Expert Evidence and Scientific Disagreement

*Mike Redmayne**

INTRODUCTION

The legal system has recently been confronted with a series of scientific evidence disputes. Protracted disagreement between scientists over the probative value of DNA evidence, and over whether the drug Bendectin[1] caused birth defects are but two of the best known examples. Many other cases have also produced some degree of scientific disagreement. This Article examines scientific disagreement over evidence used in litigation. Part I sketches a theoretical framework for understanding scientific disputes. Drawing on work in the philosophy and sociology of science, it identifies a number of common factors among scientific disputes, and asks how such disputes come to be resolved. The role that the interface of science and the legal process plays in exacerbating scientific disagreement is also examined. Part II offers a detailed examination of three case studies of disputes over scientific evidence. Although the three examples chosen are all very different, the theoretical perspectives examined in Part I allow us to see a number of links between them. In particular, the connections between uncertainty, disagreement, and policy concerns emerge as a common pattern. Additionally, in all three examples we can identify the perturbations caused by the interface of science and the legal process. Together, these factors render the closure of disputes over scientific evidence problematic. Part III looks at the implications of the case studies for the

---

* Department of Law, Brunel University, United Kingdom. I am grateful to Neil Duxbury and Paul Roberts for comments on an earlier draft.

[1] Bendectin is a drug used for the symptomatic relief of "morning sickness" in pregnant women. MODERN DRUG ENCYCLOPEDIA AND THERAPEUTIC INDEX 96 (Robert S. Goodhart et al. eds., 10th ed. 1965). It combines three drugs which have antinausea effects, but which are more effective in combination than when used alone. *Id.*

resolution of scientific disputes that affect litigation, and for admissibility standards for scientific evidence. One conclusion is that scientific disagreement is an inevitable, and often useful, product of the use of scientific evidence in litigation.

I. UNDERSTANDING SCIENTIFIC DISAGREEMENT

A. *Disagreement in Science*

The ability to generate universal agreement was once considered a central feature of science. Norman Campbell, writing in 1921, defined science as "the study of those judgments concerning which universal agreement can be obtained."[2] Although he acknowledged that scientists would sometimes disagree, Campbell thought that in such a situation universal agreement could be restored by testing the disputed proposition.[3] If Campbell's depiction of science was accurate, scientific evidence would hardly be a topic of much interest to academic lawyers. Expert witnesses would rarely disagree in court, and if they did, we would expect the scientific community to be able to supply an unequivocal answer to the disputed question.

Other writers who have theorized about science have taken a rather different view. The failure of philosophers of science to offer a watertight inductive logic for theory choice in science has undermined attempts to found the scientific enterprise on universally agreed standards.[4] Responding to the failings of inductivism, philosophers of science have given scientific disagreement varying degrees of prominence. Popper's notion of scientific development through falsification demands some degree of disagreement between scientists, for it is only through the sustained attempt to falsify a prevailing theory that science will progress.[5]

---

[2] NORMAN CAMPBELL, WHAT IS SCIENCE? 27 (1921).

[3] *See id.* at 34 (arguing that conducting another test will restore universal agreement).

[4] *See* ANTHONY O'HEAR, AN INTRODUCTION TO THE PHILOSOPHY OF SCIENCE 12-34 (1989) (arguing that flaws in inductive logic have undermined attempts to found science on universally agreed standards).

[5] *See* KARL POPPER, CONJECTURES AND REFUTATIONS 228 (5th ed. 1989) (stating that falsificationists believe that what cannot in principle be overthrown by criticism is unworthy of serious consideration, while verificationists only accept belief if it can be justified by positive evidence).

In contrast, Kuhn, for all his supposed radicalism, is sometimes seen as playing down the extent to which scientists who share a paradigm will disagree with one another.[6] Kuhn acknowledged that scientists would disagree about theories, largely because the accepted standards of a good theory — such as accuracy, consistency, and simplicity — are indeterminate.[7] They cannot by themselves determine the decisions of individual scientists.[8] This, Kuhn supposed, left room for more subjective factors to influence the choices made by scientists. However, Kuhn did not allow that subjective criteria, such as "individual biography and personality," would play an overwhelming role in scientific decisionmaking.[9] The objective standards espoused by traditional philosophers of science would still be the mainstay of theory choice.[10]

Others have not been so reticent in giving social factors a prominent role in science. A number of scholars in the field of the sociology of science have explored the social dimensions of scientific activity. This work, often supported by impressive empirical studies, stresses the contingent nature of scientific knowledge.[11] In this tradition, studies of scientists in laboratories stress the role of interpretation and evaluation in the development of even the most basic theories.[12] In a similar vein,

---

[6] *See* Michael Mulkay, *Consensus*, *in* MICHAEL MULKAY, SOCIOLOGY OF SCIENCE: A SOCIOLOGICAL PILGRIMAGE 81 (1991) (arguing that Kuhn has downplayed disagreement in scientific community).

[7] *See* Thomas S. Kuhn, *Objectivity, Value Judgment and Theory Choice*, *in* THE ESSENTIAL TENSION 320-39 (1977) (arguing that accepted standards of good scientific theory are indeterminate).

[8] *See id.*

[9] *Id.* at 329 (arguing that individual theories converge into objective theory as subjective elements of decision process are eliminated).

[10] *See id.*

[11] *See* Mulkay, *supra* note 6, at 79-89; *see generally* BARRY BARNES, ABOUT SCIENCE (1985) (discussing how science is ordered and organized); HARRY COLLINS & TREVOR PINCH, THE GOLEM: WHAT EVERYONE SHOULD KNOW ABOUT SCIENCE (1993) (examining series of experiments to show that scientific certainties come from interpretation of ambiguous results); MICHAEL MULKAY, SCIENCE AND THE SOCIOLOGY OF KNOWLEDGE (1979) (arguing that sociological evaluation of scientific community is appropriate); ANDREW WEBSTER, SCIENCE, TECHNOLOGY, AND SOCIETY (1991) (discussing current developments in science and technology).

[12] *See* KARIN D. KNORR-CETINA, THE MANUFACTURE OF KNOWLEDGE (1981); BRUNO LATOUR & STEVE WOOLGAR, LABORATORY LIFE: THE CONSTRUCTION OF SCIENTIFIC FACTS (2d ed. 1986).

historical studies of important episodes in science emphasize the impact of the political climate on scientific knowledge.[13] Some writers even go so far as to reject the idea that the physical world plays a determining role in the development of science.[14] If scientists tend to agree, this must be explained through social factors. Scientific knowledge is portrayed as the product of negotiation within a scientific community, rather than as a natural feature of a discipline in which "man proposes and nature disposes."

Sociologists of science have tended to emphasize the scope of disagreement in science, claiming that "intellectual consensus is much less usual in science than has often been maintained."[15] Moreover, because some sociologists of science hold that there are no objective standards with which to compare knowledge claims, they see no obvious way in which scientific disagreement can be resolved.[16] Studies of science have come a long way from Campbell's optimistic vision of universal agreement.

The point in outlining some of the views advanced in the sociology of science is not to argue that this is what science is always like. Some of the more extreme views put forward seem to suffer from the syndrome of the disappointed infallibilist: the view that because we can know nothing with absolute certainty, we can know nothing with any degree of probability at all.[17]

---

[13] *See generally* Steven Shapin, *History of Science and Its Sociological Reconstructions*, 20 HISTORY OF SCIENCE 157 (1982) (discussing sociology as it relates to science).

[14] *See* LATOUR & WOOLGAR, *supra* note 12, at 236. These authors claim that "reality [is] the *consequence* of the settlement of a dispute rather than its *cause*." *Id.*; *see also* Harry Collins, *Stages in the Empirical Programme of Relativism*, 11 SOCIAL STUDIES OF SCIENCE 3, 3 (1981) (stating that "the natural world has a small or non-existent role in the construction of scientific knowledge").

[15] *See* LATOUR & WOOLGAR, *supra* note 12, at 89; *see also* BARNES, *supra* note 11, at 90-112 (discussing attempts by scientists to extend scope of science beyond its boundaries); SCIENCE IN CONTEXT 233-49 (Barry Barnes & David Edge eds., 1982) (arguing that credibility of expertise cannot be established by strictly logical arguments). It is not surprising, then, that some sociologists of science take the clash of conflicting expert evidence in the courts as a satisfying model of science. *See* COLLINS & PINCH, *supra* note 11, at 145-48 (arguing that given new understanding of science, it should be expected that any piece of evidence can be examined and doubted).

[16] *See, e.g.*, SCIENCE IN CONTEXT, *supra* note 15, at 11, 237 (discussing limit of scientific resolution). The editors argue that "the credibility of expertise cannot be satisfactorily established by strictly logical arguments" and that, as a consequence, the assessment of expert knowledge "raise[s] such intractable and viciously circular problems as to strangle speech." *Id.*

[17] For this and other criticisms of some of the more relativistic views of science, see

However, there are times when we would do well to remember some of the insights provided by this rich literature. Particularly, that scientific disagreement is not unusual and that it can be extremely difficult to resolve disputed issues in science.

B. Science, Society, and the Closure of Disputes

There are certain situations in which disagreement between scientists is particularly common. One of these is when scientists are involved in applying science and technology in society.[18] In an influential article, Alvin Weinberg suggests some reasons for this.[19] Weinberg argues that many of the most important questions scientists are called upon to answer are "trans-scientific," that is, they transcend science. This is partly owing to uncertainty, especially as scientists will often be asked to deliver answers in areas where research or technology is still developing. We expect scientists to give us answers before all the evidence is in. Additionally, scientific issues will often incorporate policy questions, such as what level of risk should be accepted by society when adopting a new technology.

Although we may be able to separate scientific and trans-scientific issues, often, Weinberg suggests, the line between the two will be blurred. But if scientists will tend to disagree when they are called upon to use their knowledge for social ends, how can enough consensus be generated for society to make use of scientific knowledge? How is closure achieved in scientific disputes?

One difficulty in achieving the closure of a scientific dispute is that there is a tension between closure and good science.[20] If

---

generally LARRY LAUDAN, SCIENCE AND RELATIVISM (1990) (discounting relativist view of science); ROGER TRIGG, RATIONALITY AND SCIENCE 149-71 (1993) (same).

[18] *See, e.g.*, Anthony Barker & B. Guy Peters, *Introduction: Science Policy and Government*, *in* THE POLITICS OF EXPERT ADVICE 1, 1-33 (Anthony Barker & B. Guy Peters eds., 1993). The authors argue that "scientific advice on public policy issues may be — and almost always should be — conflicting" because this reflects the nature of both science and society. *Id.* at 5; *see also* Brian Martin, *The Sociology of the Fluoridation Controversy: A Reexamination*, 30 SOCIOLOGICAL 59 (1989) (discussing disagreement among scholars in fluoridation controversy).

[19] *See* Alvin M. Weinberg, *Science and Trans-Science*, 10 MINERVA 209 (1972) (discussing scientists' attempts to answer trans-scientific questions).

[20] *See* RICHARD H. GASKINS, BURDENS OF PROOF IN MODERN DISCOURSE 166 (1992) (arguing that "the demand for closure compromises the entire procedural legitimacy of science").

we accept that science is fallible and that there is a degree of uncertainty in all scientific knowledge claims, we can never regard the closure of a scientific dispute as final. An issue, once closed, may always be reopened.

Some valuable insights on the processes of closure in science came from the work of the Closure Project. During the 1980s, this group of philosophers, scientists, and sociologists met to examine the closure of disputes in science and technology.[21] The Project examined a number of different controversies. It is hardly surprising that the ones in which closure proved most difficult were those that were, in Weinberg's terms, the most trans-scientific.

Many scientific disputes surround controversial issues of policy. Examples studied by the Project included the question of whether homosexuality should be classified as a disease[22] and whether cancer victims should have access to the drug laetrile.[23] In a purely scientific dispute, the debate over continental drift, closure was easier to achieve. However, it might be preceded by just as much disagreement as a trans-scientific dispute.[24] Thus, in the 1960s, when the evidence clearly supported continental drift and could be placed within a coherent theoretical framework, even the most vehement supporters of the theory of fixed continents came to accept the new theory.[25]

A useful distinction can be drawn between "sound argument closure" and "negotiated closure" or "consensus closure." Sound argument closure is the type of closure achieved in the continental drift debate. Negotiated closure is the only kind of agreement that can be achieved in a truly trans-scientific dispute.[26]

---

[21] *See* SCIENTIFIC CONTROVERSIES: CASE STUDIES IN THE RESOLUTION AND CLOSURE OF DISPUTES IN SCIENCE AND TECHNOLOGY (H. Tristram Engelhardt, Jr. & Arthur L. Caplan eds., 1987) (discussing closure of disputes in science and technology) [hereinafter SCIENTIFIC CONTROVERSIES].

[22] *See id.* at 381-436.

[23] *See id.* at 315-80.

[24] *See* Henry Frankel, *The Continental Drift Debate, in* SCIENTIFIC CONTROVERSIES, *supra* note 21, at 203, 203-48 (discussing debate over continental drift); *see also* RONALD N. GIERE, EXPLAINING SCIENCE: A COGNITIVE APPROACH 227-75 (1988) (discussing scientific community's debate over, and acceptance of, continental drift theory).

[25] *See generally* Frankel, *supra* note 24, at 230-43 (discussing closure of continental drift debate).

[26] *See* Tom L. Beauchamp, *Ethical Theory and the Problem of Closure, in* SCIENTIFIC CONTROVERSIES, *supra* note 21, at 27-48 (discussing various types of closure). This Article

Another important point discussed by contributors to the Closure Project was the role played by "non-epistemic" — or non-scientific — factors in scientific controversies. One would expect non-epistemic factors to play a major role in the more trans-scientific controversies. But it appears that non-epistemic factors also play a role during the course of other scientific controversies before sound-argument closure becomes possible.[27] Thus, had earth scientists been asked, during the 1950s, to conclude whether or not the continents were fixed, non-epistemic factors would likely have played a role in the negotiated closure that would have ensued.

Some examples of the sorts of non-epistemic factors that can be decisive, even in a purely scientific dispute, are provided by the arguments used by scientists in the 1970s during a disagreement over whether high fluxes of gravity waves were detectable on earth.[28] During the debate over gravity waves, the factors used by scientists to assess the validity of various experiments included the personality and reputation of another scientist, the style in which experimental results were presented, and a scientist's nationality.[29]

The difference between the sort of closure that is possible before there is sufficient evidence to make the choice between different theories clear, and that which is achievable afterwards — in the continental drift debate around 1965 — illustrates

---

adopts a fairly simple analysis of the types of closure which occur in scientific disputes. Other participants in the Closure Project suggested different ways of categorizing closure. *See* H. Tristram Engelhardt, Jr. & Arthur L. Caplan, *Patterns of Controversy and Closure: The Interplay of Knowledge, Values, and Political Forces*, *in* SCIENTIFIC CONTROVERSIES, *supra* note 21, at 1-23 (discussing closure to scientific controversies). For example, some scientific controversies will be subject to "natural death closure" because science moves on, leaving little interest in the controversy. *Id.* at 5; *see, e.g.*, Eliot Marshall, *Disputed Results Now Just a Footnote*, SCIENCE, July 12, 1996, at 174-75 (discussing little interest in paper that detailed effect of new immune system gene).

[27] *See* Ernan McMullin, *Scientific Controversy and Its Termination*, *in* SCIENTIFIC CONTROVERSIES, *supra* note 21, at 67 (arguing that "nonepistemic factors can play a major, sometimes even a decisive, part" before one of two rival theories accumulates a better explanatory record than the other); *see also* Alvin M. Weinberg, *Science and its Limits: The Regulator's Dilemma*, 2 ISSUES IN SCI. & TECH. 59 (1985) (arguing that regulators should rely on less definitive answers).

[28] *See* HARRY COLLINS, CHANGING ORDER: REPLICATION AND INDUCTION IN SCIENTIFIC PRACTICE 79-111 (1985) (discussing gravitational radiation disagreement).

[29] *See id.* at 87 (listing factors used to assess validity of various experiments).

another important point about scientific disagreement. Disagreement is a function of uncertainty and, during periods of uncertainty, non-epistemic factors will play a more important part in scientific disagreement. This is one reason why the application of technology in society can prove controversial: it often calls for an assessment of the technology before science has reached a stage where an accurate assessment is possible.

Another conclusion that can be drawn from a study of scientific controversies is that different controversies show remarkable structural similarities.[30] One of the most interesting of these similarities is the rhetorical use of burdens of proof by competing sides in a controversy, especially as a device for responding to uncertainty and risk.[31] Formal burdens of proof are one way in which the law responds to uncertainty. However, informal burdens of proof also play a role in scientific discourse as a means of dealing with uncertainty. As one commentator has put it, they play a "critical role . . . in the scientific approach to finality."[32]

If one analyzes the arguments used by actors in different controversies about the risks posed by technology, one can find virtually identical arguments. The participants resort to the same strategy of trying to win the debate by throwing the burden of proof onto the other side.[33] Critics of a new technology will

---

[30] *See* ALLAN MAZUR, THE DYNAMICS OF TECHNICAL CONTROVERSY 13-32 (1981) (comparing controversies surrounding fluoridation of water and nuclear power).

[31] *See id.* at 14-17 (discussing rhetorical devices used by each side); *see also* Allan Mazur, *Disputes Between Experts*, 11 MINERVA 243, 245-49 (1973) (same).

[32] *See* GASKINS, *supra* note 20, at 143 (stating that presumptions play critical role in scientific finality). Gaskins suggests that burdens of proof play a particularly important role in policy-oriented debates about science because, "[f]aced with unwelcome policy conclusions, critics can play the part of objective arbiters by documenting the inevitable failure of opposing experts to verify their hypotheses." *Id.* at 162. This skeptical attitude is a safe one to play, because one can maintain it without putting forward any theories of one's own. *See id.* (stating safest role is that of hard-nose skeptic); *see also* RONALD N. GIERE, UNDERSTANDING SCIENTIFIC REASONING 115 (3d ed. 1991) (noting the role of burdens of proof in arguments about marginal science); Carl F. Cranor, *Learning From the Law for Regulatory Science*, 14 LAW & PHIL. 115, 132 (1995) (observing that "[a]nalogs to legal presumptions and burdens of proof can be used to address some of the uncertainties in the sciences").

[33] *See* MAZUR, *supra* note 30, at 14-32 (discussing examples of arguments used to shift burden of proof). For other examples of burden of proof arguments in science, see Mark MacCarthy, *Closure in Occupational Safety and Health: The Benzene and Cotton Dust Decisions*, *in* SCIENTIFIC CONTROVERSIES, *supra* note 21, at 505, 512-18 (discussing shifting burden of

argue that when there is a possibility of risk to the public, it is for the proponents of a technology to prove its safety. In response, the technology's supporters will argue that, in the absence of evidence that it is dangerous, the technology should be presumed to be safe.

To this point, this Article has examined disagreement between scientists in fairly general terms. Much of the rest of the Article will involve looking in some detail at disagreements between scientists over scientific evidence used in litigation. Before turning to do this, this Article briefly addresses the relationship between science and the legal process. This is because an understanding of expert disagreement must be informed as much by an analysis of law and of the way it deals with science, as by an analysis of science itself.

C. Science and the Legal Process

There is nothing new in pointing out that science and law sometimes appear to be at loggerheads with each other. The relationship between science and law has been described as a marriage of opposites,[34] as a conflict between rival systems,[35] and as a clash of cultures.[36] The legal process is quite often criticized by both lawyers and scientists on the grounds that legal concepts and legal procedures are unscientific.[37]

---

proof to OSHA to prove substance harmful); JANE C. KRONICK, VALUES AS THEY INFORM THE USE OF SCIENTIFIC EVIDENCE: CONTEXTUAL ANALYSIS OF CONGRESSIONAL HEARINGS ON HAZARDOUS AND RADIOACTIVE WASTE, FINAL REPORT TO THE NATIONAL SCIENCE FOUNDATION 293, 301 (1988), *quoted in* GASKINS, *supra* note 20, at 150-51.

[34] *See* Anita Ky Wonder, *Science and Law, A Marriage of Opposites*, 29 J. FORENSIC SCI. SOC'Y 75, 75 (1989).

[35] *See* Lee Loevinger, *Law and Science as Rival Systems*, 19 U. FLA. L. REV. 530, 541 (1967) (pointing out traditional tension between dialectic method of law and empirical method of science).

[36] *See* STEVEN GOLDBERG, CULTURE CLASH: LAW AND SCIENCE IN AMERICA 4 (1994) (discussing how science overwhelms literary values, including social sciences); Leslie Roberts, *Science in Court: A Culture Clash*, 257 SCIENCE 732, 732 (1992) (arguing that lawyers and scientists approach courtroom with different attitudes and methodologies).

[37] *See, e.g.*, D.J. Gee, *The Expert Witness in Criminal Trials*, 1987 CRIM. L. REV. 307, 308-14 (criticizing the adversarial system in regard to treatment of expert testimony); Geoffrey M. Stephenson, *Looking to the Future: A Psychologist's Comment on Richard Abel's Contested Communities*, 22 J.L. & SOC'Y 133, 136-37 (1995) (criticizing legal system's lack of responsiveness to psychology).

Are these criticisms of the legal system valid? It would seem natural to argue that law should be more closely reconciled to the disciplines with which it comes in contact. Some writers, however, emphasize the autonomy of law and portray the legal system's differentiation from other social systems as one of its strengths.[38] On this account, law's lack of responsiveness to science is not something to be criticized.

Law's resistance to other disciplines enables the legal system to function effectively and helps to uphold law's normative authority. When the legal process uses knowledge from other disciplines, it tends to do so selectively, for its own purposes.[39] Thus, courts will usually be receptive to scientific evidence — at least when it comes from traditional "hard" scientific disciplines. This is because such evidence can be used to bolster the legitimacy of a court's verdict and enables it to impose effective closure on a dispute. However, when it comes to dealing with scientific evidence, the legal system can also be its own worst enemy.

Whereas the legal system's normative authority derives in part from its categorical closure of disputes, the adversarial system and the fact that litigation is a zero-sum contest can make the resolution of a dispute more difficult.[40] By promoting "conflict perspectives,"[41] the legal process becomes a potent tool for un-

---

[38] *See* MICHAEL KING & CHRISTINE PIPER, HOW THE LAW THINKS ABOUT CHILDREN 28-30 (2nd ed. 1995) (focusing on demands upon legal profession to assimilate concepts from other discourses to achieve dispute resolution, depoliticization, and moral order); David Nelken, *Are Disputes Between Law and Science Resolvable?*, *in* FORENSIC EXPERTISE AND THE LAW OF EVIDENCE 104, 104-05 (J.F. Nijboer, et al. eds., 1993) (acknowledging that search for truth varies in different contexts and disagreements between law and sciences arise because questions raised in each sphere differ); *see generally* DAVID NELKEN, THE TRUTH ABOUT LAW'S TRUTH, (European University Institute Working Paper Law no. 90/1 (1990)) (discussing disagreements between law and science).

[39] *See* STANLEY FISH, THERE'S NO SUCH THING AS FREE SPEECH, AND IT'S A GOOD THING TOO 221-22 (1994) (stating that particular form in which materials from other disciplines enter law is determined by law's sense of its own purpose and usefulness to that purpose).

[40] *See* Niklas Luhmann, *The Unity of the Legal System*, *in* AUTOPOIETIC LAW: A NEW APPROACH TO LAW AND SOCIETY 12, 27-28 (Gunther Teubner ed., 1988) (suggesting that delays in dispute resolution arise because people cling to expectations of conflict when none exists).

[41] *Id.*

dermining the testimony of scientists. As it does, however, it destroys the very certainty that it needs if it is to resolve a dispute.

There are two important implications of this analysis of the relationship between science and law. First, when the legal system has to deal with knowledge from other disciplines, such as science, it will tend to do so in ways which simplify that knowledge.[42] An example of this is the way in which research on the limitations of eyewitness identification has been absorbed by the legal system.

In England, these limitations were brought to lawyers' attention after a notorious miscarriage of justice involving mistaken identification of the accused.[43] The English courts responded to the problem by laying down guidelines for cases which depend wholly or substantially on eyewitness evidence.[44] The guidelines provide that in such cases the trial judge should warn the jury in fairly general terms about the possibility of eyewitness error.[45] This obviates the need for expert witnesses to be called on eyewitness identification.

This is not a particularly scientific approach, and has been criticized by some.[46] However, it is a simple rule, and makes

---

[42] *See* KING & PIPER, *supra* note 38, at 122-23 ("this reductionism, far from being a failure of the law, is essential to its successful social functioning. Law's role is . . . to reduce complexity to manageable proportions in order that its normative communications may be acceptable to society".).

[43] *See* LORD DEVLIN, REPORT TO THE SECRETARY OF STATE FOR THE HOME DEPARTMENT OF THE DEPARTMENTAL COMMITTEE ON EVIDENCE OF IDENTIFICATION IN CRIMINAL CASES 33-66 (Her Majesty's Stationary Office ed., 1976) (discussing case in which accused was wrongly identified as criminal).

[44] These guidelines are to be found in *R. v. Turnbull*, [1977] 1 Q.B. 224, 228-31.

[45] *See id.* Though in extreme cases, the judge should withdraw the case from the jury. *See id.* at 229-30.

[46] *See, e.g.*, GEOFFREY M. STEPHENSON, THE PSYCHOLOGY OF CRIMINAL JUSTICE 155 (1992) ("The enthusiasm which psychologists have shown for research on eyewitness testimony has been matched by lawyers' and legal skepticism about the usefulness of psychologists insights."); *see also* Ayre Rattner, *Social Science v. The Judicial System: The Impact of Accumulated Knowledge of Eyewitness Identification on Criminal Procedures*, 23 INT'L J. SOC. L. 97, 104 (1995) (reviewing American law and noting that law attends to what social scientists have to say from legal angle rather than necessary scientific tests). Nelken has drawn attention to the threat that the assimilation of psychological knowledge may pose to the law, because "[i]f we accept the appropriateness of psychological criticisms of the reliability of witness identifications in the criminal process . . . , why should we not accept other psychological studies whose arguments could bring us as far as replacing the accusatorial trial with a psychological tribunal?" NELKEN, *supra* note 38, at 10.

the task of appellate courts easier. When eyewitness evidence is challenged on appeal, the court can generally restrict its review to asking whether the requisite warning was given to the jury. If it was, the conviction will be sustained.

A second implication of the legal system's autonomy from other social systems, such as science, is that there is ample scope for each discipline to misinterpret the other. Law and science are good at talking past one another. This can deepen the problem of expert disagreement. We will find that this *dialogue des sourds* is a recurring feature of scientific disputes in the courts.

With this theoretical framework to analyze scientific disagreement, and the legal system's response to it, this Article now examines three examples in some detail. The examples are partly chosen for the differences between them: one is from the English legal system, the other two from the United States; one involves civil litigation, the other two criminal litigation. All involve very different levels of controversy. But beyond these differences, there are a number of similarities between the disputes examined.

In discussing these scientific disputes, this Article attempts to avoid approaching them with the wisdom of hindsight. It is easy to look back on episodes in science where consensus has now been achieved and to wonder why there was disagreement at that time. The argument that won may appear to have been the most cogent all along. But, if we are to understand why these episodes have been so problematic, we must remember that, during them, there may have been no good reasons to prefer the arguments of one side to those of the other.[47]

---

[47] Kuhn makes a similar point about scientific revolutions. He argues that, writing with hindsight, textbook writers have been able to portray the history of science as one of rational, incremental progress, thus obscuring the profound discontinuity of theories and concepts that he sees as characterizing scientific progress. *See* THOMAS S. KUHN, THE STRUCTURE OF SCIENTIFIC REVOLUTIONS 136-43 (2d ed. 1970); *see also* Thomas Nickles, *Good Science as Bad History: From Order of Knowing to Order of Being*, *in* THE SOCIAL DIMENSIONS OF SCIENCE (Ernan McMullin ed., 1992).

II. CASE STUDIES OF DISPUTES OVER SCIENTIFIC EVIDENCE

A. *The* Maguire *Case*

In the early 1990s, the English legal system was shaken as a number of high profile miscarriage of justice cases came to light.[48] Several of these cases demonstrated vividly the fallible nature of scientific evidence.[49] In the most notorious cases, defendants who had been convicted of involvement in the IRA bombing campaign of the mid-1970s were able to show that the scientific evidence that had formed a major part of the evidence against them was deeply flawed. This section focuses on one of those cases to illustrate some of the problems that disputed scientific evidence can pose for the legal system.

In 1976, members of the Maguire family were tried for explosives offenses. Although the Maguires were unlikely terrorists, they were found guilty, almost entirely on the evidence of forensic scientists who testified that swabs taken from the defendants' hands had revealed traces of the explosive nitroglycerine.[50] The swabs had been tested using a method called thin layer chromatography (TLC). Although no explosives were ever found in the Maguires' house, scientists testified that traces of nitroglycerine found under the defendants' fingernails proved that they had actually kneaded explosives. This conclusion rested on the assumption that it is impossible to get nitroglycerine under the fingernails through innocent contamination, say by handling an object previously touched by someone else with nitroglycerine on their hands.

---

[48] *See* Joshua Rozenberg, *Miscarriages of Justice*, *in* CRIMINAL JUSTICE UNDER STRESS 91, 91-117 (Eric Stockdale & Silvia Casale eds., 1992) (detailing several high profile cases preceding calls for reform of criminal appeals process).

[49] *See* BRIAN H. KAYE, SCIENCE AND THE DETECTIVE: SELECTED READING IN FORENSIC SCIENCE 120-22 (1995) (explaining reasons for failure of scientific evidence in cases known as Birmingham Six, Guildford Four, and Maguire Seven); Alec Samuels, *Forensic Science and Miscarriages of Justice*, 34 MED. SCI. & L. 148, 148-49 (1994) (delineating cases exposing fallibility of scientific evidence); *see also* CAROL A.G. JONES, EXPERT WITNESSES: SCIENCE, MEDICINE AND THE PRACTICE OF LAW 195-223 (1994) (espousing notion that fallible scientific evidence and consequential miscarriage of justice derive from corroborative efforts of prosecution and its forensics and pathology team); Clive Walker & Russel Stockdale, *Forensic Evidence and Terrorist Trials in the United Kingdom*, 54 CAMBRIDGE L.J. 69, 74-97 (1995) (discussing fallibility of forensic evidence and possible remedies in light of United Kingdom's terrorist trials).

[50] For an account of the case, see ROBERT KEE, TRIAL AND ERROR 199-223 (1986).

During the Maguire trial, the defense called an expert to challenge the reliability of the scientific evidence. The defense expert criticized the prosecution's evidence on three grounds.[51] First, there were certain inconsistencies in the TLC results that undermined them. Second, the TLC test by itself was not reliable and a confirmatory test should have been used. Finally, the TLC test was not specific enough to uniquely identify nitroglycerine. Notwithstanding these criticisms, the Maguires were convicted.

As scientific disagreement goes, this was a small-scale affair. The disagreement between defense and prosecution witnesses certainly did not approach the status of the wide-ranging and public controversy that has recently surrounded DNA evidence. Closure of this particular dispute was, at first, achieved with little difficulty when the jury's verdict indicated that it preferred the evidence of the prosecution experts to that of the defense. This is the legal system's usual manner of resolving expert disputes. However, unease about the scientific evidence in *Maguire* did not disappear and the closure that had been imposed by the legal system became increasingly hard to sustain.

In the long run, all of the defense expert's criticisms proved to have merit. An in-depth judicial inquiry conducted by Sir John May has revealed just how unreliable the prosecution's scientific evidence was.[52] One of the interesting things revealed by the inquiry is how difficult it can be to resolve disputed scientific issues. This is so even when the dispute takes place at one removed from the adversarial structure of the legal system.

One question that Sir John May's inquiry examined was whether the prosecution scientists' conclusion that nitroglycerine could only get under a person's fingernails through the

---

[51] *See id.* at 208-10 (noting defense's arguments that TLC test had one in sixteen chance of false positive, reacted to substances other than nitroglycerine, and produced contrary evidence when tested by defense).

[52] The report on the *Maguire* case was published in two parts: Sir John May, *Return to an Address of the Honourable the House of Commons Dated 12 July 1990 for the Inquiry into the Circumstances Surrounding the Convictions Arising out of the Bomb Attacks in Guildford and Woolwich in 1974: Interim Report on the* Maguire *Case*, (H.C. 556, 1990) [hereinafter May, *Interim Report*]; Sir John May, *Return to an Address of the Honourable the House of Commons Dated 3 December 1992 for the Report of the Inquiry into the Circumstances Surrounding the Convictions Arising out of the Bomb Attacks in Guildford and Woolwich in 1974: Second Report on the* Maguire *Case* (H.C. 296, 1992) [hereinafter May, *Second Report*].

kneading of explosives was sound. The May inquiry instructed a forensic scientist to examine this "kneading hypothesis." In May's first report, it was accepted that this scientist had demonstrated just how easily a person's hands could become innocently contaminated with nitroglycerine.[53]

Significant amounts of nitroglycerine could be picked up just by drying one's hands on a towel. Moreover, this nitroglycerine could then migrate from the palms of the hands to become embedded under the fingernails. However, during the course of the inquiry, these conclusions were themselves challenged by scientists who contended that the results were invalid owing to the possibility that cross-contamination had occurred during the experiments.[54]

Sir John May then set up a scientific committee to try to settle the issue. While this committee agreed that the supposed falsification of the kneading hypothesis by the earlier experiment was inconclusive, it was not prepared to declare the kneading hypothesis correct. However, the committee members could not agree on where their work left the positive results from the original trial. While some of the committee thought that the ease with which nitroglycerine traces could be spread[55] left contamination a likely explanation, others thought that contamination did not provide a realistic explanation for the original results.[56] Some of the committee members have now published research which, they claim, explains the pattern of the original results.[57]

In science, even a well resourced and independent committee may find it difficult to achieve the closure of scientific disagreement.[58] This is in marked contrast to the ease with which the

---

[53] *See* May, *Interim Report*, *supra* note 52, at 29.

[54] *See* May, *Second Report*, *supra* note 52, at 6-7.

[55] For example, the very process of swabbing someone's hands may force some nitroglycerine under the fingernails. *See id.* at 7.

[56] *See id.* at 9.

[57] *See* E. McKenzie et al., *A Statistical Model For the Response Patterns to Chemical Tests For the Absence or Presence of Trace Materials*, 35 SCI. & JUST. 31 (1995). The conclusions of these members of the committee have been questioned on the grounds that they are too artificial an attempt to reproduce the original results. *See* May, *Second Report*, *supra* note 52, at 10 (contending that *Maguire* case experiments were designed to produce certain results).

[58] *See* Walker & Stockdale, *supra* note 49, at 86 (noting difficulty of drawing scientific conclusions).

legal system achieved a procedural closure of the disagreement between the experts at the original trial. Much of law's authority derives from the process by which the verdict of a court imposes closure on what may be bitterly contested issues. *Maguire* demonstrates the way in which the use of scientific evidence may ultimately render such closure — and in the process, law's authority — problematic. This brings us to the heart of some of the differences between science and law.

When a court is confronted with a disagreement between experts, it is not able — unlike the May Inquiry — to order investigations into the disputed issue in an attempt to reach consensus. In the words of Lord Scarman:

> "The investigation of facts and the discovery of truth . . . are subordinate as far as law is concerned to the requirement that disputes be settled and justice be done. Justice cannot wait upon truth. If the truth is not discoverable, justice still has to be done for justice delayed is justice denied."[59]

Justice and finality take precedence over truth, but science progresses. By 1982, research had been published that threw doubt on the kneading hypothesis.[60] The Home Office by 1986 had conceded that, rather than relying on a TLC test alone, "it would now be standard practice to establish the presence of nitroglycerine through confirmatory tests."[61]

Doubt over the convictions of the Maguire family meant that there was continuous pressure for the case to be referred back to the Court of Appeal. The Court of Appeal, however, has often been reluctant to grant appeals which rely solely on the ground of mistake of fact by the trial court.[62] This reluctance stems in part from the need to secure finality in the legal sys-

---

[59] Kate Malleson, *Appeals Against Conviction and the Principle of Finality*, 21 J.L. & SOC'Y 151, 158 (1994) (quoting Lord Scarman, *Truth and the Legal Process*, E.H. Young Memorial Lecture (1976)).

[60] J.D. Twibell et al., *Transfer of Nitroglycerine to Hands During Contact with Commercial Explosives*, 27 J. FORENSIC SCI. 783, 783 (1982) (describing 1977 experiment measuring persistence of nitroglycerine residue on hands after handling explosives).

[61] *See* KEE, *supra* note 50, at 204 n.* (quoting letter from the Home Office (Apr. 28, 1996)).

[62] *See* Malleson, *supra* note 59, at 152 (discussing Court of Appeal's reluctance to widely exercise its powers of review).

tem, and in part from an ideological commitment to trial by jury which leads the Court of Appeal to take a deferential attitude towards jury verdicts.

The difficulty in reconciling scientific and legal procedures is illustrated by discussions about the scientific evidence in *Maguire* that took place within the Home Office while it was deciding whether or not the case should be referred back to the Court of Appeal.[63] A scientific inquiry, like that eventually instigated by Sir John May, would have been the obvious scientific way to deal with unease about the evidence.

Officials in the Home Office were reluctant to take this course. Their concerns were influenced by the need for finality in the legal system. They did not want to create further scientific uncertainty about the weight of the original evidence, nor did they want to open the floodgates for referring other cases to the Court of Appeal where science had evolved since the time of the conviction.[64]

Officials were reluctant to form a scientific committee "to determine a question which should more properly be considered by the courts,"[65] especially as such a committee might approach the matter "with an over-concern with scientific proof and certainty."[66] Tellingly, the Home Secretary referred to a distinction between new evidence, which would justify referring the case to the appeal court, and differences of opinion about old evidence, which would not.[67] Whereas both are significant in science, it was only the former that was deemed to be legally relevant.

When, finally, *Maguire* was referred to the Court of Appeal, the Maguires' convictions were quashed.[68] The Court of Appeal allowed that the possibility of innocent contamination rendered

---

[63] Under the legislation then in force, the Home Secretary was empowered to refer cases to the Court of Appeal after an original appeal had failed. For a description of the Home Office's power to refer cases to the Court, see ROSEMARY PATTENDEN, ENGLISH CRIMINAL APPEALS 1844-1994, at 346-412 (1996) (discussing Home Office's role in post-appeal remedies).

[64] *See* May, *Second Report*, *supra* note 52, at 80 (discussing how developments in scientific technique can cast undue doubt on prior convictions).

[65] *Id.* at 81.

[66] *Id.* at 82.

[67] *See id.* at 84-85 (discussing different considerations involved in examining new and old evidence).

[68] *See* R. v. Maguire, [1992] 2 All E.R. 433, 450 (C.A. 1991).

the verdicts unsafe and unsatisfactory. This, however, was a particularly narrow ground on which to quash the convictions: the Court ignored a number of other important deficiencies in the original scientific evidence, and was criticized by Sir John May for doing so.[69] Even though the Crown chose not to dispute the innocent contamination issue, the court was reluctant to accept uncontested scientific evidence without both allowing for cross-examination and checking the coherence of the original jury verdict against the new evidence.[70]

In another appeal involving discredited scientific evidence, the Court took a similar approach. Instead of accepting that the uncertainty surrounding the original scientific evidence rendered the verdict unreliable, the Court effectively imposed a reverse onus of proof. It demanded that the appellants provide an alternative, more likely explanation for the original results.[71] Rather than accepting good scientific practice, which would admit that unreliable methodology may provide unreliable results, the Court will seek to uphold a jury verdict on any tenable ground. By being partially deaf to science, the legal process can resist the threat to its procedures that science poses.[72]

Earlier it was suggested that burdens of proof — whether the formal ones of law or the strategic ones of scientific debate — play an important role in determining approaches to uncertainty. The handling of scientific evidence in *Maguire* can be criticized because it was not scrutinized in light of the high burden of proof required in criminal cases. There were numerous examples of this, implying that, at various stages of the case, its outcome would have been very different had the approach to uncertainty been different. For example, when examining the response of Home Office officials to reviews of the scientific evidence in the case, Sir John May noted that an appropriate appraisal of the evidence could only have been made by a lawyer,

---

[69] *See* May, *Second Report*, *supra* note 52, 2-3 (discussing John May's criticism of court); Heather Mills, *Slur Remains After Maguire Appeals Won*, INDEPENDENT, June 27, 1992, at 1.

[70] *See* Richard Nobles & David Schiff, *Miscarriages of Justice: A Systems Approach*, 58 MOD. L. REV. 299, 307-10 (1995) (discussing Court's reluctance to accept scientific evidence).

[71] *See* Richard Nobles et al., *The Inevitability of Crisis in Criminal Appeals*, 21 INT'L J. SOC. L. 1, 1-19 (1993) (discussing inevitability of mistakes in criminal appeals regardless of whatever reforms are taken).

[72] *See id.* at 302-10 (discussing miscarriages of justice in legal system).

"with a lawyer's practical appreciation of evidence and the burden of proof."[73] Had this occurred, May suggested, it would have been apparent that there were serious doubts about whether the scientific evidence was still sufficient to sustain the convictions.

Another criticism involves the handling of the dispute over the specificity of the TLC test for nitroglycerine.[74] During the original trial, the scientist called by the defense argued that the TLC test was not specific for nitroglycerine, though this was denied by the prosecution. Later in the trial it was revealed that the TLC test would also give a positive result for PETN, another explosive.

The handling of this revelation has been the subject of considerable criticism. The prosecution and the trial judge (later joined by judges in the Court of Appeal) played down the significance of the revelation on the grounds that PETN was another explosive. However, it should have been apparent that, "on the evidence, the Crown [was] unable to prove a material allegation in the indictment — namely the presence of [nitroglycerine]."[75]

The question of the specificity of the TLC test was fertile ground for miscommunication between science and law. May suggested that the scientists may have failed to grasp the importance of PETN "in legal not scientific terms."[76] Likewise, the lawyers did not seem to grasp the extent to which this fact cast doubt on the reliability of the scientific test.

Finally, the disagreement about the scientific evidence in *Maguire* should be seen in the light of English law on scientific evidence. In the United States, most academic debate about

---

[73] May, *Second Report*, *supra* note 52, at 91.

[74] *See* KEE, *supra* note 50, at 212-15, 232-34 (discussing and criticizing various Justices' treatments of TLC test results).

[75] May, *Interim Report*, *supra* note 52, at 35. In fact, May's comment overstates the problem. The Crown's ability to use the scientific evidence to prove the presence of nitroglycerine would have depended on the jury's assessment of all of the other evidence in the case. However, given the almost total absence of other evidence against the Maguires, the comment is unsurprising.

[76] *Id.* at 40.

scientific evidence has revolved around the role that courts should play in screening scientific evidence before admitting it. In England, this debate is hardly recognized as an issue.

Under English law there is currently no special evidentiary hurdle that must be passed before the results of novel scientific techniques are admitted into evidence.[77] In criminal cases this deficiency is problematic. For example, it has been suggested that the absence of an exclusionary rule specifically for scientific evidence places an informal burden of proof on the accused. The defense is expected to show why a scientific technique used by the prosecution is not reliable.[78]

This is not to suggest that, had English law had a *Frye*-type exclusionary rule, the TLC evidence in *Maguire* would definitely have been ruled inadmissible. But such a rule would have given the defense a better opportunity to argue for close scrutiny of the scientific evidence. It might also have made it easier for the defense to successfully appeal when doubts about the reliability of the scientific evidence came to light.

Maguire illustrates a number of important themes of this Article. The case shows just how difficult it can be for scientists to reach agreement on disputed issues. It also exemplifies some of the differences between science and the legal process. Disagreement is normal in science; it must be accommodated in scientific practice and may be resolved over time. But this may be perceived as a threat by the legal system that seeks finality and values lay fact-finding. When confronted with changes in scientific consensus, the legal system, with its emphasis on procedural truth, may prove resistant to the new scientific truth.

---

[77] *See generally* David E. Bernstein, *Junk Science in the United States and the Commonwealth*, 21 YALE J. INT'L L. 123, 166-73 (1996) (discussing admissibility of scientific evidence in England); Ian Freckelton, *Science and the Legal Culture*, 2 EXPERT EVIDENCE 107 *passim* (1993) (discussing relationship between science and courts in common law countries). For criticism of English law on this point, see Peter Alldridge, *Recognising Novel Scientific Techniques: DNA as a Test Case*, 1992 CRIM. L. REV. 687, 688-98; Paul Roberts, *The Admissibility of Expert Evidence: Lessons From America*, *in* 4 EXPERT EVIDENCE 93, 93-99 (1996) (discussing lessons from recent United States legal developments).

[78] *See* BEVERLEY STEVENTON, THE ROYAL COMMISSION ON CRIMINAL JUSTICE, THE ABILITY TO CHALLENGE DNA EVIDENCE 37 (1993) (arguing that jury's assessment requires defense to challenge evidence).

B. The DNA Controversy

During the early stages of its development, criminal courts readily accepted evidence based on DNA profiling techniques.[79] It was only when scientists from the wider scientific community became acquainted with how DNA technology was being applied forensically that doubts about the reliability of DNA evidence were brought to the attention of the legal system. In response, courts began to rule DNA evidence inadmissible.[80] DNA's fall from grace was made even more spectacular by the degree of faith that had been placed in it by scientists,[81] judges,[82] academics,[83] politicians,[84] and the media.[85] However, as with several other technologies,[86] the viewpoint that DNA evidence was infallible concealed a number of weaknesses in the technique, now open to re-evaluation.

---

[79] *See* D.H. Kaye, *The Admissibility of DNA Testing*, 13 CARDOZO L. REV. 353, 357 (1991) (observing that "in [the] first wave of cases, expert testimony for the prosecution rarely was countered, and courts readily admitted [DNA evidence]").

[80] *See* David H. Kaye, *DNA Evidence: Probability, Population Genetics, and the Courts*, 7 HARV. J.L. & TECH. 101 *passim* (1993) (citing several criticisms); William C. Thompson, *Evaluating the Admissibility of New Genetic Identification Tests: Lessons From the "DNA War"*, 84 J. CRIM. L. & CRIMINOLOGY 22, 22-104 (1993) (reviewing and criticizing both DNA testing and legal system's response).

[81] *See, e.g.*, A.J. Jeffreys et al., *Individual-Specific "Fingerprints" of Human DNA*, 316 NATURE 76 (1985) (describing statistical accuracy of genetic testing).

[82] For judicial dicta extolling the virtues of DNA evidence, see Jonathan J. Koehler, *DNA Matches and Statistics: Important Questions, Surprising Answers*, 76 JUDICATURE 222, 222 n.1 (1993).

[83] *See* Robin M. White & Jeremy J.D. Greenwood, *DNA Fingerprinting and the Law*, 51 MOD. L. REV. 145, 145 (1988) (calling DNA evidence "indisputably accurate"); I.H. Dennis, *Reconstructing the Law of Criminal Evidence*, 42 CURRENT LEGAL PROBS. 21, 43-44 (1989) (suggesting DNA evidence provides "conclusive proof of guilt or innocence").

[84] According to the chair of Parliament's Home Affairs Committee, a national DNA database would lead to "a sharp reduction in violent crimes." Quentin Cowdry, *MPs Back Police Chiefs' Call for a DNA Register*, TIMES (London), Jan. 16, 1991, at 3.

[85] DNA evidence has been claimed to be "the advance of the century" and to be "100 per cent reliable." *See* Craig Seton, *DNA 'Advance of Century'*, TIMES (London), Nov. 14, 1987, at 3; Craig Seton, *Genetic Tests '100% Reliable'*, TIMES (London), Jan. 6, 1987, at 3; *see also* DOROTHY NELKIN & M. SUSAN LINDEE, THE DNA MYSTIQUE: THE GENE AS CULTURAL ICON 47-49 (1995) (discussing DNA reliability being embraced by media and popular culture); Koehler, *supra* note 82, at 222 n.1 (noting exaggerated claims for DNA evidence in press).

[86] *See* DONALD A. MACKENZIE, INVENTING ACCURACY: A HISTORICAL SOCIOLOGY OF NUCLEAR MISSLE GUIDANCE 340-81, 417-23 (1990) (examining claims about accuracy of nuclear missiles).

Criticisms of the use of DNA profiling in the courts have focused on the way DNA technology was transferred from diagnostic to forensic work,[87] the subjectivity of the criteria used to compare DNA profiles,[88] and the impact of error rates on match probabilities.[89] Although these are all important criticisms, this Article will concentrate on what has proved to be the most controversial issue — the one least amenable to easy resolution. That is the debate about the statistical assumptions underlying the calculation of DNA match probabilities.

A DNA "match probability" is the figure that represents the probability that a randomly selected member of the population would match the crime scene DNA. These probabilities can be tiny; figures such as one in hundreds of millions are commonly quoted. The process used by forensic science laboratories to calculate match probabilities conforms, generally, to the following pattern.[90]

"DNA profiling" samples portions of a subject's DNA at several loci (specific locations) on the genome. Each locus will usually produce two alleles (alternative forms of the DNA at the locus sampled), one from each parent. To calculate the match probability, scientists must first calculate the probability of finding the subject's genotype (the combination of two alleles) at the relevant locus. Allele frequencies are calculated using a database, and these frequencies are combined to calculate the

---

[87] *See* Jon Turney, *Gene Believers*, TIMES (LONDON) HIGHER EDUC. SUPPLEMENT, Nov. 3, 1989, at 17; Eric Lander, *DNA Fingerprinting: Science, Law and the Ultimate Identifier*, *in* THE CODE OF CODES: SCIENTIFIC AND SOCIAL ISSUES IN THE HUMAN GENOME PROJECT 191, 193-207 (Daniel J. Kevles & Leroy Hood eds., 1992) (discussing early use of DNA fingerprinting in trials).

[88] *See* William C. Thompson, *Subjective Interpretation, Laboratory Error and the Value of Forensic DNA Evidence: Three Case Studies*, 96 GENETICA 153, 165-68 (1995); William C. Thompson & Simon Ford, *The Meaning of a Match: Sources of Ambiguity in the Interpretation of DNA Prints*, *in* FORENSIC DNA TECHNOLOGY 93 (M. Farley & J. Harrington eds., 1991).

[89] *See* Edward J. Imwinkelried, *The Debate in the DNA Cases Over the Foundation for the Admission of Scientific Evidence: The Importance of Human Error as a Cause of Forensic Misanalysis*, 69 WASH. U. L.Q. 19, 24-33 (1991) (discussing effect of DNA technician error); Koehler, *supra* note 82, at 228-29 (discussing relevance of DNA test error rates); Jonathan J. Koehler et al., *The Random Match Probability in DNA Evidence: Irrelevant and Prejudicial?*, 35 JURIMETRICS J. 201, 201-19 (1995) (showing limits of DNA random match probability and arguing prejudicial effect on jurors).

[90] For more detailed accounts, see Kaye, *supra* note 80, at 107-27 (describing DNA profiling procedure); Thompson, *supra* note 80, at 26-30, 33-42, 61-84 (describing DNA technology and profiling procedures).

genotype frequency. Once all of the relevant genotype frequencies are known, they are multiplied together to calculate the probability of finding the multilocus genotype; that is, the particular combination of alleles found. The use of multiplication at this stage has been called the "product rule."

This process of match probability calculation has been controversial because it relies on several assumptions from the field of population genetics. The rule for calculating genotype frequencies assumes that the alleles at each locus are combined at random. In population genetics jargon, it assumes that the population in question is in Hardy-Weinberg equilibrium. In plain terms, it assumes that people mate, and genetic material is transmitted at random. The use of the product rule to calculate the probability of finding the multilocus genotype also rests on an assumption. It assumes that the alleles at each locus are independent of each other; in technical terms, that they are in linkage equilibrium.

Around 1990, critics began to challenge the use of these assumptions in the calculation of DNA match probabilities by forensic laboratories. It was argued that "some people may be going to jail because statistical independence has been declared in forensic applications of DNA fingerprinting without anyone ever collecting the data to justify it."[91] Critics argued that it was unrealistic to assume Hardy-Weinberg equilibrium in human populations, and also claimed that forensic databases showed evidence of an absence of Hardy-Weinberg equilibrium.[92] It was asserted that significant amounts of population substructure in racial subgroups, such as Italians and Poles, meant that "the probability of a random match . . . cannot be estimated reliably for 'Caucasians,' probably not for 'blacks,' and certainly not for 'Hispanics.'"[93] This controversy has not proved easy to resolve. Only in 1996, some six years after doubts about the reliability of

[91] Joel E. Cohen, *DNA Fingerprinting: What (Really) Are the Odds?*, 3 CHANCE 26, 26 (1990).

[92] *See* Eric S. Lander, *DNA Fingerprinting on Trial*, 339 NATURE 501, 504 (1989) (discussing equilibrium deviations in Hispanic population).

[93] R.C. Lewontin & Daniel L. Hartl, *Population Genetics in Forensic DNA Typing*, 254 SCIENCE 1745, 1749 (1991) (criticizing present DNA matching techniques and databases).

DNA profiling were first raised, are there signs that the controversy over the calculation of match probabilities is drawing to a close.[94]

Why has it proved so difficult to achieve a consensus in the scientific community over the evidential use of DNA profiling? One reason is that many of the issues underlying the DNA controversy have been, in Weinberg's terms, trans-scientific. That is, scientists have not been able to answer them unequivocally. This is because of two interconnected reasons: scientists have had to cope with a considerable degree of uncertainty surrounding the calculation of match probabilities, and they have been dealing with issues that are intimately connected to legal policy.

1. Uncertainty

When DNA evidence was first used in the courts, and for some time after that,[95] the assumptions underlying match probability calculation were underdetermined by the available empirical data. While it was quite possible that the assumptions were valid, this could not be proved to the satisfaction of some critics. Some commentators suggest that we can divide the actors in the DNA controversy into two schools, which I call the pragmatic

---

[94] I say this with some caution. However, brief reports canvassing reactions to a new report on DNA profiling (discussed *infra* notes 115-38 and accompanying text) suggest that criticism of match probability calculation is dying away. *See* Colin Macilwain, *Ceiling Principle "Not Needed" in DNA Cases*, 381 NATURE 103 (1996) (discussing report's results, implications, and reactions to them); Eliot Marshall, *Academy's About Face on Forensic DNA*, 272 SCIENCE 803 (1996) (summarizing report and several reactions). The *Science* report quotes Richard Lewontin, one of the most vehement critics of the way DNA evidence has been used in the courts, as saying that the population genetics question is "not at the center" of the debate any more. Marshall, *supra*, at 803. However, a few commentators are still critical of the assumptions underlying the way match probabilities are calculated in some cases. *See Discussion of the Paper by Balding and Donnelly*, 158A J. ROYAL STAT. SOC'Y 40 (1995) (collecting several comments on their paper regarding probability of determining suspect's guilt); Stanley Sawyer et al., *DNA Fingerprinting Loci Do Show Population Differences: Comments on Budowle et al.*, 59 AM. J. HUM. GENETICS 272, 272-74 (1996) (arguing in support of DNA fingerprinting).

[95] Commentators would doubtless disagree on the question of just when there was enough empirical data for them to be confident about the assumptions used in match probability calculation. Lempert suggests that the tide turned around 1992. *See* Richard Lempert, *DNA, Science and the Law: Two Cheers For the Ceiling Principle*, 34 JURIMETRICS J. 41, 45 n.10 (1993) (discussing results of empirical research). Lander and Budowle's discussion suggests a later date. *See* Eric S. Lander & Bruce Budowle, *DNA Fingerprinting Dispute Laid to Rest*, 371 NATURE 735, 738 (1994) (discussing FBI population surveys).

and empirical schools.[96] What divides these two schools is their approaches to uncertainty.

The pragmatic school acknowledges that there possibly may be a significant degree of substructure in relevant populations. They argue, however, that there is no convincing evidence that substructure affects forensic match probability calculation. Pragmatists are prepared to dismiss the evidence of substructure put forward by critics, often on theoretical grounds.[97] For those taking an empirical approach, the theoretical assumptions under lying the pragmatist's stance are unwarranted. The uncertainty over whether or not substructure exists, they argue, can only be reduced through empirical research.[98]

The difference between the empirical and pragmatic approaches to uncertainty maps onto another characteristic of the DNA controversy: the strategic use of burdens of proof by actors in the debate. In a debate such as this, where the validity of the assumptions at issue is underdetermined, there is a substantial advantage in throwing the burden of proof onto one's opponents.[99] Thus, pragmatists counter criticisms of the product rule

---

[96] *See* Eric S. Lander, *Lander Reply*, 49 AM. J. HUM. GENETICS 899, 901 (1991) (suggesting dividing actors in controversy into "keep it simple," "statistical," and "empirical" schools); Thompson, *supra* note 80, at 70-71 n.215 (adapting Lander's terminology to "theoretical," "statistical," and "empiricist" schools). My own terminology conflates Thompson's theoretical and statistical schools to produce the pragmatic school. It is interesting to note that there has, of late, been a reversal of such terminology, with proponents of DNA evidence claiming empirical support and denigrating critics for being too theoretical. *See* Kathryn Roeder, *DNA Fingerprinting: A Review of the Controversy*, 9 STAT. SCI. 222, 223 (1994) (stating that "[a]lthough many of the arguments put forth by the critics of current methods of evaluating DNA evidence are theoretically correct, my conclusions are that the data do not support their claims"); *Discussion of the Paper by Balding and Donnelly, supra* note 94 (noting several commentators suggest that Balding and Donnelly's paper is too theoretical to provide any serious grounds for criticizing DNA evidence).

[97] *See, e.g.*, Roeder, *supra* note 96, at 230 (arguing that Lander's claim to have found an excess of homozygotes in a forensic database is unlikely to be true, on account of improbability of such degree of homozygosity occurring).

[98] *See, e.g.*, David J. Balding et al., *Comment: Some Causes for Concern About DNA Profiles*, 9 STAT. SCI. 248, 250 (1994) (stating that answering the question of whether there is large interracial differentiation in genotype frequencies "will require substantial surveys at appropriate levels of stratification at each of the loci in forensic use"); Eric S. Lander, *Invited Editorial: Research on DNA Typing Catching up with Courtroom Application*, 48 AM. J. HUM. GENETICS 819, 821 (1991) (stating that sampling ethnically distinct populations and observing actual degree of genetic differentiation answers question of population heterogeneity).

[99] *See supra* note 33 and accompanying text (explaining advantages of shifting burden of proof).

arguing that "multiplication provides an appropriate approximation to genotype frequencies until proven otherwise."[100] In contrast, the empirical critique is that proponents of DNA evidence bear the responsibility of proving that the assumptions they are making about the genetic structure of human populations are valid. The way in which the debate has been structured by burdens of proof has been recognized by some of the more astute commentators.[101]

2. Legal Policy

The other reason why the debate over DNA match probability calculation has been trans-scientific is that the scientific issues could not easily be separated from questions of legal policy. Rather than being a simple debate about science, it became a "mixed" controversy,[102] involving both scientific and legal issues. The way in which the debate evolved is closely connected to the uncertainty affecting the assumptions used in match probability calculation.

In criminal litigation, the prosecution bears the burden of proving the case against the defendant beyond reasonable doubt. This is because of the normative structure of the criminal process, which requires that the innocent be protected from wrongful conviction. Generally speaking, defendants should be given the benefit of doubts in the prosecution's case. This then allows the argument that any doubts attaching to match probability calculation should be resolved in ways that will favor defendants. We should, then, demand that the procedures used to calculate DNA match probabilities produce conservative figures.

Critics of DNA profiling do not always make this argument explicit. Perhaps they do not wish to make it look as if their arguments are based on policy rather than on science. But it

---

[100] B. Devlin et al., *Response*, 253 SCIENCE 1039, 1041 (1991).

[101] *See* Lander, *supra* note 96, at 902 (observing that Chakraborty "appears to feel that absence of significant population structure should be *assumed* until proven otherwise (and thus that observed deviations from HWE and LE should not be taken seriously unless all possible artifacts have been eliminated)"). Lander also notes that other critics "seem to feel that proponents of statistical evidence bear the burden of unambiguously demonstrating independence (and, therefore, possible artifactual explanations for observed deviations should not be accepted until proved experimentally)." *Id.*

[102] *See* McMullin, *supra* note 27, at 75-77 (discussing mixed controversies).

undoubtedly informs the general tenor of criticism of DNA evidence, especially the demand that proponents of DNA evidence provide convincing proof of the absence of population substructure.

Occasionally, critics do refer explicitly to legal concepts. For example, one commentator criticizes laboratory matching rules by arguing: "If conclusions must be proved beyond a reasonable doubt, it might be wise to use the 99% upper limit of the confidence interval for each allele."[103] Another has connected the issue of laboratory error rates to the concept of reasonable doubt.[104] Other actors in the controversy have been, perhaps somewhat wilfully, blind to the way in which arguments about the use of DNA evidence in the courts have been structured around the values of the criminal process. For example, some contend that it is inconsistent for lawyers to argue that DNA evidence should be used to acquit defendants but not to convict them.[105]

Obviously, one cannot demand that the proponents of DNA evidence be absolutely certain that the assumptions underlying their calculation of match probabilities are correct. The fallibility of science counsels against demanding such a high level of confidence in the techniques used in the courts. Therefore, it seems that much of the debate about DNA evidence can be seen in terms of a debate about the appropriate degree of conservatism to be factored into match probability calculation. During the

---

[103] Eric S. Lander, *Population Genetic Considerations in the Forensic Use of DNA Typing*, *in* DNA TECHNOLOGY AND FORENSIC SCIENCE 143, 147 (Jack Ballantyne et al. eds., 1989). Lander's contention is certainly arguable. First, one should be wary of mapping statistical conventions onto legal standards of proof. Second, any strict matching standards tend to ignore the cumulative effect of evidence: a number of alleles that do not match exactly may amount to strong proof, especially where there is other evidence against a defendant.

[104] *See* R.C. Lewontin, *Comment: The Use of DNA Profiles in Forensic Contexts*, 9 STAT. SCI. 259, 261 (1994) (discussing correlation between reasonable doubt and potentially bad research data). Lewontin criticized current assumptions because, "in a system of justice in which defendants are to be convicted on the basis of 'no reasonable doubt,' no guidelines are offered for what error rate would constitute 'reasonable doubt.'" *Id.*

[105] *See* Rockne P. Harmon, *Legal Criticisms of DNA Typing: Where's the Beef?*, 84 J. CRIM. L. & CRIMINOL. 175, 176 (1993) (refuting argument that DNA evidence is applied inconsistently).

early 1990s, it seemed that the hypothesis put forward by critics — that there is a significant degree of substructure in human populations — might turn out to be validated.

How, then, should the legal system have dealt with this? Put another way, how far should the criminal process go in making assumptions that favor defendants? There is no easy answer to this question, but it is obvious that it is not a scientific question. Rather, it is a question of legal policy, to be decided by the legal institutions. But if the law is to adopt a conservative approach to scientific evidence, it must tread warily. The preferred solution should rest on a scientifically valid estimate of the likelihood that a phenomenon such as substructure exists.

3. Which Database?

Some of the themes discussed above can be illustrated by an issue that has dogged arguments about forensic science agencies' abilities to calculate reliable match probabilities. This is the question of whether the defendant's ethnicity is relevant to match probability calculation. It is known that genotype frequencies vary between races; a genotype uncommon in one race may be more common in another. It has also been argued that, within races, there may be subgroups in which genotype frequencies vary even more.

One example of this was *People v. Mohit*.[106] In this case, the defendant was a Shiite Muslim from a community in Iran in which, apparently, marriage between cousins was common. One would expect the defendant's DNA profile to be more common among those who share his ethnicity than among the US population in general. This concerned the court because, not surprisingly, the FBI did not have a separate database for Iranian Shiite Muslims.

The argument that the defendant's ethnicity is relevant to match probability calculation has been called "one of the most persistent fallacies" in the forensic DNA debate.[107] It is not difficult to see why. If it is presumed that the defendant is inno-

[106] 579 N.Y.S.2d 990 (Sup. Ct. 1992).

[107] B.S. Weir, *Population Genetics in the Forensic DNA Debate*, 89 PROC. NAT'L ACAD. SCI. U.S. AM. 11654, 11656 (1992).

cent, then the crime scene DNA could have been left by anyone in the population. This means that an appropriate comparison is with a database that represents the population of people who may have committed the crime — which may be called the "suspect population."[108]

According to this argument, only in rare situations should the defendant's ethnic group be the same as that used for comparison purposes. In those few cases where it is appropriate to do so it will be because of the composition of the suspect population, not the defendant's ethnic background. Thus, if the crime for which Mohit was tried was committed in a community containing a large number of Iranian Shiite Muslims, comparison with an Iranian Shiite Muslim database might have been appropriate. Presumably it was not.[109]

Some commentators have responded, however, by suggesting that the "defendant's race is irrelevant" argument is too simplistic. It is always possible, they argue, that other members of the suspect population share the defendant's specific racial ancestry.[110] Once again we are dealing with uncertainty — we are uncertain about the composition of the suspect population. While critics of DNA evidence exploit this uncertainty to suggest that match probabilities cannot be reliably calculated, there is a familiar response from the other side of the DNA debate.

Proponents of DNA evidence use an argument that throws the burden of proof onto those who raise the possibility of the sus-

---

[108] *See* Richard Lempert, *The Suspect Population and DNA Identification*, 34 JURIMETRICS J. 1, 2-3 (1993) (defining "suspect population" as group of people suspected of committing particular crime).

[109] The argument that the defendant's race is irrelevant to the calculation of match probabilities is a powerful one. It can be used to argue that, in the majority of cases, fears about population substructure are irrelevant. *See* Kaye, *supra* note 80, at 137 (distinguishing between general population cases and subpopulation cases, and arguing that defendant's race only relevant in latter). Roeder uses the argument to suggest that a much cited paper by Lewontin and Hartl, which raises fears about substructure, is flawed because its authors compared allele frequencies in one subgroup with those in another, rather than with those in a mixed database. *See* Roeder, *supra* note 96, at 236 (criticizing R.C. Lewontin & Daniel L. Hartl, *Population Genetics in Forensic DNA Typing*, 254 SCIENCE 1745 (1991)).

[110] *See* Lewontin, *supra* note 104, at 260 (discussing match probabilities within large population and subpopulations); Richard A. Nichols & David J. Balding, *Effects of Population Structure on DNA Fingerprint Analysis in Forensic Science*, 66 HEREDITY 297, 298 (1991) (same); Thompson, *supra* note 80, at 83-84 (same); Mark Webster, *DNA Profiling Evidence*, 142 NEW L.J. 1712, 1712 (1992) (same).

pect and culprit sharing racial ancestry. In the literature, we find assertions such as, "[i]n the absence of evidence to the contrary, the suspect and culprit should be assumed to be randomly drawn from a forensic population."[111] Similarly, that the defense counsel arguing that the defendant's race is relevant, "must explain why both [suspect and culprit] are credibly drawn from the same suspect population unless the suspect is the culprit."[112]

It can be argued that these are not particularly satisfying responses to the point made by the critics. If the composition of the suspect population is uncertain, as it almost always is, the possibility remains that other members of the defendant's racial subgroup are members of the population. This will likely effect the weight of the DNA evidence. While uncertainty about the suspect population would not be a good reason for ruling DNA evidence inadmissible, the defendant is free to raise the possibility of there being other members of her subgroup in the population and bears no burden of proof in respect of such issues.[113] It is for the jury to decide whether or not they find such explanations of the DNA evidence credible; it is not for scientists to assert that they are incredible.[114]

---

[111] N.E. Morton et al., *Kinship Bioassay on Hypervariable Loci in Blacks and Caucasians*, 90 PROC. NAT'L ACAD. SCI. U.S. AM. 1892, 1895 (1993).

[112] Newton E. Morton, *Genetic Structure of Forensic Populations*, 89 PROC. NAT'L ACAD. SCI. U.S. AM. 2556, 2559 (1992); *see also* D.A. Berry et al., *Statistical Inference in Crime Investigations Using Deoxyribonucleic Acid Profiling*, 41 APPLIED STAT. 499, 529 (1992) (asserting that, absent criminological research showing "tendency, when a wrong man is arrested, for him to belong to a distinct [racial] subpopulation in common with the true offender," different subpopulations should be assumed).

[113] *See* BERNARD ROBERTSON & G.A. VIGNAUX, INTERPRETING EVIDENCE: EVALUATING FORENSIC SCIENCE IN THE COURTROOM 44-46 (1995) (discussing use of hypothetical questions at trial). I pursue some of the implications of this in Mike Redmayne, *Science, Evidence and Logic*, 59 MOD. L. REV. 747, 757-60 (1996).

[114] Some critics go further, and suggest that the prosecution should incorporate the possibility that suspect and culprit come from the same subpopulation, or are closely related, in the match probability. *See* Balding et al., *supra* note 98, at 249; David J. Balding & Peter Donnelly, *Inferring Identity from DNA Profile Evidence*, 92 PROC. NAT'L ACAD. SCI. U.S. AM. 11741, 11744 (1995).

4. Attempts to Resolve the Controversy

The controversy over the use of DNA evidence caused obvious problems for the courts. Successive pre-trial hearings, during which the reliability of scientific evidence is debated, impose costs on the legal system. Whether courts choose to admit or exclude such evidence, the controversy surrounding it will undermine faith in the legal system's ability to perform its fact-finding function.

It was important to try to resolve the issues and, in 1990, a Committee of the United States National Research Council (the "Committee" or NRC1) was appointed to examine the use of DNA technology in forensic science.[115] The Committee, which included both lawyers and scientists, dealt with a number of issues. The Committee's most important recommendations addressed the controversial question of how match probabilities should be calculated. NRC1 noted the existence of a controversy over the issue of population substructure, but did not specifically endorse either side in the debate. It simply "assume[d] for the sake of discussion that population substructure may exist."[116]

The Committee proposed a method of calculating match probabilities that would take substructure into account. This method, called the "ceiling principle," was intended to produce indisputably conservative match probabilities.[117] NRC1 also endorsed an empirical response to the determination of allele frequencies. It recommended that a number of genetically homogenous populations should be examined so the range of allele frequency variation would be known with more certainty.[118]

One might have thought that a report by a prestigious committee would have ended the controversy over the use of DNA evidence in the courts. NRC1 was able to examine the issues in

---

[115] *See* COMMITTEE ON DNA TECHNOLOGY IN FORENSIC SCIENCE, NATIONAL RESEARCH COUNCIL, DNA TECHNOLOGY IN FORENSIC SCIENCE 1 (1992) [hereinafter NRC1].

[116] *Id.* at 80.

[117] "[T]he NRC committee sought to define . . . a standard of practice so conservative as to ensure that there would be no serious scientific argument that the evidence could be said to *overstate* the case against a defendant." Eric S. Lander, *DNA Fingerprinting: The NRC Report*, 260 SCIENCE 1221 (1993) (emphasis in original). Lander was a member of the committee.

[118] *See* NRC1, *supra* note 115, at 84.

a non-adversarial context, away from the heated atmosphere of the courtroom. Those critical of the adversarial system's ability to resolve disputed scientific questions have proposed just this sort of procedure as an ideal solution.[119] However, if NRC1 is judged in terms of its success in ending the controversy and in easing the reception of DNA evidence by the courts, it was a dismal failure.

The report merely succeeded in "fanning the flames"[120] of the controversy, leading many courts to rule DNA evidence inadmissible.[121] The report was criticized in strong terms by some scientists, who argued that the ceiling principle "lack[s] any logical basis." One critic accused the Committee of "neglecting established genetic principle, misleading the courts and disgracing the National Academy."[122]

The reasons why NRC1 was so controversial tell us much about the difficulties of achieving closure in scientific disputes. The controversy that engulfed the report is a reminder, if one is needed, of the difficulty in resolving such disputes. Just as the investigation resulting from the May inquiry failed to provide any definite answer as to why positive TLC results were obtained in 1974, so NRC1 was unable to give any agreed-upon answer to the question "how reliable is DNA evidence?"

Further, NRC1 was examining a truly trans-scientific issue, one that combined scientific and policy issues and in which answers to the scientific questions were underdetermined by the available evidence.[123] Unfortunately, just as the report was released, a

---

[119] *See, e.g.*, PHANTOM RISK: SCIENTIFIC INFERENCE AND THE LAW 438-42 (Kenneth R. Foster et al., eds., 1993) (arguing for expert "consensus groups" to decide questions about controversial scientific evidence); Peter Alldridge, *Recognizing Novel Scientific Techniques: DNA as a Test Case*, 1992 CRIM. L. REV. 687, 694-95 (arguing for extra-judicial committee to examine the reliability of novel scientific techniques).

[120] *See* Roeder, *supra* note 96, at 223.

[121] *See* Peter Aldhouse, *Geneticists Attack: NRC Report as Scientifically Flawed*, 259 SCIENCE 755, 755 (1993) (discussing effects of NRC report on admission of DNA evidence); Christopher Anderson, *Courts Reject DNA Fingerprinting, Citing Controversy After NAS Report*, 359 NATURE 349, 349 (1992) (stating "it's clear that message hasn't gotten through").

[122] Newton E. Morton, *Alternative Approaches to Population Structure*, 96 GENETICA 139, 142 (1995); *see also* Aldhous, *supra* note 121, at 755 (discussing claims that NRC report is scientifically flawed); B. Devlin et. al., *Statistical Evaluation of DNA Fingerprinting: A Critique of the NRC's Report*, 259 SCIENCE 748, 749 (1993) (noting that critics have argued that parts of report lack scientific justification).

[123] *See* Kenneth R. Kreiling, *DNA Technology in Forensic Science*, 33 JURIMETRICS J. 449, 486

number of studies were published that appeared to tip the balance in the scientific debate about substructure.[124] These studies suggested that substructure was not a significant problem. This ensured that the ceiling principle, based on a presumption of substructure, would be perceived with skepticism in the scientific community. In these circumstances, the Committee had little hope of achieving sound argument closure. Even closure by consensus proved to be impossible within the wider scientific community. It is the trans-scientific status of NRC1 that appears to have rendered the report especially controversial. Because the report spoke to both scientific and policy concerns, it was open to attack by scientists on the ground that it was not sufficiently scientific.

One lesson to be drawn from the reception of NRC1 is that it is important to make as explicit as possible the different scientific and policy concerns that motivate the adoption of solutions such as the ceiling principle. NRC1's failure to do this meant that its proposals were open to misinterpretation. This can be seen in the way that some courts responded to the ceiling principle. Several did not realize that the ceiling principle was based on a response to uncertainty (the presumption of substructure) and policy (the need to be conservative). Both of these were open to reassessment if, for example, new research suggested that substructure was not significant or if a lesser degree of conservatism would suffice.[125]

Owing to the failure of NRC1 to resolve the DNA controversy, in 1993 a second committee was convened under the auspices of

---

(1993) (noting that questions addressed by committees such as NRC1 "involve a political dimension which is not subject to resolution by science alone. If the answers attempt to force a consensus or, because of uncertainty or other reasons, do not provide clear solutions, the recommendations may generate more controversy than they eliminate.").

[124] *See* Lempert, *supra* note 95, at 43 (discussing release of other reports and flaws in testing models).

[125] Particularly striking in respect of the latter is the opinion of one civil court that questioned a probability of paternity because it was its not being calculated in accordance with the ceiling principle. *See* Franson v. Micelli, 645 N.E.2d 404, 410 (Ill. App. Ct. 1994) (rejecting certain DNA evidence as not yet generally accepted under *Frye* test), *vacated for lack of juris.*, 666 N.E.2d 1188 (Ill. 1996). On courts' misreadings of NRC1, see generally David H. Kaye, *The Forensic Debut of the National Research Council's DNA Report: Population Structure, Ceiling Frequencies and the Need for Numbers*, 34 JURIMETRICS J. 369, 375-82 (1994) (discussing use of NRC report in judicial opinions).

the National Research Council.[126] To date, the second report (NRC2) appears to have had a positive reception, at least so far as its recommendations on match probability calculation are concerned.[127] One of the strengths of NRC2 is its generally candid approach to the trans-scientific issues that have made the DNA controversy so difficult to resolve.

NRC2 accepts that our knowledge of a number of factors — such as the degree of substructure in human populations, allele frequencies in certain ethnic subgroups, and laboratory error rates — is limited. Indeed, in the report the word "uncertainty" appears almost as often as the letters "DNA." Acknowledging uncertainty, NRC2 proposes measures that account for uncertainty and is explicit about the policy factors that shape its response to uncertainty. Granting that its recommendations are structured by the prosecution's burden to prove its case beyond reasonable doubt, the report notes that its recommendations should not necessarily be applied in civil cases.[128] In general, NRC2 takes a conservative approach when faced with uncertainty, though it recognizes that this approach is not necessarily scientific.[129]

Throughout most of the report, care is taken to mark a distinction between science and policy. Rather than compromising its legitimacy by proposing answers to policy questions, the report frequently leaves them to be decided by the courts. This includes acknowledging that some courts may wish to continue using the ceiling principle.[130] For example, although NRC2

---

[126] *See* COMMITTEE ON DNA FORENSIC SCIENCE: AN UPDATE, NATIONAL RESEARCH COUNCIL, THE EVALUATION OF FORENSIC DNA EVIDENCE 1-2 (1996) [hereinafter NRC2] (explaining that lawyers and scientists on both sides of DNA debate criticized 1992 report, and courts misinterpreted or misapplied some statements in report). The main recommendations are reviewed in Bruce S. Weir, *The Second National Research Council Report on Forensic DNA Evidence*, 59 AM. J. HUM. GENETICS 497, 498 (1996).

[127] *See* Weir, *supra* note 126, at 497 (commenting that most important finding of 1996 report would likely be predicted time when each person could be uniquely identified (except for identical twins)).

[128] *See* NRC2, *supra* note 126, at 53-54 (noting that 1996 NRC report recommendations, like 1992 report, were motivated by legal requirement of proof beyond reasonable doubt in criminal trials).

[129] The report notes that "[s]tatistically accurate estimates . . . can yield results that overvalue the weight of evidence against the defendant, even though on average they produce values that are closer to the true frequency than those produced by conservative estimates." *Id.* at 52-53.

[130] *See id.* at 190-91 (stating that jurisdictions admitting scientific evidence on basis of

looks forward to the day when uncertainty has been so reduced that forensic scientists can declare a DNA profile to be unique, it realizes that whether a profile is uncommon enough to be called unique is a question of legal policy.[131] Similarly, the question of how match probabilities are best presented to fact-finders is left to the courts. However, NRC2 does make the important recommendation that empirical research be carried out to evaluate juror comprehension of probabilistic evidence.[132]

Occasionally NRC2's delicate balancing act between science and policy is less successful. To return to the question of whether the defendant's ethnic subgroup is relevant to match probability calculation,[133] NRC2 notes that some commentators argue that the suspect's subgroup should always be taken into account.[134] This suggestion is rejected because it would be "unnecessarily conservative."[135] Given that NRC2 is rejecting a rule that would favor defendants, a little more explanation of the Committee's reasoning and of its approach to conservatism would be appropriate.

Another criticism concerns the way NRC2 deals with the problem of laboratory error rates. Currently, the argument that the testing laboratory might have made an error appears to be the most fruitful line of attack in DNA cases. Some commentators have gone so far as to suggest that, because the probability that the laboratory declared a match owing to an error will usually dwarf the random match probability, the latter is largely irrelevant.[136] NRC2 discusses the problem of error rates at some length. While making no specific recommendation on how error

---

source-methodology standard have usually fared well using ceiling estimate).

[131] *See id.* at 137 (stating that NRC2 leaves courts to determine minimally acceptable probability to establish profile uniqueness).

[132] *See id.* at 203-04 (recommending further research as to how to reduce misinterpretation and misunderstanding of DNA evidence).

[133] *See supra* notes 106-14 and accompanying text (discussing controversy over race and ethnicity in DNA testing).

[134] *See* NRC2, *supra* note 126, at 114 (citing David J. Balding & Richard A. Nichols, *DNA Profile Match Probability Calculations: How to Allow for Population Stratification, Relatedness, Database Selection and Single Bands*, 64 FORENSIC SCI. INT'L 125 (1994)).

[135] *See id.* (arguing that subgroups are irrelevant).

[136] *See* Koehler et al., *supra* note 89, at 201 (concluding that random match probabilities contribute little when they are several orders of magniutude smaller than error rate).

rates should be dealt with in court, it rejects the most radical proposal — that the error rate be combined with the match probability — on the grounds that it is "inappropriate."[137]

In spite of its willingness to discuss error rates, NRC2 refuses to endorse NRC1's recommendation that error rates be disclosed to juries, calling this a policy issue.[138] It is difficult to reconcile these two parts of the report. Perhaps this indicates the science/policy distinction is less clear than the Committee would have us believe.

C. *The Bendectin Litigation*

The third example of scientific disagreement is the debate over whether the morning sickness drug Bendectin was[139] a teratogen, that is, whether it caused birth defects in children. The debate over Bendectin was different from the two examples of scientific disagreement previously discussed. The twenty-seven lawsuits brought against Merrell Dow, Bendectin's manufacturer, obviously made this a more wide-ranging dispute than that over the TLC results in *Maguire*. This suggests that the Bendectin dispute is closer to the DNA controversy. Yet, whereas the forensic use of DNA evidence caused controversy in the scientific community, disagreement over Bendectin's teratogenicity could hardly be called a controversy. While some scientific studies supported the allegation that Bendectin was harmful to developing fetuses, these studies were always a minority. Further, their authors usually expressed caution about their results.[140]

By the mid-1980s, after there was a considerable body of research on Bendectin, only in the courts was there any real disagreement between scientists over whether Bendectin was a teratogen.[141] This is one of the most intriguing aspects of the

---

[137] *See* NRC2, *supra* note 126, at 87 (stating that possible false matches can be directly tested).

[138] *See id.* at 185 (refusing to pass social policy judgment because purpose of report is to recommend procedures for DNA test validation).

[139] Although its ingredients are still readily available, Bendectin was withdrawn from the market in 1983. *See generally* MICHAEL D. GREEN, BENDECTIN AND BIRTH DEFECTS 180-88 (1996) (describing withdrawal of Bendectin from market).

[140] *See* Joseph Sanders, *From Science to Evidence: The Testimony on Causation in the Bendectin Cases*, 46 STAN. L. REV. 1, 25-27 (1993) (reviewing studies supporting teratogenicity of Bendectin).

[141] *See id.* at 27 (stating that substantial scientific findings disputing Bendectin's birth

Bendectin cases: while the scientific community was able to form a consensus on Bendectin relatively easily, the issue of whether Bendectin caused birth defects continued to be litigated. Around forty percent of juries returned verdicts for plaintiffs, although very few of these verdicts withstood the appeals process.[142]

The world of law seemed to be very much out of step with the world of science. Some commentators have pointed to the apparent distortion of science in these cases as the basis for strong criticism of the way the legal system handles scientific evidence.[143] However, a closer look at the Bendectin litigation reveals that things were not quite this simple.[144]

Despite the differences between Bendectin and the two other examples examined in this Article, there are also similarities among all three. The most obvious is that, once again, courts and scientists were faced with uncertainty. Even today, when Bendectin is one of the best studied drugs in history, a conclusion about its teratogenicity must be expressed in cautious terms. While it is very unlikely that Bendectin is a powerful cause of birth defects, it is possible that the drug is a mild teratogen.[145]

In the early years of Bendectin litigation, the issues were much less clear. The evidence was consistent with a two- to five-fold increase in birth defects.[146] This left room for

---

defect effects did not support verdict for plaintiff).

[142] *See id.* at 10-12 (stating that courts have struck down all five plaintiffs' federal jury verdicts and explaining that courts have dismissed other cases through summary judgment or directed verdict).

[143] *See, e.g.*, PETER J. HUBER, GALILEO'S REVENGE: JUNK SCIENCE IN THE COURTROOM 175-78 (1991) (criticizing courts for tolerating trivial or marginal scientific evidence).

[144] *See* GREEN, *supra* note 139, at 328-51 (exploring lessons and "non-lessons" of Bendectin litigation).

[145] *See* Louis Lasagna & Sheila R. Shulman, *Bendectin and the Language of Causation*, *in* PHANTOM RISK: SCIENTIFIC INFERENCE AND THE LAW 101, 109 (Kenneth R. Foster et al. eds., 1993).

> In summary, the scientific evidence seems sufficient to rule out the possibility that Bendectin is a powerful cause of birth defects. The possibility that it might cause undetectably small increases in the rate of birth defects cannot be ruled out by scientific data. Proving that Bendectin does not cause birth defects is logically impossible.

Id.

[146] *See* GREEN, *supra* note 139, at 328-32 (describing scientific knowledge and studies available to Bendectin litigants in mid-1970s).

disagreement among scientists over whether Bendectin had caused a particular plaintiff's birth defects. The question of whether Bendectin caused birth defects was, and to an extent still is, trans-scientific. Low level teratogenicity simply cannot be satisfactorily proved given available scientific resources.[147]

It is not surprising that, given the residual uncertainty, scientists disagreed as to whether Bendectin was a teratogen. Indeed, we find those involved in the debate over Bendectin employing some familiar arguments. Although the plaintiffs in a civil case bear the burden of proof, during the litigation lawyers and experts sometimes used the argument that, owing to the risk posed by a drug which might be harmful, the defendants should bear the burden of proving Bendectin's safety.[148] Indeed, an analysis of the Bendectin cases suggests that, when juries returned plaintiffs' verdicts, they may have been indulging in similar reasoning. Swayed by Merrell's lack of caution in marketing Bendectin, these juries may have been more prepared to find

---

[147] *Cf.* Weinberg, *supra* note 19, at 210 (discussing problems that are "trans-scientific" questions which cannot be answered by science). An example deployed by Weinberg is instructive: the genetic effects of low level radiation on mice. *See id.* A scientist wanting to show, at the 95% confidence level, that a dose of 150 millirems of radiation caused a 1/2% increase in the mutation rate in mice would need about 8 billion mice. *See id.* "[T]he number is so staggeringly large that, as a practical matter, the question is unanswerable by direct scientific investigation." *Id.* *See also* Wendy E. Wagner, *Trans-Science in Torts*, 96 YALE L.J. 428, 431-36 (1986) (discussing trans-science in toxic torts).

[148] Two examples are given in GREEN, *supra* note 139, at 232, 283-84. In one case the plaintiffs' attorney made the following argument during his summary:

> [Merrell] would say that it's not in our system of justice, say that it's not their burden to prove it's safe. But I ask you, shouldn't that be the starting point, that any drug given to a pregnant woman during the period of organogenesis has to be looked at with a jaundiced eye to be sure that it does not harm this helpless child at the most vulnerable part of its life?

Id. at 232. In another case, Shanna Swan, one of the plaintiff's expert witnesses, used a similar strategy. Green recounts that:

> Swan explained that while a given epidemiological study may not have found a statistically significant association, it should not be understood as proving the safety of Bendectin. . . . Taking on her role as a public health official, she explained that the upper bound of the confidence interval is the one of primary concern: "[I]f you want to protect the public health and you want to know how bad can the situation be and still be calling something safe and I think to do that you need to look at the upper confidence levels . . . ."

Id. at 283-84.

causation proved.[149] For their part, the defendants repeatedly used arguments that effectively placed too high a burden of proof on the plaintiffs.[150]

This use of burdens of proof in arguments about risk suggests a degree of similarity between arguments about both DNA and Bendectin. One contrast between the two disputes lies in the way the legal system's formal burdens of proof structured the debates. In the DNA controversy, critics of DNA evidence argued that the prosecution, typically the adducer of DNA evidence, must prove its case beyond reasonable doubt. This added plausibility to their demand that DNA's proponents prove that substructure in human populations would not undermine match probability calculation.

With Bendectin, however, the plaintiffs bore the burden of proving, on the balance of probabilities, that Bendectin caused birth defects. This meant that they could not capitalize on the strategy of skepticism used by critics of DNA evidence. Nor could appeals to the need for conservatism in the face of uncertainty play a significant role in the dispute. During the Bendectin litigation it became apparent that, as far as the scientific evidence was concerned, plaintiffs simply would not be able to meet their burden of proof.[151] Therefore, judges became increasingly interventionist in dealing with Bendectin cases.[152]

---

[149] *See* Sanders, *supra* note 140, at 52-54, 72-77 (noting that plaintiffs' strategy in many Bendectin cases was to commingle testimony on negligence and causation, which may have led juries to find against Merrell). Sanders suggests that bifurcation of trials (to separate causation and negligence issues) might lead to more accurate verdicts. *See id.* at 72-77.

[150] *See, e.g.*, GREEN, *supra* note 139, at 33. Green observed that "Merrell consistently insisted, before both juries and judges, that unless an epidemiological study found an association that was statistically significant, it could not serve as proof of causation." *Id.* For a discussion of why this approach is problematic, see Michael D. Green, *Expert Witnesses and Sufficiency of Evidence in Toxic Substances Litigation: The Legacy of Agent Orange and Bendectin Litigation*, 86 NW. U. L. REV. 643, 674-95 (1991) (discussing dangers of using epidemiologic studies to find toxic causation); *see also infra* note 175 and accompanying text (discussing *Brock v. Merrell Dow Pharms., Inc.*, 874 F.2d 307 (5th Cir.), *modified*, 884 F.2d 166 (5th Cir. 1989)).

[151] *See* Sanders, *supra* note 140, at 3 ("from a legal point of view, based on [the] scientific evidence, a plaintiff simply cannot show by a preponderance of the evidence that Bendectin caused her birth defects"); *see also* GREEN, *supra* note 139, at 328 ("it seems reasonably clear that no plaintiff should be able to satisfy the burden of proof on causation in a Bendectin case").

[152] *See generally* Joseph Sanders, *The Bendectin Litigation: A Case Study in the Life Cycle of*

Given the difficulty faced by plaintiffs, why were a substantial number able to persuade juries of the merits of their cases? One lesson of the Bendectin litigation, which has wider relevance for courts dealing with disputed areas of science, is the difficulty courts face in trying to establish the probability of a scientific proposition. This is because of the many ways in which the legal system distorts science.

In a detailed review of the Bendectin litigation, Joseph Sanders has charted the ways in which this can happen. One of his most interesting observations is that, even before science becomes evidence in the legal system, it can be molded by legal concerns.[153] Research on Bendectin was driven by legal concerns.

As the Bendectin litigation gathered momentum, Bendectin research became a "hot topic" and resources were made available for its study.[154] However, certain sorts of research, such as animal studies, were less popular. Partly because Merrell saw no advantage in such studies, resources were not available for them.[155] Articles on Bendectin probably became easier to publish, possibly even those with marginal results. Because Merrell funded much of this research, allegations were made that this undermined any results that demonstrated the safety of

---

Mass Torts, 43 HASTINGS L.J. 301, 362-84 (1992) (discussing life cycle of Bendectin cases).

[153] *See id.* at 331 ("We should anticipate that the science itself is influenced by the legal process. As the congregation of cases grows and matures, it creates its own gravity field, attracting and distorting the science that comes near it."). Similar points can be made about the DNA controversy. The Justice Department was able to award a grant of $200,000 for research intended to generate a series of peer-reviewed articles that would provide better support for the FBI's statistical methods. *See* Christopher Anderson, *Coincidence or Conspiracy?*, 355 NATURE 753 (1992) (presenting allegations of federal law enforcement interference with peer review process of paper critical of FBI statistical analyses); *see also infra* notes 180-85 (discussing adversarialism due to DNA controversy).

[154] *See* Sanders, *supra* note 140, at 346 (graphing increase in Bendectin studies).

[155] *See id.* 336-37 (giving reasons why studies presented no advantage to Merrell). In the long term, toxic tort litigation may have lessened the demand for animal studies of possibly harmful substances. Courts involved in the Bendectin and Agent Orange litigation have tended to privilege epidemiological studies over animal studies, and this may mean that the latter come to be seen as less important by the scientific community. *See* Michael D. Green, *Expert Witnesses and Sufficiency of Evidence in Toxic Substances Litigation: The Legacy of Agent Orange and Bendectin Litigation*, 86 NW. U. L. REV. 643, 668 (1992) (noting preference for epidemiological data over animal studies in Bendectin cases); Joseph Sanders, *Scientific Validity, Admissibility, and Mass Torts after* Daubert, 78 MINN. L. REV. 1387, 1407-16 (finding courts not allowing animal studies where epidemiological data exists).

Bendectin.[156] Further, because the Bendectin litigation had political overtones — a large pharmaceutical company accused of peddling an unsafe product in the search for profit, abetted by lax government regulation — certain scientists may have been drawn to the subject, and taken a particular view of the merits of plaintiffs' cases.[157]

Sanders's other illustrations of how the Bendectin litigation distorted science are more familiar. Once a body of research exists in the scientific community, it must be translated into evidence in the courtroom. During this process, however, it is subject to distortion.

Consider the selection of experts. The experts who appear during a trial will obviously not be a random selection from the relevant scientific community. They will be specifically chosen by each party because they support that party's case. This does not mean that those experts do not believe the propositions to which they are testifying,[158] merely that they are an unrepresentative sample.

In the Bendectin litigation, this problem was exacerbated by the fact that each side tended to present equal numbers of experts. This probably had a levelling effect in the eyes of the jury. To them it would seem that the scientific community was evenly divided on the question of Bendectin's teratogenicity. In fact, the weight of scientific opinion favored Merrell, but the company was not easily able to convey this to the jury. Nor was it easy to communicate to the jury that epidemiological evidence carries more weight in the scientific community than other evidence, such as animal studies.[159]

Another factor that perhaps had a levelling effect on expert testimony in the Bendectin cases was cross-examination. Cross-examination is an extraordinarily powerful tool for bringing out

---

[156] *See* Sanders, *supra* note 140, at 37-38.

[157] *See* Peter H. Schuck, *Multi-Culturalism Redux: Science, Law, and Politics*, 11 YALE LAW & POL'Y REV. 1, 9 n.27 (1993) (noting that intersection of science and law "caused scientists to become politicized on this issue").

[158] *See* Sanders, *supra* note 140, at 37 (explaining that litigants are unlikely to choose experts who must lie and prefer experts who give helpful testimony that they believe).

[159] *See id.* at 47. "If the legal process tends to cause all experts to appear equally qualified, it also causes all science to appear equally worthy. The problem confronted by the jury is basically the same, a problem of *weighing*." *Id.* (emphasis in original).

the weaknesses in expert testimony.[160] Cross-examination can trade on the uncertainty inherent in any scientific proposition, making an accepted scientific fact appear completely tenuous. Used against Merrell's experts, it threw their testimony into doubt.[161]

Sanders also suggests that the juries' decision-making process was poorly suited to assessing the strength of Merrell's case. If jurors approached their decision-making task in terms of story comparison,[162] Merrell's case would have seemed less compelling than in fact it was. In addition, the probabilistic epidemiological evidence, which favored Merrell, was not only difficult to understand[163] but also fit poorly into story-based decisionmaking.[164]

The difficulty of translating scientific consensus into a strong case in the courtroom left the legal system in an embarrassing position. Juries in different parts of the country were returning inconsistent verdicts on what was essentially the same issue: whether or not Bendectin was a teratogen. Moreover, those verdicts that found for plaintiffs were at odds with the scientific consensus.[165] With so many different cases being litigated, the

---

[160] On the deconstructive potential of the cross-examination of expert witnesses, see SHEILA JASANOFF, SCIENCE AT THE BAR: LAW, SCIENCE, AND TECHNOLOGY IN AMERICA 53-55 (1995); J.S. Oteri et al., *Cross Examination of Chemists in Drugs Cases*, *in* SCIENCE IN CONTEXT, *supra* note 15, at 250, 251-52; Brian Wynne, *Establishing the Rules of Laws: Constructing Expert Authority*, *in* EXPERT EVIDENCE: INTERPRETING SCIENCE IN THE LAW 23, 32-39 (Roger Smith & Brian Wynne eds., 1989).

[161] *See* Sanders, *supra* note 140, at 47-51 (describing how cross-examination in Bendectin litigation undermined credibility of expert witnesses).

[162] A number of commentators suggest this is how fact-finders decide cases. *See, e.g.*, Nancy Pennington & Reid Hastie, *The Story Model for Juror Decision Making*, *in* INSIDE THE JUROR: THE PSYCHOLOGY OF JUROR DECISION MAKING 192, 192-203 (Reid Hastie ed., 1993) (discussing story model for fact finder decisionmaking); WILLEM A. WAGENAAR ET AL., ANCHORED NARRATIVES: THE PSYCHOLOGY OF CRIMINAL EVIDENCE 33-43 (1993) (discussing anchored narratives theory for fact finder decisionmaking).

[163] *See* Sanders, *supra* note 140, at 45-46 (stating that complexity of information and problems with using demonstrative props make it difficult for lawyers to keep jury's attention).

[164] *See id.* at 58-60 (explaining that, since jurors choose most persuasive story as basis for decision, statistically-based causation argument is less effective).

[165] *See id.* at 82-86 (discussing actions by judges removing issues from juries).

legal system could impose no neat closure on the Bendectin issue and the cases then became a target for critics of the tort system.[166]

In stark contrast to the jury deference seen in the English Court of Appeal's reaction to scientific evidence,[167] American judges took an interventionist approach to pro-plaintiff Bendectin verdicts. Some judges granted summary judgment[168] or judgments for Merrell notwithstanding the verdict.[169] Other plaintiff's verdicts were overturned on appeal.[170] It seems that the legal process could only be reconciled with science through implicit criticism of one of the legal system's most cherished institutions — the jury.

In the process of appellate review of jury verdicts, however, we can find another example of the legal system's difficulty in interpreting science. The courts have a tendency to simplify science, and to look for simple, easy to apply rules for dealing with scientific evidence.[171] Some of the Bendectin cases have been accused of taking this desire too far.[172]

Some courts began to erect what has been called an "epidemiological threshold," ruling that without epidemiological evidence to support their cases, plaintiffs should not prevail.[173] Further,

---

[166] *See* GREEN, *supra* note 139, at 328 (calling Bendectin cases "the single most criticized piece of large-scale litigation of all time").

[167] *See* Nobles et al., *supra* note 71, at 12 (criticizing jury's ability to decide scientific evidence).

[168] *See, e.g.*, Lynch v. Merrell-Nat'l Lab. Div. of Richardson-Merrell, Inc., 646 F. Supp. 856, 867 (D. Mass. 1986), *aff'd*, 830 F.2d 1190 (1st Cir. 1987).

[169] *See, e.g.*, Richardson v. Richardson-Merrell, Inc., 649 F. Supp. 799, 804 (D.D.C. 1986), *aff'd*, 857 F.2d 823 (D.C. Cir. 1988).

[170] *See, e.g.*, Brock v. Merrell Dow Pharms., Inc., 874 F.2d 307, 315, *modified*, 884 F.2d 166 (5th Cir. 1989).

[171] *See supra* note 42 and accompanying text (noting that legal system tends to simplify knowledge from other disciplines).

[172] *See* GREEN, *supra* note 139, at 311-20 (criticizing courts' generalization and simplification of evidentiary thresholds).

[173] *See id.* at 306-11. Epidemiology is the statistical study of correlations between phenomena such as diseases and birth defects, and the phenomena alleged to cause them. In the context of Bendectin, this would involve studying the number of children with birth defects born to mothers who took Bendectin, compared to the number of birth defect children born to mothers who did not take Bendectin. In an interesting analysis, Gaskins argues that the replacement of traditional laboratory studies by epidemiology constituted a shift in the burden of proof, making it more difficult to prove a relationship between a drug and a disease. *See* GASKINS, *supra* note 20, at 148-49.

publication of research began to be used as a seal of approval. Courts held that testimony based on unpublished studies was inadmissible.[174]

In the most criticized decision, the court in *Brock v. Merrell Dow Pharmaceuticals, Inc.* ruled that only statistically significant epidemiological studies were admissible as evidence. The court then determined that because the studies relied on by the defendant were statistically significant, they were beyond criticism.[175] The *Brock* decision incorporates a misunderstanding of science. As significance tests only speak to sampling error, they cannot overcome any of the other possible methodological flaws in a study.[176] Additionally, the decision misunderstands the relationship between statistical significance and the legal burden of proof.[177]

The high level of statistical significance usually demanded in science is designed to guard against false positives. It is too stringent a test to apply in civil litigation, where a mistaken verdict for the plaintiff is not usually seen as worse than a mistaken verdict for the defendant.[178] In addition, the tendency of some courts to demand that plaintiffs furnish epidemiological studies demonstrating a relative risk of at least two is also an overly simplistic approach to scientific evidence.[179] Although such bright-line tests will always prove attractive, they are no substitute for careful and reasoned review of the evidence.

---

[174] *See* GREEN, *supra* note 139, at 304. Compare the more cautious approach to publication as a guarantee of good science in the Supreme Court's decision in *Daubert v. Merrell Dow Pharmaceuticals, Inc.*, 509 U.S. 579, 584-85 (1993).

[175] *See Brock*, 874 F.2d at 312 (discussing need for epidemiological study to be statistically significant in order to be admissible as evidence).

[176] *See* GREEN, *supra* note 139, at 317.

[177] *See id.* at 318; David H. Kaye, *Is Proof of Statistical Significance Significant?*, 61 WASH. L. REV. 1333, 1346-47 (1986) (noting that triers of fact unfamiliar with statistical terms may interpret studies with no significant differences as having no real differences); David H. Kaye, *Statistical Significance and the Burden of Persuasion*, L. & CONTEMP. PROBS., Autumn 1983, at 13-14 (1983) [hereinafter Kaye, *Statistical Significance*].

[178] *See, e.g.*, Kaye, *Statistical Significance*, *supra* note 177, at 16 (citing Speiser v. Randall, 357 U.S. 513, 525-26 (1958)).

[179] *See* Mark Parascandola, *Evidence and Association: Epistemic Confusion in Toxic Tort Law*, 63 PHIL. SCI. 168 (1996) (discussing two different uses of strength employed in deciding causation in toxic torts cases).

III. Implications of the Case Studies

What lessons can be learned from these examples of scientific disagreement? Disagreement among witnesses giving evidence in court is nothing new for the law: it is a staple of litigation. Usually, such disagreement will be left for the fact-finder to sort out. But the scientific disagreements examined here are more problematic.

The disagreement among scientists has gone beyond individual court cases, whether through examination by government departments and investigatory committees, as in *Maguire*; spilling into the scientific community, as in the DNA controversy; or becoming the subject of repeat litigation, as in the Bendectin cases. These disagreements about scientific issues pose wider problems for the law because the legal system cannot easily resolve them through its normal procedures. Further, because authoritative closure of scientific disputes can only be performed by the scientific community, scientific disagreement threatens to undermine the legal system's authority in fact-finding. This is especially so where courts come to decisions which differ from those reached by the scientific community.

A. *Science, Evidence, and Adversarialism*

In the Bendectin and DNA controversies, there is little doubt that scientific disagreement was intensified by the interaction between science and the legal system. For example, scientists' discussions of DNA evidence sometimes showed an acute awareness of the problems that criticisms of DNA technology might cause in the courts.[180] Because legal policy introduces such

[180] For instance, one geneticist commented in the following terms on an article criticizing DNA evidence: "I felt publishing the article would create a very serious problem in the legal system, and that that was [the authors'] intent." Leslie Roberts, *Was Science Fair to its Authors?*, 254 SCIENCE 1722, 1722 (1991) (quoting Kenneth Kidd). Thomas Caskey was also concerned with this article: "'[P]ublishing defence testimony in a scientific journal' gives it such weight that courts might reopen, perhaps to overturn convictions obtained on the basis of DNA evidence." Christopher Anderson, *DNA Fingerprinting Discord*, 354 NATURE 500, 500 (1991) (quoting Thomas Caskey). Another example of the extent to which science and law had become intertwined is the FBI's reaction to leaked drafts of NRC1, which led the agency to pressure the committee to change its recommendations on match probability calculation. *See* Leslie Roberts, *DNA Fingerprinting: Academy Reports*, 256 SCIENCE 300, 301 (1992) (suggesting that FBI concerns regarding recommendations were relaxed

non-epistemic factors into scientific debate, scientific consensus becomes harder to achieve. Courts also give scientists a ready public venue in which to vent opposing views, magnified because the adversarial system often has a polarizing effect, making it harder for the parties, or their experts, to reach a compromise.

In the DNA controversy, adversarialism went beyond the courtroom and characterized exchanges as scientists published chains of rebuttals and counter-rebuttals in scientific journals. In the Bendectin litigation, the legal system succeeded in fomenting a scientific dispute that had little value outside the courtroom. There is some irony in this, because the legal system, even as it searches for some consensus on the reliability of scientific evidence, plays a role in destroying the very consensus which it seeks.[181]

Protracted scientific disagreement imposes costs on the legal system. The DNA controversy meant that some courts denied themselves probative evidence, while the Bendectin dispute required the processing of a number of largely unmeritorious cases.[182] But at the same time, scientific controversies have their value. As Mazur notes, "the proper function of a controversy is the identification and evaluation of potential problems, as an informal method of technology assessment."[183]

---

after last-minute revisions).

[181] Although writing about a very different subject, Hofstadter provides a good metaphor for the process by which the legal process destroys the consensus it seeks:

> A good friend is visiting from far away and before she returns home, you want to capture her infectious smile on film. But she is terribly camera-shy. The moment you bring out your camera, she freezes: spontaneity is lost, and there is no way to record that smile. The act of trying to capture this elusive phenomenon completely destroys the phenomenon.

DOUGLAS R. HOFSTADTER, *Heisenberg's Uncertainty Principle and the Many-Worlds Interpretation of Quantum Mechanics*, *in* METAMAGICAL THEMAS 455, 455 (1985). A good example of this sort of circularity in the legal system's handling of scientific evidence is the way that, in jurisdictions that look for some degree of general acceptance before admitting scientific evidence, decisions to exclude evidence may increase the likelihood that such evidence will be excluded in the future. For example, in DNA cases, when courts ruled DNA evidence inadmissible, this would often lead to indignation among some members of the scientific community. Reading such indignation as further controversy, a *Frye* court might then be even more likely to exclude DNA evidence.

[182] *Cf.* MAZUR, *supra* note 30, at 130 ("Controversies bring their share of problems. In delaying the implementation of a technology, they may deny to society important benefits, at least for awhile.").

[183] *Id.* at 129. Mazur also argues that "[t]here have been numerous instances when the

In just this way, the adversarialisation of the dispute about DNA evidence has encouraged intense scrutiny of every aspect of DNA profiling and prompted research that might not otherwise have been undertaken.[184] Similarly, the very reason why we can now say, with a large degree of confidence, that Bendectin is not a significant teratogen, is that the litigation spurred a determined research program to study the drug. In contrast, one might speculate that had the use of the TLC tests to secure the convictions of the Maguire family sparked more controversy, it might sooner have been known that substances such as shoe and floor polish could produce a positive TLC result.[185]

B. *Resolving Scientific Disputes*

If one lesson from these case studies is that law tends to exacerbate scientific disagreement, what can be learned about attempts to establish consensus? It is evident that scientists often find it hard to reach complete agreement about scientific propositions. Even if the scientific dispute is referred to a body outside the legal system, this by no means guarantees that agreement will be reached. Even if agreement is reached, the consensus will be open to attack by outsiders.[186] In fact, one of the more pessimistic conclusions to be drawn from the case studies is that an attempt at consensus building may even increase the degree of controversy surrounding scientific evidence.[187]

---

informal process of social controversy has been more effective in identifying and explicating the risks and benefits of a technology than have been any of the formal means which are supposed to do this." *Id.* at 127.

[184] *See* Richard Lempert, *Comment: Theory and Practice in DNA Fingerprinting*, 9 STAT. SCI. 255, 258 (1994) (observing that "in this instance the importation of legal adversariness into the scientific world has spurred both valuable research and practical improvements in the way DNA evidence is analyzed and presented").

[185] This last fact was only revealed after the Maguires' successful appeal. It came to light during the hearing of the *Ward* case, another terrorist case in which TLC had been used to detect nitroglycerine. *See* R. v. Ward, [1993] 1 W.L.R. 619; *see also* Heather Mills, *Scientist Admits Ward Evidence 'Misleading'*, INDEPENDENT, May 19, 1992, at 2 (noting that misleading evidence had been given at 1974 trial of Judith Ward).

[186] *See* GASKINS, *supra* note 20, 166-67 (arguing that body such as science court "would undermine its own authority" because its attempt to speak with finality would stifle scientific inquiry).

[187] It will also tend to increase the *perception* that there is a controversy. Kaye has shown that courts considering the admissibility of DNA evidence most often referred to NRC1 as

There are, though, more positive lessons to be learned from the case studies. One difficulty that faced NRC1 was that it was studying DNA evidence at a time when research that threw valuable light on the technique was still being published. This meant that any consensus achieved was likely to be unstable. Unfortunately, the report did not contain any clear indication of how its proposals might be modified in the light of new research. Another lesson is that consensus building is more likely to be successful if science can be separated from policy.[188] This not only makes it easier for scientists to agree with each other, but also decreases the likelihood that the recommendations of a consensus building body will be misinterpreted by either scientists or lawyers.

Another means of reducing the level of dispute about scientific evidence in the courtroom is to increase the scrutiny and regulation of scientific evidence outside the courtroom. This is perhaps most appropriate for scientific evidence used in criminal litigation. When a new criminal identification technique is used in the courts,[189] judges should be able to rely on two facts. First, that the technique has already been subjected to rigorous tests and second, that the laboratories implementing it have undergone proficiency testing.

In the United Kingdom, it is remarkable that, despite repeated recommendations,[190] no body exists to oversee the work

---

evidence of controversy. *See* Kaye, *supra* note 125, at 373-74.

[188] *See generally* MAZUR, *supra* note 30, at 34-42 (discussing greater ease in resolving disputes when people evaluate facts instead of politics).

[189] There are a number of new techniques that may be developed for forensic use in the near future. As well as new methods of carrying out DNA profiling, it will soon be possible to identify characteristics such as race from DNA samples. *See* Gail Vines, *Genes in Black and White*, NEW SCIENTIST, July 8, 1995, at 34. Other novel techniques include HIV profiling, laser mass spectroscopy, facial mapping, and offender profiling. *See* Phyllida Brown, *Lawyers Look to Genetics to Prove HIV "Guilt"*, NEW SCIENTIST, July 11, 1992, at 5 (describing increasing use of genetic studies in HIV lawsuits); Jerome Burne, *Caught in the Act? Anti-crime Surveillance Systems*, TIMES (LONDON) MAG., July 16, 1994, at 14 (discussing use of video images in profiling criminals); *Mass Murder Charge Hangs by a Hair*, NEW SCIENTIST, Aug. 7, 1993, at 18 (noting use of laser mass spectroscopy in forensic science); David C. Ormerod, *The Evidential Implications of Psychological Profiling*, 1996 CRIM. L. REV. 863.

[190] *See* VISCOUNT RUNCIMAN, THE ROYAL COMMISSION ON CRIMINAL JUSTICE: REPORT 144 (1993) (proposing establishment of Forensic Science Advisory Council); BRIAN CADDY, ASSESSMENT AND IMPLICATIONS OF CENTRIFUGE CONTAMINATION IN THE TRACE EXPLOSIVE SECTION OF THE FORENSIC EXPLOSIVES LABORATORY AT FORT HALSTEAD 42-43 (1996)

and standards of the country's forensic science agencies. There are also concerns in the United States about the lack of regulation of forensic science laboratories.[191] The existence of regulatory bodies would not ensure that novel scientific evidence is problem free. The intense scrutiny engendered by adversarial litigation will often reveal difficulties that never came to light outside the courtroom. Nevertheless, regulatory bodies can play a valuable role in scrutinizing novel scientific techniques. They can also help ensure that there is a body of expertise to be consulted when problems arise.

C. *Dealing with Uncertainty*

The theme that most clearly links the examples examined here is uncertainty, and how the participants in a dispute respond to it. Uncertainty produces disagreement. Often, scientific disagreement can usefully be analyzed through the concept of informal burdens of proof. A scientific dispute may revolve around the question of which side should bear the burden of proving its case. This was the case in the DNA controversy and, to a lesser extent, in the Bendectin dispute. The lack of protracted controversy in *Maguire* meant that proof burdens did not play a major role in arguments about the scientific evidence. However, an analysis of the case demonstrates that, had those assessing the evidence taken a different approach to uncertainty, the outcome of the case would have been very different. As Sir John May observed, had the Home Office's assessment of the scientific evidence been structured by the criminal burden of proof, the Home Secretary would likely have referred the case to the Court of Appeal much earlier.[192]

One difficulty in assessing uncertainty is that courts are often called upon to make judgments about scientific evidence before thorough research has been completed. Those assessing the

---

(calling for establishment of Inspectorate of Forensic Sciences). Another relevant recommendation has been that forensic scientists should be registered. *See generally* DAINTON, SELECT COMMITTEE ON SCIENCE AND TECHNOLOGY, FORENSIC SCIENCE: REPORT 33 (1993).

[191] *See* Randolph N. Jonakait, *Forensic Science: The Need for Regulation*, 4 HARV. J.L. & TECH. 109, 124-30 (1991).

[192] *See* May, *Second Report*, *supra* note 52, at 91.

available evidence may often feel uneasy because they fear that future research may present them with very different data. Thus, in the first Bendectin case, experts called by the plaintiffs suggested that future research would provide convincing evidence that Bendectin was a teratogen.[193] Similar arguments were used in the DNA controversy. This sort of situation is particularly problematic, because, while scientists may find it difficult to agree about what the available evidence shows, they are even more likely to disagree about what evidence uncovered in the future will show. We might call this second order uncertainty. That is uncertainty about how accurate are our present assessments of uncertainty.[194]

The legal system is, it seems, responsive to second order uncertainty. This partly explains why courts involved in the Bendectin litigation became less and less sympathetic to the claims of plaintiffs. Although there had never been convincing evidence that Bendectin was a teratogen, there was in the early cases more uncertainty. This made the disagreement more appropriate for jury resolution than in later cases, however unsatisfactory this may have seemed.[195]

In the DNA controversy, the calls for conservative match probabilities were largely a response to second order uncertainty. Some scientists claimed that one could not be confident that match probabilities were accurate because empirical research on allele variation in subpopulations was lacking. This call for conservatism rightly struck a chord with many courts.[196]

---

[193] *See* Sanders, *supra* note 140, at 42.

[194] *See* Peter Gärdenfors & Nils-Eric Sahlin, *Unreliable Probabilities, Risk Taking, and Decision Making*, 53 SYNTHESE 361 (1982) (studying classes of probability distributions of decisionmaker's knowledge); *cf.* JAMES LOGUE, PROJECTIVE PROBABILITY 87-95 (1995) (developing concept of second order probability to account for partial knowledge). For an interesting account of one way of dealing with such uncertainty in policy decisions, see Richard A. Kerr, *A New Way to Ask the Expert: Rating Radioactive Waste Risks*, 274 SCIENCE 913 (1996).

[195] *Cf.* Richard D. Friedman, *The Death and Transfiguration of* Frye, 34 JURIMETRICS J. 133, 147 (1994) ("sometimes courts ought to be willing to allow juries to take advantage of scientific information even when the scientific establishment is unwilling to declare a conclusion").

[196] The argument that match probabilities should be conservative has been criticized by Brookfield. *See* John F.Y. Brookfield, *The Effect of Relatedness on Likelihood Ratios and the Use of Conservative Estimates*, 96 GENETICA 13, 17-18 (1995). Brookfield argues that "there is no obvious reason why one should seek to revise upwards a probability on no better grounds

The concept of second order uncertainty also goes some way towards explaining the attraction of the *Frye*[197] test.[198] *Frye* is often criticized for its conservatism.[199] However, the element of "wait and see" in the *Frye* test accords with the values underlying criminal litigation, at least when it is applied to evidence adduced by the prosecution.

D. *Understanding Science in the Courts*

Another attraction of the *Frye* test is its seeming simplicity. A *Frye* court only has to ask the question "is this technique generally accepted in the scientific community?" This helps to explain why *Frye* has proved so resilient. Often, though, the simplicity of

---

than its being a probability. . . . [A] probability statement is, by its nature, an encapsulation of our partial knowledge about a situation." *Id.* Brookfield's argument can be criticized on two grounds. First, it draws on a personalist conception of probability. However, the applicability of personalist probability to legal decisionmaking has been questioned, in part on the grounds that it fails to address the problems caused by partial knowledge. *See, e.g.*, L. Jonathan Cohen, *The Role of Evidential Weight in Criminal Proof*, 66 B.U. L. REV. 635 (1986) (arguing that proof assessment of completeness of facts cannot be avoided). Some personalists concede that some concept, such as second order probability, is needed to overcome such difficulties. *See* LOGUE, *supra* note 194, at 87-95, 152-56 (recommending use of second order probability to overcome difficulties). Secondly, Brookfield's argument is informed by an over-simplistic conception of what courts are doing when they make admissibility decisions. By refusing to admit non-conservative match probabilities, a court may be refusing to let a defendant bear the risk of being convicted on unreliable evidence. Such a court may also be encouraging the prosecution to produce better evidence in order to justify a small match probability. *See* Alex Stein, *The Refoundation of Evidence Law*, 9 CAN. J.L. & JURISPRUDENCE 279, 330-31 (1996) (arguing that scientific evidence failing scientific recognition offered against defendant should always be excluded); Dale A. Nance, *The Best Evidence Principle*, 73 IOWA L. REV. 227, 230-47 (1988) (explaining best evidence principle).

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

[198] Commentators have noted that, even after the Supreme Court's decision in *Daubert*, which abrogated *Frye* in the Federal jurisdiction, case law has tended to gravitate back towards *Frye*'s general acceptance test. *See* Friedman, *supra* note 195, at 133; Martin L.C. Feldman, *May I have the Next Dance Mrs.* Frye?, 69 TULANE L. REV. 793 (1995) (noting movement back towards general acceptance test). This trend was perceptively predicted by Allen. *See* Ronald J. Allen, *Expertise and the* Daubert *Decision*, 84 J. CRIM. L. & CRIMINOL. 1157, 1173 (1993) ("I predict that [the lower courts] will continue to apply the *Frye* rule under the disguise of the [Supreme] Court's new vocabulary.").

[199] *See, e.g.*, Paul C. Giannelli, *The Admissibility of Novel Scientific Evidence:* Frye v. United States, *a Half-Century Later*, 80 COLUM. L. REV. 1197, 1223-32 (1980) (discussing issues surrounding application of *Frye*). *But 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, 739-41 (1994) (arguing that actual examples of probative evidence being excluded under *Frye* "are scarce, and those most often cited do not bear close scrutiny").

Frye is taken to a ridiculous extreme, as when some courts excluded DNA evidence post-NRC1, largely on account of news articles in the journals *Science* and *Nature.*[200]

While simple rules for dealing with scientific evidence will always prove attractive, they are also problematic. Certainly, there is no substitute for a careful review of scientific evidence — though careful scrutiny is not incompatible with a rule such as *Frye.*[201] However, judges are not scientists, and they will always find careful scrutiny of scientific evidence demanding.[202]

Much recent analysis of scientific evidence stresses just this point, that it is through coming to grips with science that courts will succeed in overcoming the problems posed by scientific evidence.[203] While this trend is to be welcomed, one should be cautious about some of the depictions of science that are being put forward as part of this "science sensitive" analysis of admissibility standards. Some of these depictions are as simplistic as the bright-line rules that they are meant to replace. For example, the Supreme Court,[204] as well as some academic commentators,[205] have stressed falsifiability as a criterion of good science. Commentators have also offered episodes such as the Lysenko/Mendel controversy as suitable examples of the scientific method.[206]

None of this grand theorizing seems to offer useful lessons for judges faced with disagreement among scientists. Falsifiability is often of little use in sorting good science from bad.[207]

---

[200] *See* Weir, *supra* note 107, at 11654 (discussing debate over DNA evidence admissibility).

[201] *See* Black et al., *supra* note 199, at 743-45 (arguing that courts can extract details of expert reasoning regardless of using *Frye*). This point also emerges from the analysis in *Developments in the Law: Confronting the New Challenges of Scientific Evidence,* 108 HARV. L. REV. 1481, 1490-1509 (1995) (stressing indeterminacy of rules of admissibility applied to scientific evidence).

[202] In this context it is encouraging to see efforts to give judges the tools they need in order to understand scientific evidence. *See, e.g.,* REFERENCE MANUAL ON SCIENTIFIC EVIDENCE (Federal Judicial Center ed., 1994).

[203] *See, e.g.,* Black et al., *supra* note 199 (discussing problems of inadequate guidance on how to understand science); *see also* Bert Black, *A Unified Theory of Scientific Evidence,* 56 FORDHAM L. REV. 595, 694-95 (1988) (concluding that courts' scientific understanding will reduce their problems with scientific evidence).

[204] *See* Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 592-94 (1993).

[205] *See, e.g.,* Black et al., *supra* note 199, at 755.

[206] *See id.* at 767-73.

[207] Falsifiability, as a criterion for demarcating science from non-science, was proposed

Further, "idealized stories of [scientific] heroes and villains"[208] trade on just the sort of hindsight that is not usually available to judges confronted by a scientific dispute.[209] If judges are to learn about science, then studying the sorts of scientific disputes examined in this Article is a good starting point. There is much to be learned from these case studies about why disagreement about scientific evidence arises and why it often can be so intractable.

CONCLUSION

Both pessimistic and optimistic conclusions can be drawn from this survey of disagreements over scientific evidence. It is clear that scientific disagreement about the reliability or probative value of evidence poses very real problems for the legal system. It is also evident that there is no magical solution to these problems.

---

by Karl Popper. *See* POPPER, *supra* note 5. However, Popper's use of falsifiability has been criticized on the grounds that the asymmetry that Popper presumed to exist between corroboration and falsification is not as great as he thought. *See* Imre Lakatos, *Falsification and the Methodology of Scientific Research Programmes, in* CRITICISM AND THE GROWTH OF KNOWLEDGE 91 (Imre Lakatos & Alan Musgrave eds., 1970). Commentators have also argued that the use of falsifiability as a demarcation criterion would render much of modern science unscientific. *See* ALAN CHALMERS, SCIENCE AND ITS FABRICATION 18 (1990) ("if this aspect of [Popper's] demarcation criterion is formulated sufficiently strongly to have some force, then physics would fail to qualify as a science"). Others point out that the aspects of science to which falsification speaks are relatively unimportant, because accepting and rejecting theories plays a minor part in science. *See* IAN HACKING, REPRESENTING AND INTERVENING 3 (1983); *see also* Allen, *supra* note 198, at 1166-75. *But see* Sean O'Connor, *The Supreme Court's Philosophy of Science: Will the Real Karl Popper Please Stand Up?*, 35 JURIMETRICS J. 263, 268-69 (1995) (questioning Allen's critique of falsificationism).

[208] JASANOFF, *supra* note 160, at 210 (commenting on Francisco J. Ayala & Bert Black, *Science and the Courts*, 81 AM. SCIENTIST 230 (1993)).

[209] On the problems of hindsight, see the sources cited *supra* note 47. An example of the problems in relying on the standard histories of the Mendel/Lysenko controversy is that there remain results, produced by Western scientists, that do not fit into the Mendelian model of genetic inheritance. *See* R.C. Lewontin, *Facts and the Factitious in Natural Sciences, in* QUESTIONS OF EVIDENCE: PROOF, PRACTICE AND PERSUASION ACROSS THE DISCIPLINES 478, 487 (James Chandler et al., eds., 1994). Black, Ayala, and Saffran-Brinks also cite Pasteur's experiments on fermentation and putrefaction as an illustrative example of good science. *See* Black et al., *supra* note 199, at 766. Again, Pasteur's results were not as conclusive as they are now portrayed as having been. *See* COLLINS & PINCH, *supra* note 11, at 79-90. These are just the sort of details that get swept under the rug when traditional histories of science are written.

The legal system cannot expect scientists to come to a speedy consensus on a disputed issue. Disagreement plays an inevitable role in science, just as it does in law. Furthermore, scientists must accept that law will continue to provoke and deepen scientific disputes. This is a natural outcome of both the adversarial nature of the legal process and its need to make decisions quickly. The need for quick decisions sometimes demands reliance on scientific propositions about which most scientists would prefer to reserve judgment.

More optimistically, law and science may come to enjoy a less troubled relationship by learning lessons from past disputes. But this will occur only if each comes to understand the other's procedures, limitations, and responses to uncertainty. It is only such shared understanding that offers hope that future scientific disputes will not be deepened by the mutual miscomprehension of science and law.