mistakes, are out of place in a testing laboratory or clinical trial, where accidental contamination could have devastating consequences and sound record-keeping and comparability across cases are of paramount importance.⁵⁴ The potential for variation in standards of research methods led a panel of the National Academy of Sciences, in 1992, to define a discrete category of “questionable research practices” and to exclude these from the official definition of scientific misconduct.⁵⁵

Judgments about experimental validity, thus, cannot be made in a vacuum, independent of any wider attempt at contextualization. Cut loose from its domain of practice, an experimental claim ceases to be anchored to a communal system of belief and, hence, is susceptible to deconstructive attack by anyone who wishes to advance a contrary claim. Scientific claims acquire enough force to count as “knowledge” only through their incorporation into established, ongoing, collective routines of production and interpretation. Validity judgments within scientific subcultures, moreover, cannot be reduced to explicit rules that others can mechanically apply because, as we have noted, much of the knowledge that defines such cultures is tacit and defies complete codification.⁵⁶ We will return later to the implications of these observations for research subpoena cases.

53. *Id.*

54. See, e.g., Jon Cohen, *Clinical Trial Monitoring: Hit or Miss*, 264 SCI. 1534 (1994).

55. NATIONAL ACADEMY OF SCIENCES, RESPONSIBLE SCIENCE: ENSURING THE INTEGRITY OF THE RESEARCH PROCESS (1992). The report of the Panel on Scientific Responsibility and the Conduct of Research defined “questionable research practices” as “actions that violate traditional values of the research enterprise and that may be detrimental to the research process.” *Id.* at 5.

56. Collins, *supra* note 21. Tacit knowledge and culturally conditioned judgment present very general problems for the law, as seen in contexts where experts are required, for legal reasons, to make their scientific thought processes completely transparent. Thus, regulatory agencies have faltered in their attempts to reduce the risk assessment of carcinogenic chemicals to explicit and exhaustive principles. Every such attempt has revealed new areas of judgment not yet encompassed by principles. Controversies about the regulation of carcinogens have, therefore, begotten an almost interminable deconstruction of regulators’ technical judgments. See Sheila Jasanoff, *Science, Politics, and the Renegotiation of Expertise at EPA*, 7 OSIRIS 195 (1992).

VI

QUALITATIVE RESEARCH AND FIELD NOTES

The disclosure of information from qualitative social science research entails somewhat different consequences. Confidentiality of research subjects is perhaps the most generally recognized problem, and it is the one that courts have grappled with the most extensively.⁵⁷ The subjective (indeed, intersubjective) and non-standardized character of interviews and field notes creates more subtle problems.

Practices for recording observations in qualitative research are less standardized than in experimental science, where, as we have seen, there is already the potential for substantial variation among researchers and laboratories. Different theoretical perspectives on qualitative research may give rise to divergent objectives, sampling strategies, and ethics.⁵⁸ Researchers may or may not have received formal instruction in how to take, compose, or interpret field notes. The simultaneous creation of written notes and technological records, such as video- and audiotapes, may increase the discrepancies, depending, for example, on whether researchers see their notes as supplements to or back-ups for taped records. Rules or criteria for doing qualitative research may need continual modification in light of the observed complexity of social phenomena.⁵⁹ Thus, contrary to judicial supposition, practices in the field can vary widely.⁶⁰

Comparing qualitative research findings with the underlying data can lead to considerable confusion unless such work is handled sensitively. One experienced ethnographer, who is also a noted sociologist, summarizes the possible

57. Courts may require the removal of identifying information (such as names, addresses, and birth dates) from requested data. See, e.g., *In re Application of Am. Tobacco Co.*, 880 F.2d 1520, 1525 (1989) (describing precautions taken by tobacco company to preserve the confidentiality of epidemiological study subjects). Courts may also craft special procedural safeguards to protect the confidentiality of research subjects. For example, in 1992 a University of Alabama sociologist, J. Steven Picou, was subpoenaed by the Exxon Shipping Co. to make available his work on community stress in Alaskan coastal villages following the 1989 Exxon Valdez oil tanker disaster. The Federal District Court for the Southern District of Alabama granted Exxon's request but stipulated that the information sought could be reviewed only by the expert sociologist whom Exxon had retained to assess Picou's files. In turn, the expert was warned that he could be cited for contempt if he failed to protect confidentiality. Marshall, *supra* note 11, at 284. For a relatively unsympathetic treatment of confidentiality claims, see *In re Grand Jury Subpoena* dated January 4, 1984, 750 F.2d 223, 225 (2d Cir. 1984) (requiring justification for assurances of confidentiality made to sources in a sociological study).

58. See, e.g., LEONARD SCHATZMAN & ANSELM L. STRAUSS, *FIELD RESEARCH* 8 (1973) (describing disagreement between an anthropologist and other social scientists about whether certain Native American records constituted a "universe" within which random or stratified sampling could appropriately be used).

59. See ANSELM STRAUSS & JULIET CORBIN, *BASICS OF QUALITATIVE RESEARCH* 249-58 (1990) (calling attention to flexible nature of criteria and to differences between concepts such as "reproducibility" in natural and social sciences).

60. In *In re Grand Jury Subpoena dated January 4, 1984*, the Second Circuit Court of Appeals asked Mario Brajuha, a sociology graduate student, to establish "the nature and seriousness of the study and the methodology employed" as a precondition for seeking a scholar's privilege against disclosure. 750 F.2d at 225. Such a request appears to presuppose greater methodological uniformity in sociological fieldwork than may, in fact, exist.

difficulties in this way:

How would I react if my field notes were submitted to a panel of outsiders? I would be uneasy about what such readers might make of the unsystematic, haphazard qualities of what they would find in my notes. There might be an occasional question of confidentiality, but for the most part I would worry about assessments about the lack of order, consistency, and the like, especially if I figured that these would be held up against an idealized conception of method. A more general problem has to do with the fact that much of what I might say in a publication cannot easily be traced to specific notes taken at some occasion. My field notes tend to describe what I happen to have noted at a given time and place, and not more general, cumulative insights about the routine order of things.⁶¹

If these habits are widely shared, then a social science researcher could be extremely vulnerable to charges of drawing conclusions that are insufficiently substantiated by field records.

Drawing inferences from field notes and other observations tends, in any case, to involve complex acts of interpretation rather than simple applications of well-understood analytic principles. Not only the recording, but also the interpretation of qualitative data involves different ways of seeing and understanding social phenomena. Even the most accurate records (for example, videotapes) are subject to multiple interpretations. Thus, in the infamous Rodney King case, a single videotape, produced by a chance witness, of four white police officers beating a seemingly helpless, unresisting black man was interpreted in the courtroom to tell stories with very different moral implications and legal consequences.

Anthropologist Charles Goodwin observed in a perceptive analysis of the first Rodney King trial that legal argumentation can transform raw sensory impressions into "professional vision," that is, into particular "socially organized ways of seeing and understanding events that are answerable to the distinctive interests of a particular social group."⁶² Goodwin identified three interpretive practices, coding, highlighting, and producing material representations, that the lawyers and expert witnesses for the police officers used to impose their "professional vision" on the ambiguous videotape.⁶³ In the victim's rendition, the tape displayed a single, seamless episode of racially motivated police brutality. In the police officers' version, as recounted by a specialist in police practice, the same incident was broken into a disjointed sequence of separate mini-events that required expert decoding.⁶⁴ Frame by frame, the jerky movements of the victim's body parts were interpreted as providing a plausible rationale for "assessment periods," "escalations of force," and the delivery of strategically directed "kicks" and "blows."⁶⁵ Goodwin concluded that the fact finder's willingness to accept one reading of the evidence over another in such situations may have more to do with the professional power of the witnesses than

61. Personal interview with Michael E. Lynch, Brunel University, London (June 20, 1996) (quotation is from a letter of Professor Lynch's).

62. Charles Goodwin, *Professional Vision*, 96 AM. ANTHROPOLOGIST 606, 606 (1994).

63. *Id.*

64. *Id.* at 619-20.

65. *Id.* at 619-22.

with abstract and absolute criteria of truth.⁶⁶

The recognition of subjectivity in seeing and understanding cultures has deeply influenced the way in which social scientists, particularly ethnographers, view and practice their "science." Many modern ethnographers are intensely aware of the historically and culturally situated character of their work, as well as the frequent disparities in economic, political, and rhetorical resources between the observer and the observed.⁶⁷ Sensitized in this way, they no longer regard the observation of other cultures as objective fact finding from a value-neutral standpoint. The recording and writing of ethnography has changed to reflect these new perceptions, as described by the historian and anthropologist James Clifford:

Anthropological fieldwork has been represented as both a scientific "laboratory" and a personal "rite of passage." The two metaphors capture nicely the discipline's impossible attempt to fuse objective and subjective practices. Until recently, this impossibility was masked by marginalizing the intersubjective foundations of fieldwork, by excluding them from serious ethnographic texts, relegating them to prefaces, memoirs, anecdotes, confessions, and so forth. Lately this set of disciplinary rules is giving way. The new tendency to name and quote informants more fully and to introduce personal elements into the text is altering ethnography's discursive strategy and mode of authority.⁶⁸

Shifts such as this carry numerous implications for legal fact finding.

After admitting the intersubjectivity of ethnographic fieldwork—that "truth" in this context arises from the meshing or interpenetration of multiple subjective viewpoints—researchers may take a variety of steps to ensure the authoritativeness of their interpretations. At one extreme is the practice, noted by Clifford, of letting informants speak in their own voices. Unfortunately this opens up the further question of how typical or generalizable these voices may be in relation to the culture they allegedly represent. At the other extreme is the effort to construct an "objective" truth by accumulating many individual stories and systematizing them in accordance with a predetermined analytic framework. These stories, in turn, may be validated in different ways. For example, one researcher speaks of applying the test of "recognizability" by going back to informants with a preliminary analysis and asking if it corresponds in a recognizable way with the reality of their lives and experiences.⁶⁹

The turn toward such dialogic methods of creating accounts of social reality brings its own problems for legal practice and procedure. First, there is the mundane and recurrent issue of methodological conflict in a changing field, requiring courts to sort out claims and counterclaims about "proper" methods that have not yet been resolved by the discipline. When disciplinary rules are in transition, as they seem to be in ethnography, the possibility of distinguishing

66. *Id.* at 624-26.

67. See generally WRITING CULTURE: THE POETICS AND POLITICS OF ETHNOGRAPHY (James Clifford & George E. Marcus eds., 1986); JOHN VAN MAANEN, TALES OF THE FIELD: ON WRITING ETHNOGRAPHY (1988).

68. James Clifford, *On Ethnographic Allegory*, in WRITING CULTURE, *supra* note 67, at 109.

69. Personal interview with Dr. Keith Hawkins, University of Oxford (June 19, 1996).

truth from falsehood, reasonable from unreasonable interpretations, and fringe from mainstream practices becomes more difficult than when a discipline is in a steadier state. Methodological attacks, reflecting fundamental and unresolved theoretical disagreements, are more frequent. Under these circumstances, the law's institutional commitment to notions of unambiguous facticity and truth run up against a reality in which knowledge is transparently value-laden, political, and contested.

Second, there is the deeper issue of the cognitive status of notes and other "one-sided" accounts of more complex realities. Can an expert retained by a party or appointed by the court reliably assess the accuracy or sufficiency of a social researcher's data without steeping herself in the situations from which those data were drawn? Can "truth" be determined from field notes and the like without the participation of those whose words and actions are therein recorded? More generally, how can courts accomplish their central mission of resolving disputes and creating order in cases involving social science information if the scientific knowledge they rely on is shown, in many cases, to be both contingent and open-ended? We will return to these questions below.

VII

THE TESTIMONY OF PEERS

Because peer review can become a major issue in research subpoena cases, judges should have a thorough grasp of the role that it plays in the production of scientific knowledge.⁷⁰ Questions concerning peer review can arise in two ways. First, a litigant may ask for information about peer review in an effort to undermine the credibility of published scientific results. For example, an attorney representing plaintiffs in breast implant litigation served a subpoena on Dr. Marcia Angell, an editor of the *New England Journal of Medicine*, for all records pertaining to the peer review of an epidemiological study that reported no correlation between silicone and immune system disorders.⁷¹ Second, subpoenaed information may be turned over for review to a party expert, whose conclusions in turn are presented to the courts as evidence. For example, in a case arising from the 1989 Exxon Valdez oil spill, the court stipulated that raw and unpublished data collected by a sociologist, J. Steven Picou, should be given for evaluation to Exxon's designated technical expert, Richard Berk, a sociologist at the University of California, Los Angeles. The court warned that Berk might be cited for contempt if he failed to protect the confidentiality of Picou's data.⁷²

70. For descriptions and analyses of peer review in science, see generally John C. Burnham, *The Evolution of Peer Review*, 263 JAMA 1323 (1990); Stephen Cole, et al., *Chance and Consensus in Peer Review*, 214 SCI. 881 (1981); Marjorie Sun, *Peer Review Comes Under Review*, 244 SCI. 910 (1989). See also, BROAD & WADE, *supra* note 35, at 88-106; JASANOFF, *supra* note 10, at 64-76.

71. ANGELL, *supra* note 12.

72. Marshall, *supra* note 11, at 284.

Daubert v. Merrell Dow Pharmaceuticals, Inc.,⁷³ arguably the most influential decision concerning the admissibility of scientific evidence in the last half century, encouraged judges to ask what makes science “scientific” and to apply those criteria in discriminating good scientific evidence from bad.⁷⁴ Judicial attention was particularly drawn to peer review, which was one of the four criteria explicitly endorsed by the *Daubert* majority.⁷⁵ Peer review not only institutionalizes science’s well-known capacity for self-criticism but provides on its face a convenient procedural solution to a knotty substantive problem. Courts may be helpless when it comes to assessing scientific evidence on its merits, but even a technically illiterate fact finder can ask whether a claim offered in evidence has been peer reviewed and then devalue claims that have never been exposed to this form of scrutiny.

Deference to peer review, however, appears to be giving way to a wholesale skepticism that is equally problematic. Peer review, as courts have rightly acknowledged, is a cover term for a wide range of practices that can vary from one scientific institution (for example, journals and granting agencies) to another. More than a decade of systematic research has made it plain that peer review does not provide fail-safe guarantees of scientific integrity.⁷⁶ At best, peer review constitutes a partial screen, whose efficacy depends to some extent on the reviewer’s skill, training, and professional biases.⁷⁷ Even the Supreme Court in *Daubert* registered skepticism about the filtering capacity of peer review, noting that “publication (which is but one element of review) is not a *sine qua non* of admissibility; it does not necessarily correlate with credibility.”⁷⁸

Skepticism about peer review, however, can be carried to extremes, as a federal district court did in *Valentine v. Pioneer Chlor Alkali Company, Inc.*⁷⁹ Following *Daubert*, the district court held that the sole fact of publication in a

73. 509 U.S. 579 (1993).

74. A point insufficiently noted is that judges are permitted in the process to impose their own idealized, and for the most part unchallenged, understandings of science on legal decisionmaking. This raises obvious problems of accountability as further discussed below. On the myth of mainstream science, see JASANOFF, *supra* note 10, at 206-10. For general commentary on *Daubert*, see Bert Black, et al., *Science and the Law in the Wake of Daubert: A New Search for Scientific Knowledge*, 72 TEX. L. REV. 715 (1994); Margaret G. Farrell, *Daubert v. Merrell Dow Pharmaceuticals, Inc.: Epistemology and Legal Process*, 15 CARDOZO L. REV. 2183 (1994); Kenneth R. Foster, et al., *Science and the Toxic Tort*, 261 SCI. 1509 (1993).

75. The so-called *Daubert* criteria for addressing questions of admissibility are the following: (1) has the theory or technique underlying the proposed evidence been tested and can it be falsified; (2) has it been peer reviewed; (3) if known, what is the technique’s error rate; and (4) is it generally accepted? *Daubert*, 509 U.S. at 593-94. These criteria are, of course, merely suggestions that a trial judge can consider. The Court did “not presume to set out a definite check-list” *Id.* at 593. In the end, *Daubert* stands for the use of discretion on the part of trial judges concerning the admission of scientific evidence.

76. See DARYL E. CHUBIN & EDWARD J. HACKETT, *PEERLESS SCIENCE: PEER REVIEW AND U.S. SCIENCE POLICY* (1990); STEPHEN LOCK, *A DIFFICULT BALANCE: EDITORIAL PEER REVIEW IN MEDICINE* (1985); see also BROAD & WADE, *supra* note 35; JASANOFF, *supra* note 15.

77. Peer review has proven least effective as a screen against misconduct in the biomedical sciences. One reason is that reviewers typically do not have access to the researcher’s raw data and cannot easily judge whether results have been falsified or fabricated.

78. 509 U.S. at 593 (citing JASANOFF, *supra* note 15).

79. 921 F. Supp. 666 (D. Nev. 1996).

peer-reviewed journal did not provide sufficient guarantee of scientific reliability. Conducting its own review of the expert's scientific methods, the court ruled inadmissible a published study which concluded that the plaintiffs' claimed cognitive and emotional deficits had resulted from chlorine inhalation. In reaching this result, the court distinguished between "editorial peer review," consisting of pre-publication review of an article by selected journal referees, and "true peer review," by which the judge apparently meant all post-publication attempts to replicate or otherwise build on published scientific results. "Editorial peer review," the court opined, could be accorded relatively low respect in forensic settings because of the review's narrow scope and because "the average referee spends less than two hours assessing an article submitted to a biomedical journal."⁸⁰

Such dismissive treatment misses the important constructive purpose of prepublication peer review. Within science, the policing function of peer review is deemed on the whole subordinate to its role in holding scientists to higher, communally certified standards of reasoning and presentation. That reviewers may not discover clever or determined instances of data falsification is today a foregone conclusion; scientists, as the *Valentine* court correctly observed, rely on the more deliberate processes of replication and cross-checking to root out and discard such abuses. Instead, reviewers for scientific journals and research grants are expected to apply the standards of their field fairly and dispassionately to every submission, treating it as a potential contribution to knowledge. Peer review thus refines the analysis of scientific data, points out inadequacies in argument, sharpens the focus of conclusions, and improves the clarity of expression. Science that survives peer review enters by definition into the collaborative enterprise of knowledge-making. Its findings may still be set aside in time, as unimportant or misdirected, or be modified in important particulars.⁸¹ But peer-reviewed science is earmarked at the very least as having met a research community's threshold criteria of legitimacy, relevance, and interest. For a court to rule such work inadmissible on the basis of its own methodological inquiry, as was done in *Valentine*, raises serious questions of judicial accountability.⁸²

As the foregoing account suggests, scientific peer review is likely to differ

80. *Id.* at 675 (citing Stephen Lock & Jane Smith, *What do Peer Reviewers Do?*, 263 JAMA 1341, 1341-43 (1990)).

81. For an illuminating account of how the initial publication of the structure of DNA was mistaken in key respects, see FRANCIS CRICK, *WHAT MAD PURSUIT* (1988). See also FLECK, *supra* note 18.

82. Several points about *Valentine* are worth noting. From a legal standpoint, the court appears to have determined admissibility in accordance with criteria that would have been more appropriate for judging the weight of the evidence. See *Valentine*, 921 F. Supp. at 670-71. Further, in distinguishing between "editorial" and "true" peer review, the judge made a number of unverified (possibly unverifiable) assumptions about science. For example, he assumed that all published science gives rise to post-publication scrutiny and that the quality of pre-publication review is a function of time spent on it. *Id.* at 675. He also assumed that "obscure" journals are less reliable than better known ones. *Id.* at 670 n.3. Needless to say, neither the *Valentine* court's implicit model of science nor its specific methodological critique of the rejected evidence was subjected to the equivalent of "editorial peer review." Whether it will be subjected to "true peer review," and what that might mean in practice, remains to be seen.

markedly in its objectives and impact from review carried out by an expert in a litigation context. In legal review, the goal is neither to make good work better nor to retrieve what might be of value from work of lesser significance. It is, instead, to seek as aggressively as possible to discredit the proffered evidence and to deploy in the process all the skeptical resources that experts engaged specifically for this purpose can muster. Litigation-centered review thus differs in three crucial respects from peer review as usually practiced in science:

(1) It serves the litigants' pecuniary interests rather than the disciplinary interests of a well-defined scientific community. The reviewer may belong in fact to an entirely distinct scientific subculture from the person whose work is being reviewed.⁸³

(2) It operates on the principle of exclusion rather than inclusion by aiming to deny scientific status to questioned results. Its purpose is to expose gaps and omissions in the translation from observation to result. Hence it is inherently unforgiving. In Imanishi-Kari's case, such attitudes led some reviewers to see fraud where others had seen only error.

(3) It holds research to idealized norms of good practice that do not necessarily conform to the actual practices and tacit knowledge of workers in any field. In the Gallo-Popovic investigation, such idealized accounts of science underwrote radically different interpretations of the accused scientists' research ethics.

For these reasons, the review of subpoenaed research materials by partisan experts should not be seen as a simple extension of ordinary scientific peer review, like adding another pair of discerning eyes to an existing, but inadequate, critical apparatus, or even as "true peer review" in the terms of the *Valentine* court. The same alleged "flaws" could well come to light in pre-publication and in litigation-centered review processes, but for the reasons stated above, the normative readings placed on them would probably diverge materially. In attempting to weigh different readings of the same data, judges should try to understand the basis for such divergences and to manage them appropriately.

VIII

JUDICIAL SKEPTICISM

To review scientists' research materials in the course of litigation, therefore, is not to continue the work of science by other means, but rather to engage science in an alien field of inquiry, governed in part by different rules of the game. In the litigation context, scientists are deprived of many supports that protect their work against unbounded skepticism and forgive minor errors or devia-

83. Recent tensions over attempts to consolidate peer review panels for neuroscience at the NIH underscore the importance of such disciplinary differences. While the distinction between "biopsychology" and "psychobiology" may not be apparent to outsiders, they are real and consequential to those inside these fields. See Elliot Marshall, *Despite Anxiety, NIH Begins Merging Neuroscience Panels*, 273 SCI. 731 (1996).

tions in the interest of promoting legitimate inquiry. These supports include shared, tacit knowledge about what “works,” communal standards of good practice, cultural ties that bestow credibility within research communities,⁸⁴ the testing of promising work by other investigators, and an overriding interest in the larger enterprise of creating new knowledge. In the courtroom, moreover, science is reevaluated in accordance with idealized lay conceptions of “the scientific method” that are insufficiently attuned to its genuine complexity and diversity. The fact that scientific claims unravel and that “truth” is difficult to reconstruct under legal scrutiny is well established from studies of regulatory science.⁸⁵ Research subpoenas place science and scientists in a similar predicament, without the assurance that second-guessing will produce a more robust form of knowledge, let alone the “truth.”

Skepticism toward science, however, is not intrinsically unwarranted, and researchers should not be granted full immunity merely because their work has satisfied their own professional community’s standards of criticism and peer review. Cases of scientific fraud and misconduct offer powerful evidence that autonomy and responsibility do not necessarily go hand in hand.⁸⁶ As related by the science journalists William Broad and Nicholas Wade, “[s]cience is not self-policing. Scholars do not always read the scientific literature carefully. Science is not a perfectly objective process.”⁸⁷ Qualified external critics, moreover, may uncover problems not perceived by closed communities of practice, as happened in the case of DNA typing. To make science serve justice, it may be desirable, and occasionally essential, to step outside the charmed circles of intra-disciplinary criticism.

How should courts manage the review of subpoenaed research materials so as to advance the cause of truth-seeking but to reduce the risk of imposing, or deferring to, misleading views of scientific practice? Three approaches merit consideration as they relate to the sources of misunderstanding described above: the screening of party experts; the appointment of independent experts; and the use of review panels. Each has costs and benefits that need to be carefully weighed in the context of specific cases.

A. Screening Party Experts

A relatively low-cost approach to assuring more balanced assessment of research records is for courts to screen the expert reviewers selected by the parties. By asking appropriately targeted questions, courts can exclude wholly unqualified experts, as well as help focus the issues in dispute and make sure they are located in clearly defined and relevant disciplinary and methodological

84. Shapin, *supra* note 14; see also STEVEN SHAPIN, *A SOCIAL HISTORY OF TRUTH* (1994).

85. DAVID COLLINGRIDGE & COLIN REEVE, *SCIENCE SPEAKS TO POWER: THE ROLE OF EXPERTS IN POLICY* (1986); see also JASANOFF, *supra* note 15.

86. Patricia Woolf, *Integrity and Accountability*, in *THE FRAGILE CONTRACT* 82 (David H. Guston & Kenneth Keniston eds., 1994); David H. Guston, *The Demise of the Social Contract for Science: Misconduct in Science and the Nonmodern World*, 38 *THE CENTENNIAL REVIEW* 215 (1994).

87. BROAD & WADE, *supra* note 35, at 210.

frameworks. In general, the purpose of such screening would be to secure timely disclosure of important biases rather than to exclude expert witnesses from testifying.

Criteria that courts may use in this connection include straightforward tests of credibility—whether, for example, the expert has been a “repeat witness” in litigation, maintains an active professional life in science or medicine, or is a known professional adversary of the person whose work is to be reviewed.⁸⁸ Identification of experts who testify frequently should be assisted by recent amendments to the Federal Rules of Civil Procedure that require parties to list recent testimony, as part of the standard disclosure, when proffering expert testimony.⁸⁹ State laws can provide collateral support, for instance, by limiting the percentage of income that medical or other experts may earn from appearing as expert witnesses. Courts in any event can operationalize some of the insights gained from the sociology of science by asking party-selected experts to reveal in advance the nature and sources of the standards they will use in reviewing subpoenaed materials.

B. Appointing Independent Experts

To correct for the biases of party-selected experts, who may represent extreme viewpoints, the court can appoint independent experts. The principal disadvantage, apart from potentially significant issues of cost and inefficiency, is that the independent experts’ own biases could be veiled by the presumption of neutrality that attaches to being appointed by the court. By remaining invisible, these biases could exert undue influence on the fact-finding process.⁹⁰ In line with the analysis presented in this article, independent experts would be most helpful to courts not by resolving disputes about research information but rather by locating them within an overall topography of scientific practice. The experts can educate courts about the status of methodological controversies, make discriminations between fraud and “normal” carelessness or error, and thus help separate important from trivial issues.

Courts may face a threshold difficulty in identifying suitable experts to play this impartial, educative role. A proposal developed by the American Association for the Advancement of Science to create rosters of suitable experts, with assistance from leading professional societies, could help overcome this obstacle.⁹¹ Professional groups, however, may have their own economic and social

88. Courts may be assisted in the task of screening by state statutes establishing professional standards for expert witnesses. For example, an Illinois law requires that, to qualify as an expert witness, a medical professional must have devoted 75% of his or her time to medical practice, research, or teaching. ILL. ANN. STAT. ch. 735, para. 5/8-2501(b) (Smith-Hurd 1993).

89. See FED. R. CIV. P. 26(a)(2)(B) (1995).

90. For an example of the wide-ranging influence that can be exerted by “neutral” experts, see *Computer Assocs. Int’l, Inc. v. Altai, Inc.*, 775 F. Supp. 544 (E.D.N.Y. 1991), *vacated in part on other grounds*, 982 F.2d 693 (2d Cir. 1992) (MIT computer science professor proposed standards for incorporation into copyright law).

91. AMERICAN ASSOCIATION FOR THE ADVANCEMENT OF SCIENCE, PROPOSAL—PROVIDING SCIENTIFIC AND TECHNICAL EXPERTISE TO THE FEDERAL COURTS: A DEMONSTRATION PROJECT OF

biases to which judges should always be attentive.

C. Utilizing Review Panels

Courts can also seek advice from a panel of experts instead of, or in addition to, the testimony offered by individual party experts. A procedure like this would most closely parallel scientific peer review by drawing on a range of opinion rather than a single, narrow slice. If appropriately designed, the procedure can have the merit of holding subpoenaed research to communal and widely accepted standards of performance rather than the personal and possibly idiosyncratic standards of individual experts.

As seen above in the cases of Gallo and Imanishi-Kari,⁹² however, panel review does not automatically guard against the imposition of idealized notions of scientific practice on questioned work. This review may intensify, rather than correct, inappropriate disciplinary biases. Unless the panel adequately represents relevant viewpoints, its conclusions may be biased or flawed.⁹³ Finally, cost would be an even more salient consideration with a panel than with a single independent expert.

IX

CONCLUSION

In the end, the most effective way to integrate scientific knowledge fully and fairly into legal decisionmaking may indeed be, as the *Daubert* court proposed, for judges to develop a keener sense of how science works. Judges are best positioned to fashion case-specific procedures, orders, and management practices that respect the complexities of both scientific research and the particular litigation context. In order to acquire the necessary sixth sense for science, judges will have to reach beyond lawyers', scientists', and even the Supreme Court's idealized or wishful fictions about the conduct of research. They will have to abandon any naive faith in the existence of a single, well-demarcated, unchanging canon of good scientific practice. They will have to develop a greater sensitivity to the ambiguous, provisional, and partial nature of much science that is presented as hard, unambiguous, and conclusive by liti-

AAAS (April 16, 1996).

92. See *supra* part III.A-B.

93. A high-level expert panel that notably failed to generate a scientific consensus was the first committee appointed by the National Research Council to review the methodology of DNA typing. NATIONAL RESEARCH COUNCIL (1992), *supra* note 32. The principal point of contention was the panel's treatment of population genetics and its endorsement of a controversial statistical standard known as the "ceiling principle." See Lander & Budowle, *supra* note 31, at 735-36. Controversies over the first panel's report led to the appointment of a second panel, the Committee on DNA Forensic Science, with different membership from the first. It remains to be seen whether the second panel's report will achieve more scientific and legal authority. NATIONAL RESEARCH COUNCIL, THE EVALUATION OF FORENSIC DNA EVIDENCE (1996).

gating parties and their chosen experts. They will also have to recognize, and make appropriate allowances for, the communal or collective nature of scientific judgments and their embeddedness in particular cultures of scientific practice. This article has suggested ways in which the sociology of scientific knowledge can make available additional resources for such advances in judicial understanding.