Fordham Law Review

Volume 56 | Issue 4 Article 2

1988

A Unified Theory of Scientific Evidence

Bert Black

A Unified Theory of Scientific Evidence

Cover Page Footnote

* Associate, Venable, Baetjer and Howard, Baltimore, Maryland; B.S. 1974, University of Maryland; M.S. 1975, Georgia Institute of Technology; J.D. 1982, Yale Law School. The author wishes to express his thanks and appreciation to Benjamin R. Civiletti, the managing partner of Venable, Baetjer and Howard, for his advice and support; and to Professor M. Gordon Wolman, chairman of the Department of Geography and Environmental Engineering at Johns Hopkins University, who provided access to the Johns Hopkins Libraries, where most of the research on scientific questions was done. In addition, a number of other people provided many helpful suggestions. These people included John H. Morris, Jr., Dr. Simon Black, Christopher R. Mellott, Frederick C. Leiner, Robert G. Smith, Ben Finkelstein, Judy F. Black, Paul F. Strain, Dorothy G. Black, Dr. Roy A. Black, Dr. Alan T. Stone, Anthony M. Carey, and Thomas M. Lingan. Finally, thanks go to John S. Nixdorff, the librarian at Venable, Baetjer and Howard, who was tireless in tracking down countless esoteric references.

A UNIFIED THEORY OF SCIENTIFIC EVIDENCE

*BERT BLACK**

TABLE OF CONTENTS

* Associate, Venable, Baetjer and Howard, Baltimore, Maryland; B.S. 1974, University of Maryland; M.S. 1975, Georgia Institute of Technology; J.D. 1982, Yale Law School.

The author wishes to express his thanks and appreciation to Benjamin R. Civiletti, the managing partner of Venable, Baetjer and Howard, for his advice and support; and to Professor M. Gordon Wolman, chairman of the Department of Geography and Environmental Engineering at Johns Hopkins University, who provided access to the Johns Hopkins Libraries, where most of the research on scientific questions was done. In addition, a number of other people provided many helpful suggestions. These people included John H. Morris, Jr., Dr. Simon Black, Christopher R. Mellott, Frederick C. Leiner, Robert G. Smith, Ben Finkelstein, Judy F. Black, Paul F. Strain, Dorothy G. Black, Dr. Roy A. Black, Dr. Alan T. Stone, Anthony M. Carey, and Thomas M. Lingan. Finally, thanks go to John S. Nixdorff, the librarian at Venable, Baetjer and Howard, who was tireless in tracking down countless esoteric references.

INTRODUCTION

SCIENTIFIC evidence has always posed special problems for the law,[1] and in recent years these problems have become increasingly difficult.[2] Because an expert can be found to support almost any posi-

1. The use of scientific evidence dates from as far back as the fourteenth century. *See* Hand, *Historical and Practical Considerations Regarding Expert Testimony*, 15 Harv. L. Rev. 40, 42-43 (1901). Nonetheless, the law has yet to develop a satisfactory way to take advantage of the special knowledge offered by science while assuring the validity and reliability of the opinions offered by expert witnesses.

As Professor Korn has observed,

> [m]uch has been written about the patent weaknesses of a system which provides for the presentation of [specialized] information by party-selected and compensated expert witnesses to an otherwise untutored and unassisted lay tribunal. Attention has also been directed to impediments imposed by the formal trial process and rules of evidence, especially those governing expert testimony.

Korn, *Law, Fact, and Science in the Courts*, 66 Colum. L. Rev. 1080, 1080-81 (1966) (footnotes omitted); *see also* Elliott & Spillman, *Medical Testimony in Personal Injury Cases*, 2 Law & Contemp. Probs. 466, 466 (1935) (quoting commencement address to 1915 Harvard Medical School class in which speaker described expert medical witnesses at trial as coming from "two hostile camps, and prepared to attempt, under solemn oath, to uphold opinions diametrically opposed, yet supposedly derived from a single series of facts and observations"); Ford & Holmes, *The Professional Medical Advocate*, 17 Sw. L.J. 551, 552 (1963) ("Some courts and writers are of the opinion that [expert medical testimony] is the most unsatisfactory and unreliable part of judicial administration."); Smith, *Scientific Proof and Relations of Law and Medicine*, 23 B.U.L. Rev. 143, 146 (1943) (as long ago as 1909, 60% of cases in Suffolk County Superior Court, Massachusetts, involved expert testimony); Note, *Scientific Evidence and the Question of Judicial Capacity*, 25 Wm. & Mary L. Rev. 675, 675 (1984) (one of the most serious problems with scientific evidence "has been the inability of lay factfinders to understand complex scientific and technical evidence").

Courts and lawyers often just wash their hands of science. In Smith v. W. Horace Williams Co., 84 So. 2d 223 (La. Ct. App. 1956), for example, the court noted that it had "[n]ever . . . been confronted with a more formidable record of contradictory medical testimony." *Id.* at 225. It then simply deferred to the lower court, declaring that a "study of this tremendous record leaves us almost hopelessly confused." *Id.* at 227.

2. As one writer has recently noted:

> In the past twenty years courts have confronted an increasing variety of controversies over the safety of new technologies, the liability for harm caused by technological hazards, and the validity or sufficiency of regulatory policies for controlling science and technology. Despite the emergence of these technically complex disputes, there has been remarkably little institutional or procedural reform to ensure a smoother legal passage for such cases.

Jasanoff, *Science and the Courts: Advice for a Troubled Marriage*, 2 Nat. Resources & Env't 3, 3 (Fall 1986); *see also* M. Saks & R. Van Duizend, The Use of Scientific Evidence in Litigation 3 (1983) ("[P]ractitioners and policymakers have had to cope with the problems engendered by [scientific and technological information] in largely unsystematic ways."). The bitterly contested testimony by scientific experts that regularly surfaces in toxic tort litigation involving disease or injury allegedly caused by exposure to chemicals

tion,[3] most commentators agree on the need for judicial review and control,[4] but there is no consensus on how to achieve these objectives.[5] This

---

or radiation, or by the use of pharmaceuticals or other products, causes particular problems. *See infra* note 29.

3. *See* Ford & Holmes, *supra* note 1, at 553 (discussing problem of professional testifiers); Graham, *Expert Witness Testimony and the Federal Rules of Evidence: Insuring Adequate Assurance of Trustworthiness*, 1986 U. Ill. L. Rev. 43, 45 ("Today practicing lawyers can locate quickly and easily an expert witness to advocate nearly anything the lawyers desire."); Huber, *Safety and the Second Best: The Hazards of Public Risk Management in the Courts*, 85 Colum. L. Rev. 277, 333 (1985) ("The scientific community is large and heterogeneous, and a Ph.D. can be found to swear to almost any 'expert' proposition, no matter how false or foolish."); Weinstein, *Improving Expert Testimony*, 20 U. Rich. L. Rev. 473, 482 (1986) ("An expert can be found to testify to the truth of almost any factual theory, no matter how frivolous, thus validating the case sufficiently to avoid summary judgment and force the matter to trial. . . . Juries and judges can be, and sometimes are, misled by the expert-for-hire."); *see, e.g.*, Ladner v. Higgins, Inc., 71 So. 2d 242, 244 (La. Ct. App. 1954) (witness acknowledged that his opinion depended on which side had retained him).

4. The National Conference of Lawyers and Scientists, a joint entity of the American Bar Association and the American Association for the Advancement of Science, recently concluded an extensive review of the rules governing scientific evidence. The group reached a general consensus that some special screening of scientific evidence is necessary, but reached no agreement on exactly what rule should be used. *See Rules for Admissibility of Scientific Evidence*, 115 F.R.D. 79, 81 (1987) (symposium report); *see also* Boyce, *Judicial Recognition of Scientific Evidence in Criminal Cases*, 8 Utah L. Rev. 313, 325 (1963-64) (advocating reasonable reliability test); Giannelli, *The Admissibility of Novel Scientific Evidence:* Frye v. United States, *a Half-Century Later*, 80 Colum. L. Rev. 1197, 1245-50 (1980) (advocating a reliability test that shifts the burden of proving reliability to the party offering scientific testimony into evidence); Graham, *supra* note 3, at 90 (advocating common law barriers to admissibility, including the *Frye* standard, the requirement of disclosure of the basis for an expert's opinion, the reasonable certainty test, and the prohibition of opinions on ultimate issues to ensure trustworthiness); Lipton, *The Results of Scientific Techniques as Evidence in Federal Courts: Evolution of the* Frye v. United States *Standard in the Period 1969-1977*, 8 Envtl. L. 769, 769-70 (1978) (comparing traditional general acceptance approach with modern approach, which focuses on accuracy); McCormick, *Scientific Evidence: Defining a New Approach to Admissibility*, 67 Iowa L. Rev. 879, 911-12 (1982) (advocating eleven-factor balancing test focusing on accuracy and reliability); Moenssens, *Admissibility of Scientific Evidence—An Alternative to the* Frye *Rule*, 25 Wm. & Mary L. Rev. 545, 573-74 (1984) (emphasizing focus on reliability, citing the eleven-factor test proposed by Judge McCormick); Note, People v. Murtishaw: *Applying the* Frye *Test to Psychiatric Predictions of Dangerousness in Capital Cases*, 70 Calif. L. Rev. 1069, 1089 (1982) (advocating general acceptance test); Note, *Expert Testimony Based on Novel Scientific Techniques: Admissibility Under the Federal Rules of Evidence*, 48 Geo. Wash. L. Rev. 774, 777, 790 (1980) (advocating general acceptance test as best guarantee of an available pool of experts and of scientific validity); Case Comment, *Changing the Standard for the Admissibility of Novel Scientific Evidence:* State v. Williams, 40 Ohio St. L.J. 757, 768-69 (1979) (advocating flexible reliability test).

5. *See supra* note 4. Three reports published in Federal Rules Decisions since 1983 provide an excellent summary of current legal scholarship on scientific evidence. The first report, *Symposium on Science and the Rules of Evidence*, 99 F.R.D. 187 (1983), is concerned primarily with the *Frye* test of general acceptance and alternatives to it. *See id.*; *see also infra* notes 22-24 and accompanying text, and *infra* Part II. The second report, *Symposium on Science and the Rules of Legal Procedure*, 101 F.R.D. 599 (1983), discusses the relationship between science and the rules of legal procedure. The third report, *Rules for Admissibility of Scientific Evidence*, 115 F.R.D. 79 (1987), contains four proposals for a model rule on admissibility of scientific evidence. One of these proposals is essentially procedural. *See id.* at 102-07. A second proposal advocates a reliability test,

Article argues that solving the problems surrounding the use and interpretation of scientific evidence requires a unified, coherent approach to deciding admissibility that covers all areas of science and all kinds of cases. In order to develop such an approach, this Article proposes a theoretical framework based on distinguishing two aspects of relevancy: (1) the *validity* of the reasoning leading to a conclusion, and (2) the *reliability* of the conclusion.

Validity and reliability, though intertwined, are very different concepts. One normally speaks of "valid" rather than "reliable" reasons or theories, and of "reliable" rather than "valid" instruments or machines. Results, conclusions, or techniques may be either valid or reliable. Behind these simple examples of everyday usage lie largely overlooked conceptual distinctions and relationships that are fundamental to a coherent legal theory of scientific evidence.

As used in this Article, reliability means that a successful outcome, or a correct answer, is sufficiently probable for a given situation.[6] A baseball player who gets a hit forty percent of the time is an extremely reliable batter,[7] but a lie detection device that correctly indicates falsehood ninety percent of the time may not be reliable enough for use at trial.[8] In contrast to reliability, validity means that which results from sound and cogent reasoning.[9] An invalid conclusion cannot be reliable, yet valid reasoning does not necessarily lead to reliable conclusions.[10] Reliability is the ultimate legal concern,[11] but when it hinges on controversial and

---

see id. at 84-88; a third, a relevancy test, determined in accordance with a Rule 403 balancing process, *see id.* at 89-91; and a fourth, a validity test. *See id.* at 92-101. These proposals are presented as *alternate* modifications to Rule 702. Thus, all three of the nonprocedural proposals miss the point that reliability and validity simply are two aspects of relevancy. *See infra* Part I.B. None of the proposals provides any guidance on how to determine validity.

6. According to the dictionary, the term reliable "describes what can be counted on or trusted in to do as expected or to be truthful." Webster's Third New International Dictionary of the English Language Unabridged 1917 (1961).

7. Indeed, Tommy Henrich, an outfielder with the New York Yankees from 1937 until 1950, was known as Old Reliable, even though his lifetime batting average was .282. He batted over .300 only three times in eleven Major League seasons. *See* J. Reichler, The Baseball Encyclopedia 959 (5th ed. 1982).

8. *See, e.g.*, People v. Davis, 343 Mich. 348, 371, 72 N.W.2d 269, 282 (1955) (refusing to admit lie detector test results because "[t]here is still a percentage of error in the results, estimated at being from less than 10% to 25% by various authorities").

9. The dictionary defines the term "valid" as "well grounded or justifiable [and] applicable to the matter at hand." As an example of how the word is used, the dictionary describes a valid argument or principle as one which is "supported either by objective truth or a generally accepted standard or authority." Webster's Third New International Dictionary of the English Language Unabridged 2529-30 (1961).

10. It is possible for incorrect reasoning to produce a correct answer by chance, but there is no way for a court or a jury to know when a lucky guess has been made. If a method is based on incorrect premises, but still provides accurate predictions, it may be a valid predictor. For some purposes, it would thus meet the validity portion of the standard proposed by this Article.

11. *See* Saltzburg, Frye *and Alternatives*, 99 F.R.D. 208, 212 (1983).

contested reasoning, the validity of that reasoning must be addressed.[12]

Distinguishing between validity and reliability is important because it permits the separation of scientific questions from legal questions. For example, when an expert concludes that there exists a five percent probability that a plaintiff's disease was caused by a product made by a defendant, there is the scientific question whether or not the expert is correct, and the legal question whether this probability level warrants presenting the conclusion to the jury.[13] This Article views the scientific question as a matter of *validity*, with the answer depending on accepted scientific practice and the soundness and cogency of the entire pattern of reasoning leading to the expert's conclusion.[14] In contrast, the legal question relates to how much *reliability* the law requires, with the answer depending on legal standards.[15]

Part I of this Article establishes valid reasoning and reliable conclusions as the two focal points for the legal analysis of scientific evidence, and demonstrates how the legal concept of relevancy embraces both validity and reliability. Courts, however, usually ignore the validity aspect of relevancy, even though the Federal Rules of Evidence are flexible enough to permit them to address it.[16] This Article therefore proposes a

12. *See infra* Part I.B.

13. There is also the legal question of how a given level of probability translates into liability. Under most legal theories, for example, a causal relationship between the defendant's culpable conduct and the plaintiff's injury must be proven by a preponderance of the evidence. That is, it must be shown that, more likely than not, the relationship exists. *See* Black & Lilienfeld, *Epidemiologic Proof in Toxic Tort Litigation*, 52 Fordham L. Rev. 732, 749-50 (1984).

14. *See infra* Part I.C. (discussing the modern view of how science validates theories); Part II.D.4 (discussing how scientists reason by interweaving theory and observation).

15. *See infra* notes 52-70, 176-81 and accompanying text.

16. The Federal Rules of Evidence contain no explicit requirement that the reasoning that underlies a scientific opinion conform to any standard. Rule 702 deals with the qualifications of an expert, *see* Fed. R. Evid. 702, and Rules 703 and 705 deal with the factual basis for an expert's conclusion, *see* Fed. R. Evid. 703 & 705, but no rule directly addresses the reasoning that must connect the basis to the conclusion. Some courts have used Rule 702 or Rule 703 to exercise control in this area, though very few decisions explicitly consider the fundamental question whether the connective reasoning is valid. *See, e.g.*, Lynch v. Merrell-National Labs., Inc., 830 F.2d 1190, 1194-95 (1st Cir. 1987) (rejecting extrapolation of animal test results to humans, and rejecting an epidemiologic analysis that had never been refereed or published); Marder v. G.D. Searle & Co., 630 F. Supp. 1087, 1090 (D. Md. 1986) ("The foundation for the [expert] testimony about [relationship between the Cu-7 IUD and disease] was virtually nonexistent; it was not backed by sound scientific evidence that was related specifically to the Cu-7 or IUD's which were comparable to the Cu-7."), *aff'd mem.*, Wheelahan v. G.D. Searle & Co., 814 F.2d 655 (4th Cir. 1987); *In re* "Agent Orange" Prod. Liab. Litig., 611 F. Supp. 1223 (E.D.N.Y. 1985) (discussed *infra* notes 482-96), *aff'd on other grounds*, 818 F.2d 187 (2d Cir. 1987); Johnston v. United States, 597 F. Supp. 374 (D. Kan. 1984), *cert. denied*, 108 S. Ct. 694 (1988) (discussed *infra* notes 501-05).

Chief Judge Jack B. Weinstein observed that "[t]he very flexibility of the current rules regulating expert testimony requires judges to exercise more control in some cases." Weinstein, *supra* note 3, at 473. Professor Imwinkelried advocates the use of Federal Rules of Evidence 901, 611, and 403 to reach the validity of theories or techniques. *See* Imwinkelried, *Judge Versus Jury: Who Should Decide Questions of Preliminary Facts*

clarification of Rule 702[17] that would explicitly require consideration of both scientific validity and the legal significance of reliability in deciding whether to admit scientific evidence.[18] In order to provide guidance on how the law can evaluate the validity of scientific reasoning, Part I also discusses the way in which science itself determines validity. Contrary to longstanding misperceptions, science does not generate exact knowledge with logical certainty. Instead, it relies on the give and take of criticism, testing, experimentation, and review to determine what is valid.[19] Widespread consensus and acceptance, therefore, is the central test that scientists use to decide the validity of theories and reasoning in any given context,[20] which is a test that the law can adopt and use successfully.

Part II uses the proposed validity-reliability analysis to resolve a longstanding debate about whether scientific evidence should be subject to the same rule of relevancy as any other evidence[21] or to the test of *Frye v. United States*,[22] which is still followed in most jurisdictions.[23] Under *Frye*, courts admit "expert testimony deduced from a well-recognized

---

Conditioning the Admissibility of Scientific Evidence?, 25 Wm. & Mary L. Rev. 577, 606-16 (1984) [hereinafter Imwinkelried I]. *See generally* Fed. R. Evid. 901, 611 & 403. Other commentators discuss how active control of scientific evidence fits within the Federal Rules. Professor Graham, for example, analyzes the way courts have applied the Rules. He does not advocate any changes in the Rules themselves, but he does suggest that they be interpreted in light of common law requirements, such as general acceptance, which were developed to ensure trustworthiness. *See* Graham, *supra* note 3, at 89-90; *see also* Rothstein & Crew, *When Should the Judge Keep Expert Testimony From the Jury?*, 1 Inside Litigation 19, 26 (Apr. 1987) (noting that strict scrutiny of scientific evidence "seems to be an accelerating modern movement and is the direction of the future").

17. *See infra* notes 79-81 and accompanying text.

18. Sufficiency and admissibility are two distinct issues. Evidence can be admissible but not sufficient to sustain a party's burden of proof on a particular issue. *See* Martin, *The Uncertain Rule of Certainty: An Analysis and Proposal for a Federal Evidence Rule*, 20 Wayne L. Rev. 781, 797-802 (1974); Musslewhite, *Medical Causation Testimony in Texas: Possibility Versus Probability*, 23 Sw. L.J. 622, 622 (1969); Note, *Causation in Disease: Quantum of Proof Required to Reach the Jury*, 53 Nw. U.L. Rev. 793, 796-98 (1959).

Scientific evidence can be controlled using either an admissibility or a sufficiency test. The distinction is not one of great concern for the purposes of this Article. The difference is more a question of whether the evidence is central enough to the litigation to warrant judgment against the offering party if it is deficient. In terms of this Article's two-step approach, admissibility is more closely tied to the validity of reasoning, while sufficiency is more closely related to the question of whether testimony is reliable enough to warrant presentation to the trier of fact. Sufficiency is also related to the question whether a conclusion is reliable enough to satisfy substantive legal requirements, such as the preponderance of the evidence test.

19. *See infra* Part I.C.3.

20. *See id.*

21. *See infra* note 158 and accompanying text.

22. 293 F. 1013 (D.C. Cir. 1923).

23. *See* Note, *The* Frye *Doctrine and Relevancy Approach Controversy: An Empirical Evaluation*, 74 Geo. L.J. 1769, 1769 (1986). *But see* Rossi, *Modern Evidence and the Expert Witness*, 12 Litigation 18, 20 (Fall 1985) (noting that "[w]ithin the last decade, courts in more than 15 jurisdictions have rejected *Frye*"). Both courts and commentators have attacked *Frye*. For example, in United States v. Downing, 753 F.2d 1224 (3d Cir. 1985), the court observed:

scientific principle or discovery, [only if] the thing from which the deduction is made [is] sufficiently established to have gained *general acceptance* in the particular field in which it belongs."[24] Although based on acceptance, *Frye* does not indicate whether it is the conclusions that result from reasoning or the reasoning itself that must be accepted. *Frye* makes acceptance the exclusive test for admissibility, whereas this Article's proposed mode of analysis looks to acceptance only as the best indication of the validity of the reasoning supporting a particular scientific opinion.[25] Part II further argues that when reasoning is not at issue, acceptance should have no bearing on legal disputes about the reliability of scientific conclusions.[26]

Part III illustrates how the proposed approach would greatly improve the evaluation of medical testimony, to which the test of "reasonable medical certainty"[27] is now usually applied. Use of the traditional standard generally means that courts do no more than consider the qualifications of an expert, the factual basis for his or her conclusion, and the degree of confidence expressed in his or her opinion.[28] As a result, courts scrutinize neither the validity of the expert's reasoning nor the reliability of his or her conclusions. In recent years, the traditional test has been stretched to the breaking point by the kind of medical controversies that regularly surface in toxic tort litigation.[29] When experts clash in this

---

> In sum, the *Frye* test suffers from serious flaws. The test has proved to be too malleable to provide the method for orderly and uniform decision-making envisioned by some of its proponents. Moreover, in its pristine form the general acceptance standard reflects a conservative approach to the admissibility of scientific evidence that is at odds with the spirit, if not the precise language, of the Federal Rules of Evidence.

Id. at 1237; *see also Symposium on Science and Rules of Evidence*, 99 F.R.D. 187, 188 (1983) (noting the "intractable ambiguity" of *Frye*'s general acceptance test); Giannelli, Frye v. United States, *Background Paper Prepared for the National Conference of Lawyers and Scientists*, 99 F.R.D. 189, 191-93 (1983) (summarizing criticisms of the *Frye* test). *But see* Graham, *supra* note 3, at 55 (advocating the continued use of *Frye*).

24. *Frye*, 293 F. at 1014 (emphasis added).

25. *See infra* notes 155-56 and accompanying text.

26. *See infra* notes 161-85 and accompanying text.

27. *See* Nesson, *Agent Orange Meets the Blue Bus: Factfinding at the Frontier of Knowledge*, 66 B.U.L. Rev. 521, 527 (1986) (traditional standard for admitting a medical opinion "requires the expert to be qualified, to have used respected [though not necessarily scientific] methodology in arriving at his opinion, and to express confidence in his opinion to a reasonable degree of medical certainty"); *see also infra* note 371.

28. *See infra* notes 370-71 and accompanying text.

29. This Article loosely defines toxic tort cases as those in which plaintiffs seek compensation for harm allegedly caused by exposure to a substance that increases the risk of contracting a serious disease, but generally involve a period of latency or incubation prior to the onset of the disease. Exposure to radiation and the use of pharmaceutical drugs or products fall within this loose rubric. Many kinds of injuries are involved, but most toxic tort cases involve cancer and the issue of carcinogenesis. *See* Black & Lilienfeld, *supra* note 13, at 739-49. An excellent discussion of the variety of cases that fall within the toxic tort category is provided in Feinberg, *The Toxic Tort Litigation Crisis: Conceptual Problems and Proposed Solutions*, 24 Hous. L. Rev. 155 (1987).

On the problems toxic tort cases create for the law, see Nesson, *supra* note 27, at 528 n.24 (noting that the "ill-defined standard" of "reasonable certainty" now used for medi-

context, courts relying on the traditional medical certainty test have often been unable or unwilling to differentiate between opinions based on science and those based on mere speculation by individual scientists.[30]

The analytical framework proposed in this Article calls upon courts to evaluate scientific reasoning, which raises the question of how nonexperts can judge the way in which scientists reach their conclusions. Both courts and lawyers generally have shunned this kind of review as distasteful and beyond their abilities.[31] Part IV addresses this concern by

---

cal evidence can lead to "striking inconsistencies" in the way courts respond to medical experts' conclusions); *see also* Institute for Health Policy Analysis, Georgetown University Medical Center, Conference Panel Report, Causation and Financial Compensation 41 (1986) (observing that many judges permit expert testimony that many scientists would consider unsupported by data and logically or methodologically unsound); Riley, *Toxic Shock Syndrome: Proving Causation Before Science Has*, 6 Am. J. Trial Advoc. 15, 15 (1982) (noting disagreement among competent experts); Rosenberg, *The Causal Connection in Mass Exposure Cases: A "Public Law" Vision of the Tort System*, 97 Harv. L. Rev. 849, 855-59 (1984) (discussing problem of proving causation); Weinstein, *supra* note 3, at 473 (questioning whether we ask too much of expert witnesses, especially in toxic tort and psychiatric cases); *cf.* Nesson, *supra* note 27, at 529 (suggesting that opinions about disease causation that are not scientifically supported should still be admissible and sufficient).

Typical recent toxic tort cases include Wells v. Ortho Pharmaceutical Corp., 788 F.2d 741 (11th Cir.) (birth defects allegedly caused by contraceptive foam), *reh'g denied en banc*, 795 F.2d 89 (11th Cir.), *cert. denied*, 107 S. Ct. 437 (1986); Ferebee v. Chevron Chem. Co., 736 F.2d 1529 (D.C. Cir.) (lung disease allegedly caused by exposure to herbicide), *cert. denied*, 469 U.S. 1062 (1984); Lynch v. Merrell-National Labs., Inc., 646 F. Supp. 856 (D. Mass. 1986) (Bendectin case), *aff'd*, 830 F.2d 1190 (1st Cir. 1987); Marder v. G.D. Searle & Co., 630 F. Supp. 1087 (D. Md. 1986) (infection and other injuries allegedly caused by IUD), *aff'd mem.*, Wheelahan v. G.D. Searle & Co., 814 F.2d 655 (4th Cir. 1987); *In re* "Agent Orange" Prod. Liab. Litig., 611 F. Supp. 1223 (E.D.N.Y. 1985) (health problems allegedly caused by Agent Orange), *aff'd on other grounds*, 818 F.2d 187 (2d Cir. 1987); and Johnston v. United States, 597 F. Supp. 374 (D. Kan. 1984) (cancer allegedly caused by low-level radiation exposure), *cert. denied*, 108 S. Ct. 694 (1988).

30. *See infra* Part III.B.

31. Professor Graham has observed that "[a]n underlying problem is that lawyers do not understand science, including the fundamentals of the scientific method and the techniques by which scientific evidence is generated. Unfortunately, many of the lawyers who might benefit most by overcoming that deficiency exhibit a reluctance to try." *Symposium on Science and the Rules of Evidence*, 99 F.R.D. 187, 232 (1983) (remarks during plenary session).

Joseph Nicol has observed that a very important consideration in dealing with scientific evidence

> is the inability of the defense bar to handle scientific matters. The sad truth is that those attorneys simply are incapable by education, and all too often by inclination, to become sufficiently familiar with scientific evidence to discharge their responsibilities toward the administration of justice. The scientific illiteracy of nearly all lawyers is a disgrace to their profession. The fault lies equally with the individual lawyers and with the legal profession, including law schools and bar associations.

Id. at 221.

In an article calling for a more detailed evaluation of the weight of scientific evidence, Professor Imwinkelried has written that "[f]or their part, in this new era of scientific evidence attorneys must not only be familiar with the legal standards governing the admissibility of scientific proof; attorneys must also familiarize themselves with the data

examining cases that establish that courts do have the ability to deal with science. These cases show that when expert witnesses are forced to make their reasoning explicit, courts can evaluate it against accepted scientific practice.[32]

Part IV also considers the way in which application of the proposed unified theory would actually affect the admissibility of scientific evidence. Finally, it reviews the procedural options available to courts for controlling testimony based on invalid reasoning. These options range from pretrial *in limine* exclusion to posttrial judgment notwithstanding the verdict, but this Part argues that the pretrial procedures generally are preferable because they permit more thorough and efficient review.[33] This Article concludes that the problems surrounding scientific evidence can be resolved by separating scientific concerns from legal ones and by excluding scientific conclusions based more on a particular scientist's personal biases than on sound scientific reasoning.[34]

---

bases and interpretive standards that facilitate evaluating the weight of scientific evidence." Imwinkelried, *A New Era in the Evolution of Scientific Evidence—A Primer on Evaluating the Weight of Scientific Evidence*, 23 Wm. & Mary L. Rev. 261, 289-90 (1981) [hereinafter Imwinkelried II].

The call for a more scientifically informed legal profession is not new. *See* Cavers, *Introduction to Science and the Law Symposium*, 63 Mich. L. Rev. 1325, 1334 (1965) ("the law may seem condemned to a chronic catching-up process"); Loevinger, *Law and Science as Rival Systems*, 8 Jurimetrics J. 63, 70 (Dec. 1966) (noting traditional hostility of legal profession to study of science or its methods); *cf.* Goldberg, *The Reluctant Embrace: Law and Science in America*, 75 Geo. L.J. 1341, 1350 (1987) ("It would be a mistake to believe, however, that these differences between law and science prevent members of these professions from understanding each other. There is no reason why lawyers and scientists cannot comprehend the different nature of the other's work and appreciate when it is being done well.").

Professor Goldberg views the primary difference between law and science as being rooted in the fact that science focuses on *progress* while the law focuses on *process*. *See id.* at 1345-46. While this may be true, it should not be forgotten that scientists work towards their goal of progress through the scientific process. It is the premise of this Article that when scientists testify as expert witnesses, they should base their testimony on this scientific process.

32. *See infra* Part IV.B. As Professor Ladd once put it,

> [i]n most cases . . . scientific problem[s] [are] not so far beyond the reach of the ordinary person that the reasons and theory back of an opinion serve no purpose in creating an intelligent understanding. Quite to the contrary, qualified experts may be used to throw a surer light upon a case by refinement of the average juryman's knowledge.

Ladd, *Expert Testimony*, 5 Vand. L. Rev. 414, 429 (1952).

33. *See infra* Part IV.D.

34. On the problem of personal biases of scientists, see *infra* note 214. These biases may manifest themselves in the willingness of witnesses to tailor testimony to their client's needs, *see supra* note 3, or to give opinions about ultimate issues clearly beyond their expertise. For example, in Commonwealth v. Hart, 348 Pa. Super. 117, 501 A.2d 675 (1985), a doctor testified that, based on the nature of the wounds on a child's corpse, the child had been murdered. *See id.* at 121, 501 A.2d at 677.

I. AN ANALYTICAL FRAMEWORK BASED ON DISTINGUISHING THE VALIDITY OF REASONING FROM THE RELIABILITY OF CONCLUSIONS

A. *Reasoning and Conclusions as the Focal Points for the Legal Analysis of Scientific Evidence*

The law's contact with science almost always comes through expert witness testimony. An expert testifies about his or her *conclusions*, which are based on facts generally relied upon in the expert's field[35] and on the facts in evidence. *Reasoning* connects the facts to the conclusions.[36] A legal analysis centered on reasoning and conclusions therefore addresses the most fundamental concerns about the admissibility of scientific evidence. Most commentators, however, have focused their analyses on discrete scientific techniques,[37] devices,[38] principles,[39] or theories.[40] This diffuse approach masks the basic conceptual issues at the core of all evidentiary disputes about science.[41]

---

35. Fed. R. Evid. 703.

36. *See* McCoid, *Opinion Evidence and Expert Witnesses*, 2 UCLA L. Rev. 356, 356 (1955) (experts allowed to give opinions because they are capable of drawing inferences from proved facts).

37. *See, e.g.*, Giannelli, *supra* note 4, at 1198 & nn. 1-6 (giving as examples neutron activation analysis, sound spectrometry, psycholinguistics, atomic absorption, remote electromagnetic sensing, and bitemark comparisons); Imwinkelried I, *supra* note 16, at 577 (referring to theories underlying scientific techniques); Imwinkelried II, *supra* note 31, at 262 (referring to forensic techniques); Lacey, *Scientific Evidence*, 24 Jurimetrics J. 254, 255 (1984) (focusing on validity of principle underlying a technique, on validity of technique applying the principle, and on proper application of the technique); Moenssens, *supra* note 4, at 545 (referring to *Frye* as an obstacle to evidence based on novel scientific techniques).

38. *See, e.g.*, Boyce, *supra* note 4, at 315-23 (referring to various devices).

39. *See, e.g.*, *id.* at 313; Giannelli, *supra* note 4, at 1198; Lacey, *supra* note 37, at 255.

40. *See, e.g.*, Boyce, *supra* note 4, at 313 ("Enforcement agencies, both public and private, have been applying new devices and theories of a scientific nature in an effort to stem the rising crime rate."); Imwinkelried I, *supra* note 16, at 579-82 (discussing importance of determining validity of underlying scientific theory).

41. Professor Imwinkelried has noted that the "typical text [on scientific evidence] is a catalogue of scientific techniques. . . . These texts do not even attempt to synthesize the voluminous material and do not offer the attorney an overall approach to evaluating the probative weight of the myriad techniques mentioned in the texts." Imwinkelried II, *supra* note 31, at 273. For a recent example, see J. Tarantino, Strategic Use of Scientific Evidence (1988), which, after a brief introduction, launches into discussions of specific techniques ranging from fingerprinting to thermography. The introductory chapter discusses both reliability and validity, but neither clearly defines these terms nor articulates their relationship within an analytical framework. *See id.* at 3-20.

Disputes about reliability, which is the law's ultimate concern, may involve the question of how well a device or a technique works, or questions about the validity of a principle or a theory, but addressing these questions comes down to the available empirical evidence and the expert's reasoning about that evidence. Even for a relatively simple device such as a radar speed detector, the legal determination of reliability depends on reasoned, expert conclusions about the "construction, the operation and the purpose of [the instrument in question], its margin of error if properly functioning, and the ways and means of testing its accuracy." State v. Moffitt, 48 Del. 210, 213, 100 A.2d 778, 779 (Super. Ct. 1953). *See generally* Annotation, *Proof By Radar or Other Mechanical or*

Disputes about techniques or devices usually involve questions about the accuracy of what they measure or determine,[42] or about the interpretation of measurements or results.[43] Whether it is a question of measurement or interpretation, however, the issue comes down to reasoning and conclusions. What *reasons* lead to the *conclusion* that a breath analyzer consistently measures the alcohol content of a person's breath, and what *reasons* lead to the *conclusion* that such a measurement provides an accurate and consistent enough indication of blood alcohol level upon which to base a conviction for drunk driving?[44] Likewise, what *reasons* underlie the technique of HLA blood type matching, and how reliable are the *conclusions* it provides about paternity?[45]

Evidentiary issues related to scientific principles and theories also are best analyzed within a framework that examines the reasoning connecting facts to conclusions and separately examines the conclusions themselves. Both principles and theories result from reasoning, and can be viewed as general propositions or conclusions upon which further reasoning is based.[46] In addition, the word "theory" may have a less universal meaning. When a scientist refers to a theory, he or she may mean nothing more than an explanation for a specific event, such as why a rocket exploded.[47] In either the general or the specific sense, however, the evaluation of theories reduces to an evaluation of reasoning. Indeed, in the specific sense, an expert's theory *is* his or her reasoning.

Past commentary on scientific devices, techniques, principles, or theories[48] therefore indirectly reflects concerns about the same basic concep-

---

Electronic Devices, of Violation of Speed Regulations, 47 A.L.R. 3d 822 (1973). Resolving legal issues related to more complex and controversial areas of science surely requires no less attention to the reasoning behind an expert's conclusions, or to the legal significance of those conclusions.

Professor Graham has outlined six factors that govern the reliability of evidence derived from a scientific principle: "(1) the validity of the underlying scientific principle, (2) the validity of the technique or process that applies the principle, (3) the condition of any instrumentation used in the process, (4) adherence to proper procedures, (5) the qualifications of the person who performs the test, and (6) the qualifications of the person who interprets the results." Graham, *supra* note 3, at 52 (citing E. Imwinkelried, P. Giannelli, F. Gilligan & F. Lederer, Criminal Evidence 83-90 (1979)). This list certainly touches upon the kind of relationship that must exist between principle and application, but it overlooks the fact that the reasoning connecting factual evidence to expert conclusions is the more fundamental concern. The validity of a scientific principle depends on the reasoning behind it, as does the validity of any application of such a principle.

42. For example, does a radar speed detector accurately measure speed, or does a blood test accurately determine blood alcohol level? Similar questions can be raised about almost any measuring or analytical technique.

43. For example, does a given blood alcohol level imply inebriation, or do nondestructive metallurgical test results imply that steel is of a given tensile strength?

44. *See, e.g.*, State v. Johnson, 717 S.W.2d 298, 300 (Tenn. Crim. App. 1986).

45. *See generally* Ellman & Kaye, *Probabilities and Proof: Can HLA and Blood Group Testing Prove Paternity?*, 54 N.Y.U. L. Rev. 1131 (1979).

46. An accepted theory, like the theory of relativity, or electromagnetic theory, can be viewed as "pre-packaged" reasoning.

47. *See infra* note 96 and accompanying text.

48. *See supra* notes 37-40.

tual issues that this Article addresses directly. With the validity of reasoning and the reliability of conclusions as the focal points for analysis, a single, unified approach for all evidentiary applications of science becomes possible. Furthermore, this approach accords with the traditional law of evidence.[49]

B. *The Reliability and Validity Aspects of Relevancy*

The law of evidence incorporates both validity and reliability into the conceptual touchstone of relevancy. Evidence is relevant if it has any tendency to make the existence of a fact at issue more or less probable than it would be without the evidence.[50] Whether a relationship exists between a piece of evidence and a fact at issue "depends upon principles evolved by experience or science, applied logically to the situation at hand."[51] Relevancy thus encompasses questions about both the validity of the reasoning that links evidentiary facts to conclusions *and* the reliability of those conclusions. This Article's proposed analytical framework addresses the same issues.

In deciding issues about scientific evidence, however, courts generally have fixed on the reliability of conclusions. The validity of the reasoning leading to conclusions usually is ignored, often because the issue of validity is not raised. *United States v. Williams*[52] provides an excellent example of the conceptual predominance of reliability, and demonstrates how reliability connotes probability of correctness. One of the issues in *Williams* involved the identity of a telephone caller whose voice had been recorded during a call he had made to an undercover officer about a proposed heroin sale.[53] The trial court allowed the government to introduce testimony about a voice spectrograph (or voiceprint) comparison[54] to help establish that the defendant had been the caller.[55] After his conviction, the defendant appealed, arguing that evidence based on the technique should have been excluded.[56]

The United States Court of Appeals for the Second Circuit affirmed the district court's decision to admit the disputed evidence, noting that the admissibility of scientific evidence, like any other evidence, should be determined by weighing probativeness, materiality and reliability against

49. *See infra* notes 50-51 and accompanying text.

50. *See* Fed. R. Evid. 401.

51. Fed. R. Evid. 401, notes of advisory committee on Proposed Rules.

52. 583 F.2d 1194 (2d Cir. 1978), *cert. denied*, 439 U.S. 1117 (1979).

53. *See id.* at 1195-96.

54. Voice spectrograph, or voiceprint analysis, presumes that everyone's voice is different, and that a graphical representation of the voice produced by an electrical device known as a sound spectrograph is unique for any individual. *See* Annotation, *Admissibility and Weight of Voiceprint Evidence*, 97 A.L.R. 3d 294, 298 (1980); *Spectrogram Voice Identification*, 19 Am. Jur. Proof of Facts 423 (1967); *see also* Note, *The Voiceprint Dilemma: Should Voices Be Seen and Not Heard?*, 35 Md. L. Rev. 267, 271 (1975).

55. *See* United States v. Williams, 583 F.2d 1194, 1196 (2d Cir. 1978), *cert. denied*, 439 U.S. 1117 (1979).

56. *See id.*

any tendency to mislead, prejudice, or confuse the jury.[57] The circuit court's main concern was the reliability of voiceprint identification and its possible tendency to mislead.[58] In deciding the question of reliability, several indicia were considered, including a low error rate, the maintenance of standards, the care with which the technique had been employed, similarity to other tests, and "fail-safe" characteristics.[59] The court's concentration on these five factors, and its holding in favor of admissibility show that its decision turned upon a finding that the technique was, for legal purposes, likely enough to provide a correct identification,[60] and that the jury would not perceive a higher degree of reliability than in fact existed.[61]

Williams also illustrates how judicial concentration on reliability often reflects the failure of lawyers to question the reasoning of experts. Because the reasoning underlying the voice spectrograph comparison apparently was not contested, the court had to consider only whether a less than certain identification was probative enough to outweigh its potential for confusing or biasing the jury. Had the court considered the validity of the reasoning, it might have decided differently,[62] but on the evidence of the case, the decision related to the legal significance of a scientifically uncontested fact—namely the probability that the voice on the tape was that of the defendant. Indeed, the opinion explicitly notes that "courts cannot . . . surrender to scientists the responsibility for determining the reliability of . . . evidence."[63]

Evidence based on the science of epidemiology,[64] often at issue in toxic tort cases,[65] further illustrates how the distinction between the validity of

57. *See id.* at 1198.

58. *See id.*

59. *See id.* at 1198-99.

60. *Williams* was a case in which the *Frye* test was rejected, *see id.* at 1198, freeing the court to focus on reliability. Courts that apply *Frye* generally are less prone to admit voiceprint evidence than courts that have rejected *Frye*, but their analysis also tends to focus primarily on reliability issues. *See, e.g.*, United States v. Addison, 498 F.2d 741, 744-45 (D.C. Cir. 1974) (explicitly distinguishing reliability from acceptance, but in applying acceptance test, focusing on probable success rate and problems in determining error rate).

61. For a discussion of the issue of weighing the probative value of proferred evidence against its prejudicial effect, see United States v. Baller, 519 F.2d 463, 466 (4th Cir.) ("Unless an exaggerated popular opinion of the accuracy of a particular technique makes its use prejudicial or likely to mislead the jury, it is better to admit relevant scientific evidence in the same manner as other expert testimony and allow its weight to be attacked by cross-examination and refutation."), *cert. denied*, 423 U.S. 1019 (1975); *see also* Fed. R. Evid. 403.

62. *See infra* notes 270-82 and accompanying text.

63. United States v. Williams, 583 F.2d 1194, 1198 (2d Cir. 1978), *cert. denied*, 439 U.S. 1117 (1979).

64. "Epidemiology is concerned with the patterns of disease occurrence in human populations and the factors that influence these patterns." A. Lilienfeld & D. Lilienfeld, Foundations of Epidemiology 3 (2d ed. 1980).

65. A substantial body of legal literature on this topic has developed in recent years. *See* Black & Lilienfeld, *supra* note 13; Dore, *A Commentary on the Use of Epidemiological*

reasoning and the reliability of conclusions parallels a separation of legal and scientific issues. Epidemiologists study the patterns of disease occurrence in human population groups, and, from these patterns, infer relationships between diseases and factors suspected of causing them.[66] In the case of infectious diseases, a microorganism or virus responsible for all cases of a disease usually can be identified with great assurance.[67] For latent diseases like cancer, however, causal factors cannot be identified with similar precision and certainty, either in general or in specific cases.[68] This imprecise causal attribution has generated considerable legal debate,[69] but epidemiologic reasoning is rarely attacked. Legal choices about epidemiology usually do not involve scientific questions

---

Evidence in Demonstrating Cause-in-Fact, 7 Harv. Envtl. L. Rev. 429 (1983); Hall & Silbergeld, *Reappraising Epidemiology: A Response to Mr. Dore*, 7 Harv. Envtl. L. Rev. 441 (1983); Jacob, *Of Causation in Science and Law: Consequences of the Erosion of Safeguards*, 40 Bus. Law. 1229 (1985); McElveen & Eddy, *Cancer and Toxic Substances: The Problem of Causation and the Use of Epidemiology*, 33 Clev. St. L. Rev. 29 (1984); Novick, *Use of Epidemiological Studies to Prove Legal Causation: Aspirin and Reye's Syndrome, A Case in Point*, 22 Tort & Ins. L.J. 536 (1987); *see also In re* "Agent Orange" Prod. Liab. Litig., 611 F. Supp. 1223, 1239-40 (E.D.N.Y. 1985) (citing *In re* Swine Flu Immunization Prod. Liab. Litig., 508 F. Supp. 897, 907 (D. Colo. 1981), *aff'd sub nom.* Lima v. United States, 708 F.2d 502 (10th Cir. 1983), for the proposition that when the exact organic cause of a disease cannot be scientifically isolated, epidemiologic data becomes highly persuasive), *aff'd on other grounds*, 818 F.2d 187 (2d Cir. 1987).

66. *See* Black & Lilienfeld, *supra* note 13, at 750-51.

67. *See* S. Epstein, The Politics of Cancer 38-39 (1978).

68. *See id.* at 39.

69. A relatively recent student Note captures the flavor of this debate. *See* Note, *Causation in Toxic Torts: Burdens of Proof, Standards of Persuasion, and Statistical Evidence*, 96 Yale L.J. 376 (1986). The student author raises interesting philosophical questions about the use of probability estimates to resolve the issue of causation. The author distinguishes between "fact probability" and "belief probability," pointing out that just because statistical evidence establishes a certain level of probability for a population, it does not necessarily follow that one must conclude that this was the probability of causation in an individual case. *Id.* at 386-92. The Note asserts that it is inappropriate to apply population based probability estimates directly to the probability of causation in any individual case. *See id.* at 382-84.

The question that the Note raises relates to a choice that must be made between alternative philosophical concepts of what the Note terms "fact probability": the population frequency concept, and the chance or wagering concept. *See id.* at 383-84. Under the frequency view, the rate of occurrence of a disease in a population cannot be used as the basis for estimating the probability that an individual member of the population contracted the disease as a result of exposure to a suspected factor. Under the second view, however, the population frequency rate provides the best information to use if one were making a rational wager on whether or not an individual contracted the disease because of exposure. *See* I. Hacking, The Logic of Statistical Inference 1-12 (1965); *see also* Kaye, *The Laws of Probability and the Law of the Land*, 47 U. Chi. L. Rev. 34, 35 (1979) (pointing out that "skepticism about the application of probability theory to legal factfinding rests on fundamental misconceptions about the philosophical debate over the meaning of probability, the character of rational decisionmaking, and the values of the legal system").

For a more detailed discussion of "belief probability," see generally L. J. Cohen, The Probable and the Provable (1977). This is reviewed in Schum, *A Review of a Case Against Blaise Pascal and His Heirs*, 77 Mich. L. Rev. 446 (1979). *See also* Cohen, *On Analyzing the Standards of Forensic Evidence: A Reply to Schoeman*, 54 Phil. of Sci. 92 (1987);

about the validity of epidemiologic methods, but rather what to do with the probabilistic scientific answers that the methods provide.[70]

If all scientific evidence derived from noncontroversial, well-accepted reasoning, a pure reliability rule would suffice, and there would be no question about the validity aspect of relevancy. For controversial science that involves new theories or applications, however, reliability often depends on disputed reasoning. Ignoring validity under these circumstances makes it impossible for a court to determine if an opinion is based on science or merely reflects the personal views of an individual scientist.

Jones v. State,[71] a recent Texas case, shows how the validity of the reasoning underlying a conclusion, instead of the conclusion itself, can become the principal concern. In *Jones*, a nurse allegedly had killed one of her patients by injection with succinylcholine, a poison very difficult to detect.[72] To establish that this chemical had been the cause of death, tissue samples were sent to a research institute in Sweden, where scientists used special methods to extract the small quantities of poison in the samples.[73] A technique known as gas chromatography mass spectrometry[74] was then used to identify the drug.[75] When the defendant chal-

---

Schoeman, *Cohen on Inductive Probability and the Law of Evidence*, 54 Phil. of Sci. 76 (1987).

For other views on the appropriate use of epidemiologic evidence, see Black & Lilienfeld, *supra* note 13, at 764-66 (advocates combining epidemiologic principles with the legal burden of proof); Delgado, *Beyond* Sindell: *Relaxation of Cause-In-Fact Rules for Indeterminate Plaintiffs*, 70 Cal. L. Rev. 881, 899-902 (1982) (advocating proportional recovery and reversal of burden of proof); Farber, *Toxic Causation*, 71 Minn. L. Rev. 1219, 1243-51 (1987) (advocating a scheme in which defendant pays total damages based on population analysis, but money is distributed only to plaintiffs "most likely" injured by defendant's conduct); Robinson, *Multiple Causation in Tort Law: Reflections on the* DES *Cases*, 68 Va. L. Rev. 713, 755 (1982) (advocating apportionment scheme); Note, *Establishing Causation in Chemical Exposure Cases: The Precursor Symptoms Theory*, 35 Rutgers L. Rev. 163, 191 (1982) (advocating use of epidemiologic evidence that a chemical causes precursor symptoms); Note, *The Inapplicability of Traditional Tort Analysis to Environmental Risks: The Example of Toxic Waste Pollution Victim Compensation*, 35 Stan. L. Rev. 575, 614 (1983) (advocating a probabilistic approach to causation and the use of rebuttable presumptions); Note, *Tort Actions for Cancer: Deterrence, Compensation, and Environmental Carcinogenesis*, 90 Yale L.J. 840, 855-59 (1981) (advocating scheme based on determining what exposure levels double risk); *cf.* Huber, *supra* note 3, at 278 (arguing that courts are ill-equipped to deal with "public risks").

70. *Cf.* Sulesky v. United States, 545 F. Supp. 426, 430-31 (S.D.W. Va. 1982) (court declined to consider epidemiologic evidence in Swine Flu case, relying instead on testimony of treating and evaluating physicians).

71. 716 S.W.2d 142 (Tex. Ct. App. 1986).

72. *See id.* at 144.

73. *See id.* at 145.

74. Gas chromatography is a method for the separation and analysis of complex mixtures of volatile organic and inorganic compounds. It is based on the fact that different components move at different rates through a column containing a solvent liquid or solid. *See* 6 McGraw-Hill Encyclopedia of Science and Technology 61-64 (5th ed. 1982). After separation of the mixture with gas chromatography, each component can be analyzed further with a mass spectrometer, which determines molecular structure based on the fragmentation of ions formed when the molecules are ionized. *See* 8 McGraw-Hill Encyclopedia of Science and Technology 234-41 (5th ed. 1982).

lenged this procedure, the state's expert explained that he had confidence in it primarily because the principles upon which it was based were well-established.[76] The reliability of his findings, therefore, derived almost solely from the validity of the reasoning behind those principles and the expert's reasoning from them.[77] As the *Jones* decision illustrates, however, it is difficult to pinpoint the validity issue when applying the traditional general acceptance analysis.[78]

To accommodate cases such as *Jones*, this Article proposes a clarification of Rule 702 of the Federal Rules of Evidence to require explicitly that testimony be based on valid scientific reasoning *and* that it be reliable enough to satisfy threshold legal requirements for admissibility. As modified, Rule 702 would read:

> If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise. When the witness offers testimony based on scientific knowledge, such testimony shall be admitted only if the court determines that the opinion:
>
> 1) is based on scientifically valid reasoning; and
>
> 2) is sufficiently reliable that its probative value outweighs the dangers specified in Rule 403.[79]

The two steps of the proposed clarification of Rule 702 constitute independent requirements,[80] one essentially based on scientific considera-

---

75. *See Jones*, 716 S.W.2d at 145.

76. *See id.* at 148-50.

77. *Id.* at 148. For a case that involved a fact pattern similar to that in *Jones*, see Coppolino v. State, 223 So. 2d 68, 69 (Fla. Dist. Ct. App. 1968) (expert combined new and old tests and procedures to develop method for detecting traces of succinylcholine chloride in exhumed body), *appeal dismissed*, 234 So. 2d 120 (1969), *cert. denied*, 399 U.S. 927 (1970).

78. The *Jones* court noted that while scientific and legal periodicals sometimes are used to establish general acceptence, this inquiry does not lead to a consideration of validity. *See Jones*, 716 S.W.2d at 147.

79. Currently, Rule 702 provides: "If scientific, technical, or other specialized knowledge will assist the trier of fact to understand the evidence or to determine a fact in issue, a witness qualified as an expert by knowledge, skill, experience, training, or education, may testify thereto in the form of an opinion or otherwise." Fed. R. Evid. 702.

80. Other commentators have discussed the use of a reliability standard, or a validity standard, but none has presented validity and reliability as two complementary aspects of relevancy. Professor Giannelli views reliability as the principal test and sees the validity of the underlying principle and of the technique applying the principle as issues subsidiary to reliability, *see* Giannelli, *supra* note 4, at 1200-01, an approach somewhat similar to that taken in this Article. He is not very clear, however, about the meaning of validity. He notes in a footnote that although courts use validity and reliability interchangeably,

> the terms have distinct meanings in scientific jargon. "Validity" refers to the ability of a test procedure to measure what it is supposed to measure—its accuracy. "Reliability" refers to whether the same results are obtained in each instance in which the test is performed—its consistency. Validity includes reliability, but the converse is not necessarily true.

tions, the other on legal considerations.[81] Nonetheless, validity and reliability are closely related. Invalid reasoning cannot lead to reliable

---

Giannelli, *supra* note 4, at 1201 n.20.

This definition of validity, however, is quite confusing, for it seems to reverse the relative roles of validity and reliability that Professor Giannelli sets out in the main text. What is more, equating validity with accuracy is simply at variance with the normal usage of the word "validity." Professor Giannelli himself does not really use validity in this way. In discussing voice spectrograph comparisons, he notes that "[i]f the theory of voice uniqueness is not valid, voiceprint evidence is not reliable." *Id.* at 1202. If "accurate" were substituted for "valid" in this sentence, it would not make sense, because one does not properly refer to a theory as accurate. Still later in his article, Professor Giannelli maintains that the reliability of a technique can be established through empirical validation, which indicates still another meaning for validity. *See id.* at 1213. For a very different view of the relationship between accuracy, reliability, and validity, see Epstein, *The Risks of Risk/Utility*, 48 Ohio St. L.J. 469, 470 (1987) ("Validity and reliability stand in inverse relation to each other. Validity seeks to identify the ultimate thing to be measured, while reliability seeks to insure that the measurements undertaken are done accurately.").

A 1974 article on psychiatric evidence also distinguishes reliability from validity. *See* Ennis & Litwack, *Psychiatry and the Presumption of Expertise: Flipping Coins in the Courtroom*, 62 Cal. L. Rev. 693 (1974). Like Professor Giannelli, Ennis and Litwack define validity in terms of accuracy. *See id.* at 697. They take reliability to mean "the probability or frequency of agreement when two or more independent observers answer the same question." *Id.* Once again, equating validity with accuracy simply does not accord with the usual usage of the word "validity." Nonetheless, Ennis and Litwack make a distinction not unlike the distiction made in this Article. If they were to reverse their usage of reliability and validity, their formulation would approximate this Article's formulation, except that they do not tie the two concepts together as complementary aspects of relevancy. *See also* Skolnick, *Scientific Theory and Scientific Evidence: An Analysis of Lie-Detection*, 70 Yale L.J. 694, 694 (1961) (validity of a technique relates to whether it actually tests what it claims to test).

Perhaps the most revealing indication of how the law has failed to grasp the interrelationship between reliability, validity, and relevancy is provided by a symposium that appeared recently in Federal Rules Decisions. *See Rules for Admissibility of Scientific Evidence*, 115 F.R.D. 79 (1987) (symposium report). This report contains four proposals for a model rule on admissibility of scientific evidence. One of these, advanced by Professor Giannelli, is more procedural than evidentiary. He proposes that expert testimony be inadmissible "unless the proponent gives the adverse party sufficient advance written notice of intent to use such evidence." *Id.* at 102. The other three proposals, however, each go off in different substantive directions. Professor Berger focuses on relevancy, *see id.* at 89, Professor Starrs focuses on validity, *see id.* at 92, and Professor Lederer focuses on reliability. *See id.* at 84. The clarification to Rule 702 proposed in this Article, *see supra* text accompanying note 79, essentially combines the proposals of Professors Lederer, Berger and Starrs. This merging of their ideas is possible because all three proposals fit within the analytical framework advocated herein.

81. On how scientific factual issues can be separated from policy or legal issues, see Martin, *The Proposed "Science Court"*, 75 Mich. L. Rev. 1058, 1078 (1977). While agreeing with Professor Martin's premise, this Article does not agree that a separate science court is required. *See also* Harter & Schlenker, *In Quest of Agreement: Non-Adversarial Processes for Making Scientific Judgments Concerning Public Health*, Monograph No. 102, Institute for Health Policy Analysis, Georgetown University Medical Center, at 6 (1988) (noting that "[t]he clarity of the science will do nothing to resolve the difficult political decision as to whether [a given level] of risk is acceptable"); Steinbock, Richman & Ray, *Expert Testimony on Proximate Cause*, 41 Vand. L. Rev. 261, 281-82 (1988) (arguing that experts should not be allowed to testify about the legal question of proximate cause, and that such testimony should be excluded under Rule 702).

results,[82] which, in some cases, means that the legal inquiry need go no further than the first step. In other cases, validity may be established by reference to a history of success in applying a theory or method. For example, the fact that virtually all bridges succeed in carrying the loads for which they were designed provides persuasive evidence for the validity of the principles of structural engineering and the reasoning of structural engineers.

Reliability, in other words, can be a very good indicator of validity, and when it is, the legal analysis of the two concepts may collapse into one step. Indeed, to the extent that courts rely on a history of success in reaching decisions about the admissibility of scientific evidence, they merely incorporate the traditional view of the scientific method into the law.[83] If, as once thought, all science proceeded in this fashion, there would be no need to distinguish validity from reliability, but in actual scientific practice, decisions about validity are not based purely on confirmation.

C. *Acceptance as the Scientific and Legal Basis for Determining Validity*

Unfortunately, the traditional view of the scientific method implies an exactness and certainty that simply cannot exist. Over the last twenty-five years, philosophers of science have abandoned tradition for a more realistic account of the scientific enterprise,[84] but the law persists in the

82. This point has been made by Professor Imwinkelried, who has advocated judicial screening of techniques and theories to establish their validity as a predicate to admissibility. *See* Imwinkelried I, *supra* note 16, at 581. He views theories in the sense of broad, general propositions, however, and validity in terms of accuracy. *See id.* at 600-01. He assumes an essentially positivist version of the scientific method, *see id.* at 101, which does not accurately reflect actual scientific practice. *See infra* Part I.C.1 (in-depth discussion of Logical Positivism).

83. *See infra* Part I.C.

84. To understand what makes science scientific, one must turn to the philosophy of science. Scientists themselves generally have been remarkably uninterested in what warrants their work as scientific. *See, e.g.*, T. Kuhn, The Structure of Scientific Revolutions 47 (2d ed. 1970) ("Though many scientists talk easily and well about the particular individual hypotheses that underlie a concrete piece of current research, they are little better than laymen at characterizing the established bases of their field, its legitimate problems and methods."); MacLure, *Popperian Refutation in Epidemiology*, 121 Am. J. Epidem. 343, 343 (1985) ("Practically minded epidemiologists may tend to believe that philosophy of science is irrelevant, abstract and dull."); McMullin, *Alternative Approaches to the Philosophy of Science*, in Scientific Knowledge 6 (J. Kourany ed. 1987) ("[W]hen contemporary scientists set out to give an account of the nature of the scientific method, they can sometimes be as remote from scientific practice as were Aristotle or Descartes."); Ziman, *What is Science?*, in Introductory Readings in the Philosophy of Science 35, 38 (E.D. Klemke, R. Hollinger, & A.D. Kline eds. 1980) (the "Rule of Order" that most scientists would state is not always obeyed in practice).

Philosophers, on the other hand, have long had an interest in science as the quintessential example of epistemic (knowledge-generating) activity. *See* S. Toulmin, Foresight and Understanding 13-14 (1961) ("[C]ertain general questions about science remain which, though not immediately urgent for the working scientist, are nonetheless worth asking. In dealing with these questions . . . the scientist must co-operate with the historian and

old view.[85] Understanding this misperception, and its implications for the law,[86] requires a brief review of how the philosophy of science has

with the philosopher."); Suppe, *Afterword* to The Structure of Scientific Theories 716 (F. Suppe ed. 2d ed. 1977) ("Over much of the history of philosophy a central aim of epistemology has been to vindicate the epistemic claims of then contemporary science; such were principal motives of . . . the Positivists for developing [their view of science].").

85. *See* Lansing, *The Motherless Calf, Aborted Cow Theory of Cause*, 15 Env. Law 1, 9 (1984) (advancing a proposal that would counter "the harshness of plaintiff's having to prove a true scientific cause-in-fact"); Nesson, *supra* note 27, at 523 (adopting the position that "the legal standards of proof must be more liberal than scientific standards because courts must seek justice as well as ultimate, unassailable, scientific truth"); Comment, *Scientific Evidence — Admissibility* Fryed *to a Crisp*, 21 S. Texas L.J. 62 (1981) (viewing scientific method as providing complete objectivity); Note, *Of Reliable Science: Scientific Peer Review, Federal Regulatory Agencies, and the Courts*, 7 Va. J. of Nat. Res. L. 27, 30-31 (1987) (outlining five-step "method" that reflects Logical Positivist view); *cf.* Brennan & Carter, *Legal and Scientific Probability of Causation of Cancer and Other Environmental Disease in Individuals*, 10 J. Health Politics, Policy & Law 33, 35 (1985) (noting that most discussions of the relationship between legal and scientific causation usually center on "platitudes about lawyers and scientists thinking differently"); Danner & Sagall, *Medicolegal Causation: A Source of Professional Misunderstanding*, 3 Am. J. Law & Med. 303, 303 (1977) (arguing that judges and attorneys view causation quite differently than do members of the medical community); Goldberg, *supra* note 31, at 1342-43 (noting that most lawyers think of science in terms of "progress", a very positivist view, but pointing out that there "are more than a few objections to this formulation"); Large & Michie, *Proving that the Strength of the British Navy Depends on the Number of Old Maids in England: A Comparison of Scientific Proof With Legal Proof*, 11 Envtl. L. 555 (1981) (essentially contrasting the Logical Positivist view of science with the law). In their article, Danner and Sagall focus on the difference between what physicians usually do and what they are asked in court. This, however, begs the question of how a physician can answer the questions posed by the law in a scientifically appropriate way. The article distinguishes scientific proof from the legal standard of "more likely than not." Danner & Sagall, *supra*, at 307. The real question is how one proves scientifically that something, more likely than not, is true.

86. This misperception of science has been noted by others. *See* Brennan & Carter, *supra* note 85. Brennan and Carter see judges as continuing to view science as requiring mechanistic explanation, *see id.* at 35, 53, despite the fact that neither Mechanistic Materialism nor Logical Positivism is now recognized as an accurate picture of how science actually is conducted. *See id.* at 38-39. In this regard, they agree with the view advocated in this Article, but they do not examine how science determines validity in the absence of Positivist confirmation. Also, they assume that courts have proven totally incapable of dealing with probabilistic population-based evidence. *See id.* at 54. In fact, this is not the case.

As early as 1919, the New York Court of Appeals upheld, without any difficulty, a plaintiff's verdict based on population-based evidence. In Stubbs v. City of Rochester, 226 N.Y. 516, 124 N.E. 137 (1919), the defendant city inadvertently had contaminated a portion of its water supply system. *See id.* at 518-19, 124 N.E. at 137. Though there was a typhoid outbreak in all parts of the city, the incidence rate in the contaminated part was much higher than elsewhere. *See id.* at 525, 124 N.E. at 140. Based on this probabilistic, population-based type of evidence, the plaintiff, who lived in the contaminated area and had contracted typhoid, successfully claimed against the city. *See id.* at 527, 124 N.E. at 139-40. For a discussion of other precedents favoring the use of epidemiologic evidence, see Black & Lilienfeld, *supra* note 13, at 769-76. *But see* Johnston v. United States, 597 F. Supp. 374, 412 (D. Kan. 1984) (rejecting probabilistic evidence), *cert. denied*, 108 S. Ct. 694 (1988).

The real point Brennan and Carter seem to be making is that courts should accept the *non*-probabilistic, qualitative opinions often offered by plaintiffs in toxic tort cases. They start out arguing for probability, *see* Brennan & Carter, *supra* note 85, at 34-35, but in

developed since the turn of the century. The current philosophical view of science further demonstrates how the law can evaluate the validity of scientific reasoning without detailed or specific knowledge of science. Stripping science to its philosophical core reveals that validity derives, not from applying pure logic to purely objective facts, but from the give and take process of testing, experimentation, criticism, and review.

1. The Traditional View of Science

The misconceptions about the scientific method that permeate legal thinking reflect a once deeply entrenched philosophical tradition known as Logical Positivism.[87] The philosophers who developed this tradition sought to sustain a vision of science as the source of absolute truth. Until the early 1900's, most scientists thought in terms of a purely objective world—independent of those who perceive it and governed by mechanical laws that allow precise descriptions and predictions.[88] They believed that if a person knew the positions and velocities of all particles in the universe at a given instant, he or she could predict all future events and perfectly reconstruct all past occurrences.[89] At the heart of their belief

---

effect advocate "trans-science." *See* Note, *Trans-Science in Torts*, 96 Yale L.J. 428 (1986) (discussed *infra* notes 506-08 and accompanying text).

87. *See generally* The Legacy of Logical Positivism (P. Achinstein & S. Barker eds. 1969); Suppe, *The Search for Philosophic Understanding of Scientific Theories*, in The Structure of Scientific Theories 3, 6-62 (F. Suppe ed. 2d ed. 1977). The name for this philosophy derives from the fact that it represents a combination of the rigorous mathematical *logic* of Bertrand Russell and the *Positivist* or Empirical philosophy of Ernst Mach. *See* Toulmin, *The Structure of Scientific Theories*, in The Structure of Scientific Theories 600, 600 (F. Suppe ed. 2d ed. 1977).

The Logical Positivist picture of science is best explained in terms of an example of a theory formulated in accordance with it. Consider the theory that relative density determines whether or not a solid object will float in a liquid, an example taken from C. Hempel, *The Theoretician's Dilemma*, in Aspects of Scientific Explanation and Other Essays in the Philosophy of Science 180-82 (1965). Observations of the floating properties of various objects placed in a liquid, such as water, reveal general patterns such as the fact that wood floats and iron does not. This does not account, however, for the observation that a properly made hollow iron sphere also floats. Dealing with such an anomaly forces the development of a theory, which might be described as follows. First, a theoretical term, density, is defined as the quotient of two observation terms, weight and volume. With this theoretical term, a theoretical statement, or postulate can be formulated, namely that if the density of a solid body is less than the density of a liquid, the body will float on the liquid. From this it can be inferred, or predicted, for any specific combination of solid and liquid whether or not the body will float. Test procedures for specific combinations are given by correspondence to rules similar to the rules that define the theoretical terms. *See id.* More complex theories would involve several interrelated theoretical statements, but the basic idea remains the same. *See* Feigl, *The Origin and Spirit of Logical Positivism*, in The Legacy of Logical Positivism 3, 15-16 (P. Achinstein & S. Barker eds. 1969) (explaining how Logical Positivists axiomatized theories).

88. *See* Suppe, *The Search for Philosophic Understanding of Scientific Theories*, in The Structure of Scientific Theories 3, 8 (F. Suppe ed. 2d ed. 1977).

89. *See* Salmon, *Why Ask, "Why?"? An Inquiry Concerning Scientific Explanation*, in Scientific Knowledge—Basic Issues in the Philosophy of Science 51, 51 (J. Kourany ed. 1987). An excellent description of the world view of 18th and 19th century scientists is provided in D. Bohm, Causality and Change in Modern Physics 36 (1957):

was the assumption that all phenomena could be explained in terms of a few fundamental and absolute laws. This mechanistic and self-contained view foreclosed concern about reasoning or the nature of abstract theoretical constructs.[90]

Around the turn of the twentieth century, however, advances in physiology and psychology and the advent of the quantum and relativity theories in physics destroyed simple, mechanistic certainty.[91] Quantum theory tells us that certainty is a physical impossibility, relativity that time is not absolute, and psychology that preconceptions color supposedly objective accounts of the natural world. The Logical Positivists strove to rescue this situation by divorcing theories from the world of empirical observation and examining only their logical structure.[92] Theories explained observational phenomena,[93] but supposedly existed apart

---

> The very precision of Newton's laws [of motion] led . . . to new problems of a philosophical order. For, as these laws were found to be verified in wider and wider domains, the idea tended to grow that they have a *universal* validity. Laplace, during the eighteenth century, was one of the first scientists to draw the full logical consequences of such an assumption. Laplace supposed that the *entire universe* consisted of nothing but bodies undergoing motions through space, motions which obeyed Newton's laws. While the forces acting between these bodies were not yet completely and accurately known in all cases, he also supposed that eventually these forces could be known with the aid of suitable experiments. This meant that once the positions and velocities of all the bodies were given at any instant of time, the future behaviour of everything in the whole universe would be determined for all time.

Id. (emphasis in original).

90. Mechanistic Materialists believed that:

> The scientific method yields immediate and objective knowledge . . . and is capable of doing so by empirical investigation without any recourse to philosophical speculation. Thus there is no place for a priori elements in natural science or in empirical knowledge. Observation of the world is immediate in the sense that no a priori or conceptual mediation is involved in obtaining observational knowledge; observation in accordance with the procedures of natural science is sufficient to yield knowledge of the world's mechanistic nature.

Suppe, *The Search for Philosophic Understanding of Scientific Theories*, in The Structure of Scientific Theories 3, 8 (F. Suppe ed. 2d ed. 1977).

91. *See id.* (noting that "work . . . on the physiology of the senses indicated that an adequate philosophy must make provision for the activity of the thinking subject in the growth of scientific knowledge") (citation omitted). For a nontechnical account of how relativity and quantum mechanics made it clear that events cannot be known with complete precision, see J. Gribbin, In Search of Schrodinger's Cat (1984).

92. A very readable account of the Logical Positivist view by one of the founders of Logical Positivism is provided in Feigl, *The Origin and Spirit of Logical Positivism*, in The Legacy of Logical Positivism 3 (P. Achinstein & S. Barker eds. 1969); *see also supra* note 87.

93. As Logical Positivism developed, its adherents came to view a theory's explanatory power as a prerequisite to its validity. Three forms of explanation were recognized: deductive-nomological (D-N); deductive-statistical (D-S); and inductive-statistical (I-S). *See* C. Hempel, *supra* note 87, at 335-47, 376-93. The density comparison theory described in note 87, *supra*, illustrates D-N explanation. A conclusion about any solid-liquid combination could be rigorously deduced from the theory, and the predicted result (floating or sinking) would be expected with law-like certainty. Prediction of the result of tossing a coin illustrates D-S explanation. One can deduce that a given result will occur with a given probability based upon knowledge about what a coin is like. One need not

from them.[94]

By requiring all theories to conform to a single, logical pattern,[95] the Positivists focused on only one of several ways in which scientists and nonscientists use the word "theory." Sometimes theory refers to a broad field of established knowledge, like the theory of electricity, sometimes to a mere conjecture or hypothesis. For scientists and for purposes of this Article, however, a theory is best understood as an intellectual framework or device that allows one to describe and understand natural phenomena.[96] In this sense, theories incorporate the patterns of reasoning that tie every valid scientific conclusion to empirical evidence and to recognized principles.

The Positivists, however, made the reasoning leading to a theory irrelevant in determining the theory's validity. Once put into the standard format, a theory's validity depended only on whether consequences deduced from it could be observed.[97] Known as the hypothetico-deductive

---

collect data beforehand to make this prediction. Predicting disease rates, on the other hand, requires that data be determined in order to calculate probability. This illustrates I-S explanation. It should be noted that even the Logical Positivist view did not require that science always lead to absolute and certain explanations, though the infusion of probability was an afterthought forced upon the philosophy. *See* Brennan & Carter, *supra* note 85, at 38. It should also be noted that Positivists did not distinguish between explanation and prediction. *See* Suppe, *The Search for Philosophic Understanding of Scientific Theories*, in The Structure of Scientific Theories 3, 28 (F. Suppe ed. 2d ed. 1977).

94. Logical Positivists sometimes offered a metaphorical picture according to which a theory's system of postulates hovers over the plane of empirical reality, acquiring meaning only through correspondence rules that tether it to empirical observations. *See* Feigl, *The "Orthodox" View of Theories: Remarks in Defense as Well as Critique*, in IV Minnesota Studies in the Philosophy of Science—Analyses of Theories and Methods of Physics and Psychology 3, 5-6 (M. Radner & S. Winokur eds. 1970).

95. Logical Positivists recognized that theories initially might not be in the proper form, but they thought that a valid theory had to be such that it could be "rationally reconstructed" to make its logical structure and empirical content as clear as possible. *See* P. Achinstein, Concepts of Science 68 (1968). This procedure supposedly provided an understanding of the structure of theories and the logic of scientific explanation, as well as the method for assessing claims of scientific knowledge. "With the addition of appropriate information about particular theories and experiments, such rational reconstructions [were] supposed to enable one to carry through an analysis of the kinds of claims particular theories make about the world and to evaluate the worth of such claims." Burian, *Scientific Realism and Incommensurability: Some Criticisms of Kuhn and Feyerabend*, in Methodology, Metaphysics and the History of Science 1, 2 (R. Cohen & M. Wartofsky eds. 1984).

Logical Positivists thought that science develops through a process of accumulation in which prior facts and theories are never discarded but, instead, are incorporated into later theories as special cases. *See* D. Shapere, Reason and the Search for Knowledge 60 (1984). This absolutist version of science offers the appeal of logical certainty, but it has fallen from favor because many well-recognized developments and advances in science do not conform to its strict rules. *See* Doppelt, *Relativism and Recent Pragmatic Conceptions of Scientific Rationality*, in Scientific Explanation and Understanding 107, 107 (N. Rescher ed. 1983).

96. *See* Suppe, *Afterword* to The Structure of Scientific Theories 615, 658 (F. Suppe ed. 2d ed. 1977).

97. *See* Feigl, *The Origin and Spirit of Logical Positivism*, in The Legacy of Logical Positivism 3, 5 (P. Achinstein & S. Barker eds. 1969) (discussing rejection of metaphysics

method,[98] this version of science implies that a theory, once verified, becomes an indelible part of our knowledge, though it eventually may become incorporated into a broader, more inclusive theory.[99] At the same time, however, hypothetico-deduction, if strictly applied, invalidates a theory entirely if a single prediction proves false.[100] While such a result is perhaps appropriate in some cases, neither prong of this rigid and self-contradictory view could survive as a universal description of either science or scientific theories.[101]

2. Why the Traditional View of Science Is Incorrect

The idealized Logical Positivist recipe for all of science fails because it entails two unsupportable bright-line distinctions. The first has to do with the supposed absolute distinction between observation and theory. According to Logical Positivism, observations can be described in a completely theory-neutral "observation language."[102] Logical Positivists would have it that red is red, no matter what theory of spectroscopy one subscribes to, but this misses the fact that preconceptions influence all observations.[103] The second bright line concerns the origin of theories.

---

by the Logical Positivists and their requirement that meaning derive from observable consequences of a theory); *see also* Suppe, *The Search for Philosophic Understanding of Scientific Theories*, in The Structure of Scientific Theories 3, 11-12 (F. Suppe ed. 2d ed. 1977).

98. Salmon, *Inquiries Into the Foundations of Science*, in Probabilities, Problems, and Paradoxes 139, 141-42 (S. Luckenbach ed. 1972).

99. *See supra* note 95.

100. *See* Feyerabend, *Against Method: Outline of An Anarchistic Theory of Knowledge*, in IV Minnesota Studies in the Philosophy of Science—Analyses of Theories and Methods of Physics and Psychology 17 (M. Radner & S. Winokur eds. 1970). Feyerabend, one of the most radical critics of positivism, shows that a logical consequence of the positivist view is that a theory must agree completely with experience, *see id.* at 42, which effectively makes all theories quite useless. Feyerabend bases his argument on a strict interpretation of Logical Positivism. According to him, the slightest change in a theory to accommodate an observational anomaly means that an entirely new theory has been developed, and that the old theory is completely invalidated. *See id.* at 42-43. In practice, of course, an old theory (like Newtonian mechanics) still may be used for many applications even after being shown to be invalid for other applications for which a new theory (for example, Einstein's theory of relativity) is valid.

101. This rigidly structured process does resemble some areas of science, but even as to these, it fails to account fully for the way scientists reason. For example, the epidemiologic method for determining the cause of a disease depends heavily on statistical reasoning, but not to the exclusion of other considerations. Even in this rigorously statistical branch of medicine, the reasoning cannot be reduced to a logical algorithm. Instead, a number of postulates must be satisfied. *See* A. Lilienfeld & D. Lilienfeld, *supra* note 64, at 13 (epidemiologists use a two-stage sequence of reasoning that involves "[t]he determination of a statistical association between a characteristic and a disease," and then the drawing "of biological inferences from such a pattern of statistical associations"); Black & Lilienfeld, *supra* note 13, at 750-64 (explaining epidemiologic principles); Evans, *Causation and Disease: The Henle-Koch Postulates Revisited*, 49 Yale J. Biology & Med. 175, 192 (1976) (listing ten criteria).

102. Suppe, *The Search for Philosophic Understanding of Scientific Theories*, in The Structure of Scientific Theories 3, 45-49 (F. Suppe ed. 2d ed. 1977).

103. *See supra* note 91 and accompanying text.

The reasoning process leading to a theory supposedly had nothing to do with the validity of that theory. Only confirmation in accordance with the axiomatic logical format could make the theory valid.[104]

The complete separation of theory from observation fails because it does not account for the interaction that necessarily takes place between the two.[105] Theories serve as tools for eliminating confusion about the empirical world, but they shape perception as well.[106] For example, two observers waiting by a highway at night may have very different theoretical ideas about what they will see. One might expect a UFO, and the other an automobile. When two lights appear in the distance, each observer will see something different. Observations, however, also shape theories. As the UFO/automobile draws closer, the observer who believes in extraterrestrial visitors will have to reshape his or her thinking to accommodate the empirical evidence. At some point, he or she will have to abandon the theory of UFOs, or at least modify it so that it allows for the fact that cars, as well as flying saucers, move along highways. Based on the evidence, a scientist would conclude that the car theory is far better supported and far more valid. Science thus proceeds by weaving observation and theory together, not by viewing them as parallel but separate threads.[107]

Even if theory and observation were separable, the rigid logic of hypothetico-deductive confirmation suffers from intrinsic limitations. While hypothetico-deductivism seems quite straightforward, it does not allow for the fact that the same data can confirm an infinite number of

104. *See* Nickles, *Heuristics and Justification in Scientific Research: Comments on Shapere*, in The Structure of Scientific Theories 571, 575 (F. Suppe ed. 2d ed. 1977). Karl Popper, a well-known early critic of Positivism, rejected the idea that any theory could be verified through a process of confirmation. *See* K. Popper, The Logic of Scientific Discovery 40-42 (1959). Instead, he took the view that scientific theories had to be in falsifiable form, *see id.*, and that if a theory withstood many attempts at falsification, it would be well "corroborated." *Id.* at 251-81. Like other philosophers of the Positivist period, however, Popper thought that he could make "explicit the rules of evidential inference which scientists use implicitly for making choices of theory." Laudan, *Two Puzzles About Science: Reflections On Some Crises in the Philosophy and Sociology of Science*, 20 Minerva 253, 256 (1982).

105. *See* D. Shapere, *supra* note 95, at 181 ("In many ways, scientific method is more like military strategy than it is like the rules of chess: the strategy shapes the course of the campaign, but is itself responsive to the lay of the land . . . and it adjusts to new situations and new devices."); Achinstein, *The Problem of Theoretical Terms*, in Readings in the Philosophy of Science 234 (B. Brody ed. 1970) (noting that the division of expressions used by scientists into theoretical and observational sets depends on the particular theory in question and on the context of the observation).

106. *See* Buckhout, *Eyewitness Testimony*, 231 Sci. Am., Dec. 1974, at 23, 28 (noting that "a theory can be a powerful tool for clarifying confusion, but it can also lead to distortion and unreliability if people attempt, perhaps unconsciously, to make fact fit theory").

107. By the mid 1960's, it was established clearly that no reasonable criteria for complete separation of theoretical from non-theoretical terms could work. Proposed criteria were all shown to be "tied, in most cases rather specifically, to a particular context of observation or to a particular theory." Achinstein, *The Problem of Theoretical Terms*, in Readings in the Philosophy of Science 234, 247 (B. Brody ed. 1970).

hypotheses.[108] The hypothesis that a dog pulls the sun across the sky each day is consistent with every sunrise and sunset. In fact, contemporary astronomical observations confirmed the Ptolemaic geocentric theory of the universe[109] at least as well as they confirmed the heliocentric theory[110] of Copernicus.[111] Galileo's choice of the Copernican view, therefore, could not have derived solely from confirmation—other considerations must have come into play.[112]

In practice, claims that purport to be scientific often are rejected without any resort to hypothetico-deductive testing, and often are accepted as at least worthy of further development despite hypothetico-deductive disconfirmation.[113] Astrological predictions, for example, are rejected as unscientific without empirical tests. A scientist who undertakes such tests "would, at the very least, be deemed by the scientific community to be eccentric and exhibiting questionable judgment."[114] At the other extreme, physicists continue to pursue and develop theories that predict the existence of esoteric particles despite repeated failures to find them.[115] Logical Positivism fails as the basis for judging both astrology and modern particle theory because it does not describe or explain how scientists

108. As explained in R. Pirsig's best-selling book:

> If the purpose of scientific method is to select from among a multitude of hypotheses, and if the number of hypotheses grows faster than experimental method can handle, then it is clear that all hypotheses can never be tested. If all hypotheses cannot be tested, then the results of any experiment are inconclusive and the entire scientific method falls short of its goal of establishing proven knowledge.

R. Pirsig, Zen and the Art of Motorcycle Maintenance 115 (1974); *see also* Salmon, *Inquiries Into the Foundations of Science*, in Probabilities, Problems, and Paradoxes 139, 142-43 (S. Luckenback ed. 1972).

109. The geocentric theory assumes the earth is the center of the solar system, and indeed, of the universe. *See* McGraw-Hill Dictionary of Scientific and Technical Terms 676 (3d ed. 1984).

110. The heliocentric theory maintains that the sun is the center of the solar system. *See id.* at 737.

111. *See* T. Kuhn, *supra* note 84, at 75-76.

112. *See* Drake, *Ptolemy, Galileo, and Scientific Method*, 9 Stud. in Hist. & Phil. of Sci. 99, 99 (1978). Galileo wrote about a method not very different from hypothetico-deductivism, *see id.* at 103, but he persisted in trying to justify Copernicus' view despite a lack of evidence. The fact that the heliocentric view allowed for a mechanical explanation of tidal action apparently was a factor. *See id.* at 115.

113. *See* Laudan, *A Problem-Solving Approach to Scientific Progress*, in Scientific Revolutions 144, 144 (I. Hacking ed. 1981). Indeed, for some sciences, rigorous, prospective hypothetico-deductive testing is virtually impossible. Geologists, for example, cannot do experiments to confirm directly that processes of long duration occurred in accordance with their theories.

114. Weyant, *Protoscience, Pseudoscience, Metaphors and Animal Magnetism*, in Science, Pseudo-Science and Society 77, 80 (1980).

115. J. Gribbin, *supra* note 91, at 270 (discussing "supergravity" theory, which precisely predicts the existence of 162 different kinds of particles, some of which have never been detected, and noting that the "problems are immense, but theories like supergravity are at least consistent, finite, and not in need of renormalization. There is a *feeling in the wind* that physicists are on the right track.") (emphasis added).

actually do their work. As a result, it cannot provide universal criteria for evaluating claims of knowledge supposedly based on science.

The failure to account for all of science does not mean that Positivism is completely unrelated to scientific practice. As already noted, epidemiology closely resembles the Positivist view of science.[116] Nor does the failure mean that universal criteria cannot be formulated. Some philosophers, however, have taken the extreme position that there can be *no* criteria for judging science and that all judgment is subjective and relative.[117] Paul Feyerabend, perhaps the most extreme of these "Relativists,"[118] has taken the position that anarchy is the best foundation for both epistemology[119] and the philosophy of science.[120] He maintains that if one looks at the history of science, the only rule is that "anything goes,"[121] though he freely admits that this eliminates any rational basis for distinguishing even witchcraft from science.[122] According to another version of Relativism, science normally does proceed within the confines of accepted "paradigms," but when a new paradigm (such as Einstein's theory of relativity) replaces an old one (such as Newtonian mechanics), criteria change as well.[123]

3. The Legal Implications of the Current View of Science

To what, then, should science, and the law, turn if Logical Positivism is dead and if Relativism provides no independent universal standard for deciding whether a scientific claim or theory is valid? In the actual practice of science, "hypotheses are [in fact] confronted with the test of expe-

---

116. *See supra* note 101.

117. *See, e.g.*, T. Kuhn, *supra* note 84, at 4 (noting that when different schools of science compete, the distinction between them is not based on any failure of method, but on "incommensurable ways of seeing the world and of practicing science in it"); Feyerabend, *Problems of Empiricism*, in 2 Beyond the Edge of Certainty 145, 146 (R. Colodny ed. 1965) ("Thinkers who accept different views concerning the nature of experience and the way in which theories are to be related to experience have proceeded in different ways and thereby given rise to different kinds of scientific knowledge.") (footnotes omitted).

118. The term "relativism" derives from the fact that taking a subjective view of science means that the validity of a theory is relative to the viewpoint of the person determining validity. Relativists take an historical view of science and argue that "every particular standard of scientific rationality is violated by some major historical transition in scientific life." Doppelt, *Relativism and Recent Pragmatic Conceptions of Scientific Rationality*, in Scientific Explanation and Understanding 107, 107 (N. Rescher ed. 1983).

119. Epistemology is "the study of the method and grounds of knowledge esp[ecially] with reference to its limits and validity." Webster's Third New International Dictionary of the English Language Unabridged 765 (1961).

120. *See* Feyerabend, *Against Method: Outline of an Anarchistic Theory of Knowledge*, in IV Minnesota Studies in the Philosophy of Science—Analyses of Theories and Methods of Physics and Psychology 17, 17 (M. Radner & S. Winokur eds. 1970).

121. *Id.* at 26.

122. *See id.* at 23 n.27.

123. T. Kuhn, *supra* note 84, at 5-7 (pointing out that "normal science" is based on the assumption that the scientific community knows what the world is like, but that a new theory can imply a change in the rules governing the prior practice of normal science).

rience and subjected to comparative evaluation."[124] To this extent the Logical Positivists are right; but the Relativists also are right when they point out that the criteria used for evaluation are neither universal nor changeless over time.[125] When the Relativists argue, however, that a scientist who objects to the rejection of a theory can arbitrarily articulate new criteria to achieve acceptance, they go too far. While there is no universal algorithm[126] for verifying scientific validity, this does not mean that no basis exists for judging theories and claims to scientific knowledge.

a. *The Scientific Commitment to Objectivity: Acceptance as Evidence of Validity*

Since the late sixteenth and early seventeenth centuries, science has always been rooted in a commitment to objectivity, even if the evidence used is not purely objective, and to justification in terms of reasons and evidence.[127] The goal for scientists is "to obtain systematic knowledge that provides understanding of the world we live in."[128] Scientists evaluate their ideas against criteria that involve testability, objectivity, impartiality, and a belief in a deep and obvious connection between evidence and reason.[129] Guidelines for the practice of science, however, are not fixed; they evolve along with scientific knowledge.[130]

---

124. I. Scheffler, Science and Subjectivity 10 (1967).

125. *See* Laudan, *A Problem-Solving Approach to Scientific Progress*, in Scientific Revolutions 144, 144 (I. Hacking ed. 1981).

126. An algorithm is "[a] set of well-defined rules for the solution of a problem in a finite number of steps." McGraw-Hill Dictionary of Scientific and Technical Terms 50 (3d ed. 1984).

127. *See* Suppe, *Afterword* to The Structure of Scientific Theories 650 (F. Suppe ed. 2d ed. 1977). Professor Suppe observes:

> [F]ar more of science is concerned with reasoning, argument, and marshaling evidence than with manipulating nature in the laboratory. In short, a central and characteristic activity of science is *the use of reason* in the suggestion and development of hypotheses and theories and in evaluating the knowledge claims made by those who advance such hypotheses and theories.

Id. (emphasis in original).

128. *Id.*

129. *See* I. Scheffler, *supra* note 124, at 1; Siegel, *What is the Question Concerning the Rationality of Science?*, 52 Phil. of Sci. 517, 528-32 (1985).

130. *See* M. Harris, Cultural Materialism 26 (1979) ("The task of specifying precisely what kinds of guidelines ought to be followed falls to the practitioners."); D. Shapere, *supra* note 95, at 214 ("What is counted as 'evidential' and 'observational' in science is not something given once and for all, but evolves along with scientific knowledge."); Toulmin, *The Evolutionary Development of Natural Science*, 55 Am. Scientist 456, 460 & 462 (1967) (noting that those who carry on an intellectual tradition determine if it is scientific, and that the intellectual situation within a given branch of science determines which general types of hypotheses are taken seriously).

Modern physicists, for example, have developed criteria for deciding when they have observed a subatomic particle and what kind of particle they have observed. *See* Achinstein, *The Problem of Theoretical Terms*, in Readings in the Philosophy of Science 234, 236 (B. Brody ed. 1970). Scientists of Isaac Newton's era neither required, nor could have conceived of, such criteria.

The key to rationality and consistency lies in the *process* by which science maintains its commitment to objectivity. Typically, as scientists work with a theory and develop it in detail, the theory tends to become more and more accepted as valid, and confidence grows in its application. Validity is not an all-or-nothing proposition, as Logical Positivism would have it, but rather a matter of degree. Observation and experimentation are used to find shortcomings, to determine how to make improvements, and "to discover how to eliminate known artificialities, distortions, oversimplifications, and errors in the descriptions, explanations, and predictions of reality that the theory affords."[131] Only after a theory has survived a period of this kind of testing, review and refinement can it be used without significant questions,[132] and even then, it remains open to renewed doubt.[133] One philosopher has written that this process not only reflects the scientific method, but that "it *is* the scientific method."[134]

The question of scientific validity, therefore, hinges on acceptance,[135] and the answer depends on the purpose for which a theory is advanced and on the extent to which it has been tested. Acceptance may imply no more than a commitment to do research. Indeed, scientists often use a theory for some purposes while general acceptance remains quite tentative.[136] When invalidity has serious consequences, however, acceptance requires more complete validation. This higher standard has developed because the history of science contains countless examples of well-reasoned hypotheses that proved to be wrong. For example, early successes with vaccines led Robert Koch, one of the nineteenth century founders of modern bacteriology, to reason that a vaccine could be developed against any disease once the causative bacterium was isolated. He proceeded to produce a vaccine for tuberculosis, which the German government quickly started to use. Unfortunately, the reasoning was invalid, and vaccination actually caused dormant cases to become active.[137]

---

131. Suppe, *Afterword* to The Structure of Scientific Theories 706 (F. Suppe ed. 2d ed. 1977).

132. *See* Ziman, *What is Science*, in Introductory Readings In the Philosophy of Science 35, 40 (E.D. Klemke, R. Hollinger & A.D. Kline eds. 1980); *see generally* Merton, *Behavior Patterns of Scientists*, 38 Am. Scholar 197, 198 (1969) (noting the competition among scientists to be first in declaring a new discovery).

133. *See* K. Popper, The Logic of Scientific Discovery 47 (2d ed. 1968) ("there can be no ultimate statements in science") (emphasis omitted).

134. Ziman, *supra* note 132, at 40 (emphasis in original); *see id.* at 35-40.

135. *See* D. Shapere, *supra* note 95, at xxiii (science aims to determine a range of possible solutions and to establish criteria for what counts as an acceptable solution).

136. *See* van Fraassen, *Glymour on Evidence and Explanation*, in X Minnesota Studies in the Philosophy of Science—Testing Scientific Theories 165, 168 (J. Earman ed. 1983).

137. For an interesting modern account of this event, see R. Dubos & J. Dubos, The White Plague 100-10 (1987). For a contemporary account, see K. Wezel, Robert Koch 30-53 (1912) (in German).

Koch gained fame in 1870 by isolating the bacterium that causes anthrax, following it through its life cycle, and producing anthrax by injecting bacteria into laboratory animals. He went on to develop laboratory techniques still used in bacteriological laborato-

Beyond experimental testing, acceptance of a theory also depends on the theory's level of development and sophistication.[138] At the primitive stage of development when a theory consists of little more than observed patterns of regularity, traditional inductive confirmation is more central to scientific reasoning than it is for more refined theories. For example, there exists at present no overarching concept of carcinogenesis;[139] therefore, confirmation of an hypothesis about a relationship between exposure to a chemical and cancer must depend heavily on observed correlations and the methods of epidemiology.[140] For a comprehensive theory such as the general theory of relativity, however, much less evidence is required. Only six or seven direct tests of the basic tenets of the theory have been attempted,[141] yet it has gained almost universal

---

ries around the world. On March 24, 1882, he electrified the scientific world by announcing that he had discovered the organism responsible for tuberculosis. In 1890, he again stunned the world by announcing that he had isolated a substance that could protect against tuberculosis, and even cure the disease. An English doctor by the name of A. Conan Doyle (of Sherlock Holmes fame) carefully reviewed Koch's work and "concluded his analysis by expressing much doubt as to the usefulness of tuberculin [the substance Koch had isolated] in therapy." Dubos & Dubos, *supra*, at 106. Unfortunately, "Doyle's anticipations were amply confirmed by subsequent events. It soon became obvious that tuberculin killed many more patients than it helped, and the treatment fell in [sic] discredit almost everywhere." *Id.*

Precisely because of such occurrences, federal law requires scientific clinical evaluation of most drugs before they are marketed. *See* Federal Food, Drug and Cosmetic Act, 21 U.S.C. § 355 (1982 & Supp. III 1985). For an interesting history of legislation in this area, see 39 Fed. Reg. 33229, 33231-32 (Sept. 16, 1974).

138. *See* Suppe, *Afterword* to The Structure of Scientific Theories 724 (F. Suppe ed. 2d ed. 1977).

139. Carcinogenesis is "the origin or production of cancer." Stedman's Medical Dictionary 223 (24th ed. 1982).

140. *See* Peto, *The Preventability of Cancer*, in Cancer Risks and Prevention 1, 13 (M. Vessey & M. Gray eds. 1985). As Peto points out:

> [T]here may be many qualitatively different ways of turning normal cells into cancer cells, with the ones that matter being those that underlie the common cancers; and, there is no guarantee whatever that the mechanisms of cancerous change that are currently being investigated in the laboratory do not differ so greatly from those that importantly affect people as to make the related laboratory findings of no direct human relevance. To underline these difficulties, laboratory studies have thus far failed to identify reliably which are the important carcinogens and co-carcinogens in tobacco smoke, and have consistently failed to reproduce in animals the carcinogenic effects of alcohol in humans.

Id. The author goes on to contrast the mechanistic laboratory-based approach with the epidemiologically based "black box" strategy that "tries to identify as many correlates, or inverse correlates, of human cancer risk as possible, without understanding exactly how cancer arises." *Id.* He further observes:

> Not all correlates point to true causes, of course. Indeed, most of them do not. But, having identified many correlates or inverse correlates, it is reasonable to hope that further investigation of them . . . will lead to identification of a few truly causative or truly protective factors with respect to which humans already differ.

Id. Finally, Peto points out that the epidemiology approach has yielded by far the most important findings to date. *See id.* at 14.

141. *See* Suppe, *Afterword* to The Structure of Scientific Theories 724 (F. Suppe ed. 2d ed. 1977).

acceptance.

b. *Publication as Evidence of Acceptance*

Because acceptance is predicated on a process of refinement and critical review, communication among scientists is a necessity; hence the importance of publication to scientific inquiry. "The audience to which scientific publications are addressed is not passive; by its cheering or booing, its bouquets or brickbats, it actively controls the substance of the communications that it receives."[142] Publication does not by itself confer validity, but it does serve as an evidentiary threshold of validity. If a theory is not accepted anywhere in the literature of science, strong doubts must arise.

Reference to the literature also raises ancillary questions about the relative quality of journals, and about the scientific fields in which publications typically appear on particular issues. The question of quality relates to the amount of critical review that precedes publication, and can be resolved fairly easily by emphasizing the importance of the peer review system. The level of scrutiny involved in peer review varies widely from journal to journal, but individual journals develop reputations for the degree to which they are strict or lax.[143] Although not perfect, peer review generally serves at least as an initial screening process.[144]

c. *Selecting the Relevant Field of Science*

Resolving the evidentiary question of whether a journal publication adequately covers the scientific claims at issue requires a more detailed examination of the way in which science is conducted within investigative fields such as biology, biochemistry, or oncology.[145] Fields develop and grow as scientists organize their work around central problems in an effort to fill gaps in their knowledge and understanding of the world.[146] Because acceptance of a theory or scientific claim depends on the beliefs

142. Ziman, *supra* note 132, at 41; *see also* Lakatos, *History of Science and its Rational Reconstructions*, in Scientific Revolutions 107, 119 (I. Hacking ed. 1981) (discussing the way in which science progresses, and noting that "[t]he scores of the rival sides . . . must be recorded and publicly displayed at all times") (emphasis omitted).

143. *See* Note, *supra* note 85, at 34 n.26.

144. *See* Cole, Rubin & Cole, *Peer Review and the Support of Science*, 237 Sci. Am. 34, 34 (Oct. 1977) (concludes that "the NSF peer-review system [used to distribute grant money] is in general an equitable arrangement that distributes the limited funds available for basic research primarily on the basis of the perceived quality of the applicant's proposal. . . . [T]he NSF does not discriminate systematically against noneminent scientists in ways that some critics have charged."); Merton & Zuckerman, *Institutionalized Patterns of Evaluation in Science*, in The Sociology of Science 460, 461 (W. Storer ed. 1973) ("Although the referee [peer review] system has its inefficiencies, practicing scientists see it even in its current form as crucial for the effective development of science."); Note, *supra* note 85, at 27-28 ("For years, the scientific community has validated experimental results through a process known as peer review.").

145. *See* Darden & Maull, *Interfield Theories*, 44 Phil. of Sci. 43, 43-44 (1977).

146. *See id.* at 44 (other identifying features of a field include "facts related to [the central] problem, general explanatory factors and goals providing expectations as to how

of scientists within a particular scientific field, using acceptance as evidence of validity works only if there is a way to address problems related to the choice of which field is applicable to the legal issue at hand.

If a dispute about a scientific claim reduces to the issue of which field is appropriate, how is the choice made? For example, must the law choose between toxicology and pharmacology as the relevant field when deciding if a drug has caused illness or death? In other instances, the issue may be whether there is any appropriate field, or whether the validity of a conclusion that derives from the accepted practices and principles of several fields can readily be determined. Finally, what prevents improper self-validation by a group that attempts to create its own field with its own publications?[147]

These questions do not pose insurmountable problems because of the extent to which scientific fields overlap and interact. Logical Positivists, in keeping with their commitment to logical structure, actually saw science as a hierarchy of disciplines that eventually would reduce to basic, universal principles,[148] a position that has gone the way of their view of theories. A less rigid version of scientific unity, however, has survived. Broad fields such as physics, astronomy, or biology include subfields or specialities like high energy physics or solid state physics, and even subsubfields in which relatively few scientists work.[149] Specialists, however, seek wider recognition. They will endeavor, for example, to place papers in journals covering broader, more inclusive fields. This results in wider publication and recognition for the individual scientists[150] and in peer review and criticism from outside the subspecialty. This process provides a way to avoid self-validation by a small group. As one writer has put it, a scientist cannot create his own niche.[151]

Another indication of how well a field fits into science is citation of its publications by neighboring fields.[152] Often, one branch of science, as a part of its own development, will generate "important conceptual, theoretical and problem-solving links with other branches."[153] Scientists

---

the problem is to be solved, techniques and methods, and, sometimes . . . concepts, laws, and theories") (footnotes omitted).

147. The current debate about creation science exemplifies this problem. *See infra* Part IV.A.2.

148. *See* Oppenheim & Putnam, *Unity of Science as a Working Hypothesis*, in II Minnesota Studies in the Philosophy of Science—Concepts, Theories and the Mind-Body Problem 3, 27-28 (H. Feigl, M. Scriven & G. Maxwell eds. 1958).

149. *See* T. Kuhn, *supra* note 84, at 177; D. Shapere, *supra* note 95, at 320-21.

150. *See* Doreian, *A Measure of Standing of Journals in Stratified Networks*, 8 Scientometrics 341, 342 (1985) (pointing out that citations leave a pattern of ties between journal articles and between journals; also citing examples of journal networks).

151. *See* Lugg, *Theory Choice and Resistance to Change*, 47 Phil. of Sci. 227, 239 (1980).

152. *See* Weinberg, *Criteria for Scientific Choice*, in Science Observed 134, 136 (F. Jevons ed. 1973) ("Relevance to neighboring fields of science is . . . a valid measure of the scientific merit of a field of basic science.").

153. Maull, *Unifying Science Without Reduction*, 8 Stud. Hist. Phil. Sci. 143, 143 (1977); *see also* Crane, Invisible Colleges 99-101 (1972) (pointing out how scientific com-

often work in several specialty fields,[154] and they freely look to other fields for information. If a theory from one field leads to a different conclusion than a theory from a second field, scientists in both become aware of the conflict. Both theories are then cast in doubt, and the area of divergence becomes a problem for investigation in both fields.

4. The Need for an Acceptance Test Based Upon a Correct View of Science

When a dispute about the admissibility of scientific evidence hinges on the validity of an expert's reasoning, an acceptance test based on the current, realistic view of science provides the only way to reach a rational and consistent decision. A court should determine which scientific fields are relevant to the dispute and then turn to the peer-reviewed literature from those fields for guidance. Experts should be required to make their reasoning clear and explicit, and to explain how their conclusions derive from accepted scientific practice. An expert who cannot demonstrate that his or her reasoning conforms to the standards of science should not be allowed to testify. A complete legal analysis of scientific evidence also requires consideration of the reliability of the expert's conclusions, but because reliability properly involves only legal standards, scientific acceptance should have no bearing on that issue.

II. RESOLUTION OF THE *FRYE* VERSUS RELEVANCY DEBATE

An acceptance test focused on reasoning and scientific standards differs from the traditional test of *Frye v. United States*.[155] Where *Frye* makes acceptance an exclusive test, this Article's proposed approach uses it only when the reasoning underlying an expert's conclusions is called into question. Unlike *Frye*, this approach clearly establishes what must be accepted, and it provides a minimal evidentiary threshold by reference to publications.

Distinguishing the validity of reasoning from the reliability of conclusions also permits the resolution of the debate that has dominated legal thinking about scientific evidence for some sixty years.[156] Two polar positions have defined the contours of this debate: the general acceptance test of *Frye*[157] and the pure relevancy test advocated by Professor Mc-

munities interact and how all research areas must rely to some extent on other fields of science).

154. *See* T. Kuhn, *supra* note 84, at 178 (noting that individual scientists generally will belong to several communities "either simultaneously or in succession").

155. 293 F. 1013 (D.C. Cir. 1923).

156. Although *Frye* was decided 65 years ago, courts still frequently cite it, and legal commentators still focus on it, as evidenced by the fact that the 1983 Symposium on Science and the Rules of Evidence published in Federal Rules Decisions essentially involved a debate about *Frye* and its alternatives. *See Symposium on Science and the Rules of Evidence*, 99 F.R.D. 187 (1983). For an excellent summary of the case law and commentary on *Frye*, see generally Reed v. State, 283 Md. 374, 391 A.2d 364 (1978).

157. *Frye*, 293 F. at 1014. For the text of the *Frye* test, *see infra* note 171.

Cormick and others, according to which scientific evidence is treated no differently than other evidence.[158] When the latter test is used, the focus is usually on the reliability aspect of relevance. This Article's proposal encompasses both poles, as well as the multifactor balancing approaches that one finds nestled in between.[159] These balancing approaches essentially are refined versions of the relevancy test, but they reflect the way in which *Frye* is often actually used.[160]

Within the analytical framework proposed in Part I, the extreme positions of the debate become complementary rather than antagonistic. The relevancy standard, especially its multifactor variants, easily equates with the legal considerations portion of the proposed analytical framework; and *Frye*'s general acceptance test, when confined to the question of validity, fits neatly within the scientific portion. Unfortunately, the ambiguity of *Frye* has obscured this basic harmony. The cases since *Frye* reveal extreme incoherence and inconsistency in the application of the general acceptance test. In most of them, the validity of reasoning is not at issue, and the *Frye* test merely masks legal decisions about reliability.[161] The standards advanced in place of *Frye* further cloud the matter, either by ignoring acceptance completely or by demoting it to one of several undifferentiated items on a list of factors to be weighed and balanced.[162] Because no analytical structure relates the items to one another, the significance of acceptance as evidence of validity is lost.

Despite this generally murky picture, some courts have distinguished acceptance from reliability,[163] and several have recognized that acceptance properly relates to the validity of an underlying theory or its application.[164] A few decisions outline an approach very close to the one advocated here,[165] but their significance has not been widely recognized.

158. *See* E. Cleary, McCormick on Evidence § 203, at 608 (3d ed. 1984) ("Any relevant conclusions supported by a qualified expert witness should be received unless there are distinct reasons for exclusion."); D. Louisell & C. Mueller, 1 Federal Evidence § 105, at 826 (1977) ("The *Frye* approach as such should be abandoned, and the term 'general scientific acceptance' absorbed into the judicial notice concept, under which the relevance of scientific proof can be established without taking time to lay an elaborate scientific foundation through testimony describing the principles involved.").

159. *See, e.g.*, 3 J. Weinstein & M. Berger, Weinstein's Evidence, ¶ 702(03), at 702-18 to 702-21. For a more detailed discussion, see *infra* Part II.B.3.

160. At least one court has acknowledged this relationship. In State v. Free, 493 So. 2d 781 (La. Ct. App. 1986), *cert. denied*, 499 So. 2d 83 (1987), the court used a balancing test to determine that voice spectograph identification evidence was inadmissible. *See* 493 So. 2d at 787. The court explicitly posed the question, however, whether in reality, it just was applying another version of the general acceptance test. *See id.* at 787 n.9. It noted that the Fifth Circuit had applied a balancing test, and reached the result that the value of hypnosis in refreshing a witness's memory had not been "clearly established," and that "[t]his language sounds close to the general acceptance standard." *Id.*

161. *See infra* Part II.A.

162. *See infra* Part II.B.3.

163. *See infra* note 268.

164. *See infra* note 267.

165. *See, e.g.*, People v. Collins, 94 Misc. 2d 704, 706-07, 405 N.Y.S.2d 365, 367 (Sup. Ct. 1978); Phillips v. Jackson, 615 P.2d 1228, 1233-35 (Utah 1980).

The need for wider recognition becomes particularly apparent in cases involving forensic psychology and psychiatry,[166] one of the few medical areas to which *Frye* or its alternatives has been applied. The importance of the scientific reasoning in these cases varies greatly, but, as a group, they highlight the need for this Article's proposed mode of analysis and for at least a rudimentary understanding of the philisophical foundations of science.

A. Frye's *Legacy: The Wrong Half of the Right Approach*

Few decisions have dominated an area of the law or caused as much confusion as *Frye v. United States*.[167] Decided by the United States Court of Appeals for the District of Columbia Circuit in 1923, the case marked the first judicial recognition of the need for special rules for scientific evidence.[168] The appellant in *Frye* had been convicted of a murder he claimed he had not committed.[169] To bolster his defense, he had sought to introduce evidence that he had passed a systolic blood pressure deception test,[170] a precursor of the modern polygraph lie detector. The trial court rejected this evidence, and in affirming that decision, the D.C. Circuit articulated its now famous "general acceptance" rule.[171]

Frye's general acceptance test quite clearly was meant to apply to scientific principles, which is not inconsistent with the validity portion of this Article's proposal. The ambiguous reference to "the thing from which the deduction is made"[172] has, however, led many courts to con-

166. *See infra* Part II.D.

167. 293 F. 1013 (D.C. Cir. 1923).

168. *See* E. Cleary, *supra* note 158, § 203, at 605. It should be noted that many issues related to scientific evidence have been decided without any special rule, both before and after *Frye*. *See* Reed v. State, 283 Md. 374, 400-504, 391 A.2d 364, 377-428 (1978) (Smith, J., dissenting) (discussing cases not employing the *Frye* test).

169. For an interesting historical account of *Frye*, see Starrs, "*A Still-Life Watercolor*": Frye v. United States, 27 J. of Forensic Sciences 684 (1982). Professor Starrs points out that the two-page *Frye* opinion provides very little insight into the facts of the case. *See id.* at 685. He has, however, done considerable research beyond the opinion. James Alphonso Frye was accused of murdering a doctor. *See id.* at 687-88. At first he confessed, but later he recanted. *See id.* at 688. At trial he attempted a weak alibi, but was convicted. *See id.* at 689. He served eighteen years, and after his release he married a woman who, in 1981, told Professor Starrs that Frye had caused her nothing but unremitting trouble. *See id.* at 692. Devotees of scientific evidence have long felt the same way.

170. *See* 293 F. 1013, 1013 (D.C. Cir. 1923).

171. *See id.* at 1014. The test, as articulated by the court, reads as follows:

> Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized, and while courts will go a long way in admitting expert testimony deduced from a well-recognized scientific principle or discovery, the thing from which the deduction is made must be sufficiently established to have gained general acceptance in the particular field in which it belongs.

Id.

172. *Id.*

centrate on devices or techniques,[173] rather than on theories or reasoning—a problem greatly compounded by the fact that reasoning was not really at issue in *Frye*. The theory that emotions correlate with involuntary physiological manifestations was accepted for some purposes at the time of *Frye*,[174] as was the idea that lying might trigger emotions.[175] Though the court did not cite a single legal or scientific reference in support of its holding, a substantial body of literature about the systolic blood pressure test in fact was available.[176] Under some circumstances the test had proven useful, and based on acceptance, it surely represented a valid, scientific application of a valid, scientific theory.[177]

The *Frye* court's real concern must have been whether the specific conclusion about Mr. Frye was, as a legal matter, reliable and understandable enough to warrant presentation to the jury, yet it blurred this crucial point by turning a legal decision *about* a scientific conclusion into a decision about whether the conclusion *was scientific*.[178] Lie detector deci-

173. See, for example, State v. Washington, 229 Kan. 47, 53-55, 622 P.2d 986, 991-92 (1981), discussed *infra* notes 194-98, in which focus on the electrophoresis technique, which is recognized and accepted for some purposes, led the court to overlook the way in which the technique was being used in the case before it. In Medley v. United States, 155 F.2d 857, 859-60 (D.C. Cir. 1946), discussed *infra* notes 191-93 and accompanying text, the same problem arose relative to the spectroscopic technique of metallurgical analysis. As examples of cases in which the court focused on a device, see City of Seattle v. Peterson, 39 Wash. App. 524, 527, 693 P.2d 757, 758 (1985) (regarding Doppler radar speed detector, the court noted the "issue here is not the reliability of a scientific principle . . . but whether the particular machine employing the principle is so designed and constructed that the results produced by proper operation are reliable"); *see also* People v. Magri, 3 N.Y.2d 562, 566, 147 N.E.2d 728, 730, 170 N.Y.S.2d 335, 337-38 (1958) ("We think the time has come when we may recognize the general reliability of the radar speedmeter . . . and that it will no longer be necessary to require expert testimony in each case as to the nature, function or scientific principles underlying it . . . ."); People v. Tilley, 120 Misc. 2d 1040, 1047, 466 N.Y.S.2d 983, 987 (Erie County Ct. 1983) (following review of theory behind breathalyzer and discussion of questions related to a particular type of instrument, court held that "[i]t is not necessary that the perfection of the machine be proven beyond all doubt. Reasonable proof of its accuracy has been established."). *But see* People v. Morse, 325 Mich. 270, 275, 38 N.W.2d 322, 324 (1949) (excluding "drunkometer" test results).

174. *See* Comment, *The Use of Psychological Tests to Determine the Credibility of Witnesses*, 33 Yale L.J. 771, 773 n.13 (1924) ("it is possible to influence expressive measurements somewhat by emotional control") (citing Burtt, *The Inspiration—Expiration Ratio During Truth and Falsehood*, 4 J. of Experimental Psychology 1, 23 (1921)).

175. *See* Starrs, *supra* note 169, at 686 (" 'few of us would doubt, or need any evidence other than experience, that conscious lying produces in the ordinary man emotional disturbances' ") (quoting McCormick, *Deception Tests and the Law of Evidence*, 15 Calif. L. Rev. 484, 484 (1927)).

176. *See* Annotation, *Physiological or Psychological Deception Test*, 34 A.L.R. 147 (1925). In commenting on *Frye*, this annotation notes that among other publications available at the time were a 1922 article in American Bar Association Reports, a 1921 article in the Journal of Experimental Psychology, a 1917 article in the Journal of Experimental Psychology, and two 1921 articles in the Journal of Criminal Law. *See id.* at 147-48; *see also* Case Note, 37 Harv. L. Rev. 1138 (1924).

177. *But see* Skolnick, *supra* note 80, at 699 (pointing out problems of verifying accuracy of lie detection).

178. At least one commentator previously has noted that courts essentially are making

sions since *Frye*, despite going into far more detail about the specific techniques involved, by and large have perpetuated this misuse of general acceptance.

In *People v. Davis*,[179] for example, the Supreme Court of Michigan acknowledged "the proven value of the polygraph in the fields of crime detection and criminal interrogation,"[180] but it was troubled by the possibility of erroneous results, "estimated at being from less than 10 percent to 25 percent by various authorities."[181] Neither the validity of reasoning nor acceptance was at issue. The evidence before the court established the underlying principle that a definite relationship exists between willful lying and an elevation of blood pressure, changes in respiration, and variations in the electrical resistance of the skin.[182] The evidence also made it clear that applying this principle by recording these reactions on the polygraph was an accepted way to determine truthfulness.[183] The court's real concern was the "tremendous weight which such tests would necessarily carry in the minds of a jury. . . ."[184] Nonetheless, citing *Frye* and related cases, it purported to ground its opinion on acceptance.[185]

---

policy judgments in their decisions about lie detectors. *See* Trautman, *Logical or Legal Relevancy—A Conflict in Theory*, 5 Vand. L. Rev. 385, 395-97 (1952). It should be noted that at least one court has in fact found lie detector results to be sufficiently reliable for legal purposes. *See* United States v. Ridling, 350 F. Supp. 90, 99 (E.D. Mich. 1972) (holding that defendant in perjury case could introduce polygraph evidence provided that defendant submit to additional polygraph tests administered by court-appointed expert).

One very significant result of fixing on the wrong issue is that it causes courts to ignore totally the constitutional question of whether the evidence presented by a defendant in a criminal case has to reach the same level of reliability as evidence offered by the prosecution. *See* Giannelli, *supra* note 23, at 196-98. The recent Supreme Court decision in Rock v. Arkansas, 107 S. Ct. 2704 (1987), indicates that this is a genuine concern. The appellant in *Rock* was convicted of manslaughter for shooting her husband. *See id.* at 2705. To refresh her memory as to the precise details of the shooting, she had undergone hypnosis. *See id.* Because of serious questions about the scientific validity of refreshing memory in this fashion, a number of states, including Arkansas, have held that the testimony of any witness who has undergone hypnosis is inadmissible. *See infra* Part II.D.1. In *Rock*, however, this rule conflicted with the constitutional right of a criminal defendant to testify on his or her own behalf. The Supreme Court held that under these circumstances,

> [a] State's legitimate interest in barring unreliable evidence does not extend to *per se* exclusions that may be reliable in an individual case. Wholesale inadmissibility of a defendant's testimony is an arbitrary restriction on the right to testify in the absence of clear evidence by the State repudiating the validity of all posthypnosis recollections.

107 S. Ct. at 2714.

179. 343 Mich. 348, 72 N.W.2d 269 (1955).

180. *Id.* at 371, 72 N.W.2d at 282.

181. *Id.*

182. *See id.* at 369, 72 N.W.2d at 280-81.

183. *See id.*, 72 N.W.2d at 281.

184. *Id.* at 372, 72 N.W.2d at 282.

185. *See id.* at 371-72, 72 N.W.2d at 281-82. For other decisions that employ a similar analysis, see State v. Lowry, 163 Kan. 622, 629-30, 185 P.2d 147, 151-52 (1947) (device has utility, but questions related to accuracy, reliability and ease of interpretation exist);

A number of undesirable results stem from the *Davis* kind of analysis, which forces the use of acceptance as an ultimate test, rather than as evidence of validity, and which confuses validity with reliability. In addition to using *Frye* to mask legal decisions about reliability with a simple label, courts sometimes combine acceptance and reliability into a version of *Frye* that requires evidence to be "accepted as reliable." This fused version in effect cedes to scientists the decision about what constitutes sufficient reliability for legal purposes. When *Frye* is used to decide questions related to the validity of reasoning, the acceptance test works better, but the analysis still often falls short by failing to include any consideration of reliability.

1. Acceptance as a Label

At least one court has explicitly acknowledged the problem of using *Frye* to decide reliability issues. The issue in *United States v. Franks*,[186] as in *United States v. Williams*,[187] was the admissibility of the voiceprint method for comparing a suspect's voice with an incriminating recording.[188] The court in *Franks* held that there was no evidence to rebut "the government's claim that voiceprint analysis is sufficiently accurate to be admissible."[189] Apparently realizing, however, that accuracy does not necessarily imply general acceptance, the court simply declared that "[i]f a scientific process is reliable, or sufficiently accurate, courts may also deem it 'generally accepted.' "[190] It is at this point that acceptance becomes a mere label rather than a mode of analysis.

2. Reliability and Validity Both Ignored Because of the Acceptance Test: The Problem of Self-validation

When the acceptance rule is applied in a rigidly legalistic way, it often

State v. Kolander, 236 Minn. 209, 221-22, 52 N.W.2d 458, 465 (1952) (lie detector acknowledged to be useful, but inadmissible for three reasons: probable impact on jurors; not of sufficient scientific and psychiatric accuracy; and difficulties in interpreting results); State v. Arnwine, 67 N.J. Super. 483, 491-93, 171 A.2d 124, 129 (1961) (citing *Frye*, but really basing rejection on the margin of error intrinsic to the test; court also bothered by need for the well-trained expert to interpret results); Henderson v. State, 230 P.2d 495, 501-04 (Okla.) (citing *Frye*, but really basing decision on reliability and accuracy, court found medical consensus on physiologic manifestation of emotion, but that polygraph, even when properly interpreted, does not work for about 10% of population), *cert. denied*, 342 U.S. 898 (1951).

186. 511 F.2d 25 (6th Cir.), *cert. denied*, 422 U.S. 1042 (1975).

187. 583 F.2d 1194 (2d Cir. 1978), *cert. denied*, 439 U.S. 1117 (1979). For a further discussion of *Williams* and voiceprint analysis, see *supra* notes 52-63 and accompanying text.

188. *See Franks*, 511 F.2d at 30 (taped conversations with government informants).

189. *Id.* at 33.

190. *Id.* at 33 n.12; *see also* United States v. Lewellyn, 723 F.2d 615, 619 (8th Cir. 1983) (finding reliability one of the most important factors to consider in determining acceptance) (citing United States v. Alexander, 526 F.2d 161, 163 (8th Cir. 1975)). *But see* United States v. Addison, 498 F.2d 741, 744 (D.C. Cir. 1974) (distinguishing reliability from general acceptance).

becomes a simplistic requirement for scientific precedent, and both validity and reliability may fade into the background. This wooden reliance on precedent can lead to an acceptance test that ignores the purpose or application for which a method or device has been accepted. In *Medley v. United States*,[191] for example, the prosecution introduced evidence based on spectroscopic metallurgical analysis to connect bits of metal found on a file belonging to the defendant with filed-down bullets that had killed a murder victim.[192] Although this technique was generally accepted in scientific research and for certain industrial uses, the opinion gives no indication that it had any track record for the forensic application at issue. Nonetheless, the court in *Medley* held that testimony based in part on this analysis was admissible.[193]

In some cases, such an uncritical approach to acceptance allows a group that advocates a technique or method to self-validate it simply by declaring acceptance. For example, in *State v. Washington*,[194] results of a "multi-system" blood-typing test used only by crime labs were used to link the defendant to a rape-murder.[195] The murderer apparently had cut himself during the commission of his crime, and the state claimed that the dried drops of blood found at the crime scene matched the defendant's blood.[196] Though based on methods widely used and accepted for analysis of fresh blood in clinical, research, and genetic laboratories, and in paternity testing clinics, the validity of the state's forensic extension of these methods to a single sample of dried blood raised a number of unanswered questions.[197] Kansas' highest court, however, upheld the

191. 155 F.2d 857 (D.C. Cir.), *cert. denied*, 328 U.S. 873 (1946).

192. *See id.* at 860.

193. *See id.*; *see also* People v. Haggart, 142 Mich. App. 330, 370 N.W.2d 345 (1985). In *Haggart*, the court held admissible testimony regarding the analysis of blood using a technique known as serological electrophoresis, even though the technique admittedly had virtually no track record for such a use. *See* 142 Mich. App. at 343-45, 370 N.W.2d at 352-53. For the court, it sufficed that electrophoresis was accepted in the scientific community for other applications. *See id.* at 344, 370 N.W.2d at 353. *Haggart*'s uncritical acceptance of electrophoresis soon was reversed by the Michigan Supreme Court in People v. Young, 425 Mich. 470, 505, 391 N.W.2d 270, 286 (1986), a case involving a form of electrophoresis known as the multi-system test. For a more detailed discussion of the problems surrounding the forensic use of multi-system electophoresis, see Black, *Evolving Legal Standards for the Admissibility of Scientific Evidence*, 239 Sci. 1508, 1509 (1988); *infra* notes 194-202 and accompanying text.

Another case properly distinguishing acceptance for one use from acceptance for another is United States v. Kilgus, 571 F.2d 508, 510 (9th Cir. 1978) (acceptance of radar system for determining location of aircraft did not mean acceptance of system for determining kind of airplane that had been located).

On the general problem of admitting evidence based on acceptance for the wrong purpose, see Starrs, Frye v. United States *Restructured and Revitalized: A Proposal to Amend Federal Evidence Rule 702*, 115 F.R.D. 92, 96 (1987).

194. 229 Kan. 47, 622 P.2d 986 (1981).

195. *See id.* at 49, 55, 622 P.2d at 988-89, 992.

196. *See id.*

197. *See* Grunbaum, *Physiological Stain Evidence: Guidelines to Assure Quality Analysis*, 1 Calif. Defender 20, 21 (1985); *see also* Black, *supra* note 193, at 1508-09; Jonakait, *Will Blood Tell? Genetic Markers in Criminal Cases*, 31 Emory L.J. 833, 852-54 (1982).

evidentiary use of the forensic blood test results without considering these questions, having been particularly impressed with the fact that the test was used "in over 100 criminal laboratories . . . and . . . the FBI research laboratory."[198] The court did not even mention the analytical theory underlying the multi-system test, much less consider how the multi-system differed from other applications.[199]

The Supreme Court of Kansas unfortunately overlooked the fact that use does not necessarily imply either valid reasoning or reliable results.[200] Unlike hospitals that test undeteriorated, uncontaminated blood using established techniques designed to assure safe transfusions, crime labs have no way of knowing if they are right or wrong in finding a match based on deteriorated, contaminated blood samples examined with a technique developed by police scientists.[201] In the hospital, patients would die if the blood test did not work, but complaints from innocent defendants cannot be distinguished from the protestations of the guilty. For this reason, acceptance of the forensic blood test by criminologists without controlled testing amounts to nothing more than self-validation,

198. 229 Kan. at 55, 622 P.2d at 992. For other examples of this problem of self-validation, see People v. Williams, 164 Cal. App. 2d Supp. 858, 862, 331 P.2d 251, 253-54 (1958) (nalline test for narcotics accepted by those who apply it); Robinson v. State, 47 Md. App. 558, 573-76, 425 A.2d 211, 219-21 (1981) (electrophoretic testing of blood) (discussed *infra* notes 215-17 and accompanying text).

199. *See* Black, *supra* note 193, at 1508-09.

200. *See* Jonakait, *supra* note 197, at 849.

201. *See infra* note 202. To some extent, even the developers of the multi-system test have acknowledged the need for more data to establish its validity and reliability. In 1986 they published an article that drew criticism from three FBI scientists who, though supporters of the technique, felt that the authors, Wraxall and Stolorow, had done "little to strike at the heart of the criticisms confronting its use. The paper did not support their scientific claims nor the reliability and validity of the . . . method." Murch, Kearney & Budowle, Discussion of "The Simultaneous Separation of the Enzymes Glyoxalase I, Esterase D, and Phosphoglucomutase," 32 J. of Forensic Sci. 1498, 1499 (1987). Wraxall and Stolorow responded that release of FBI data on the technique was long overdue. *See id.* at 1500.

Other forms of electrophoretic analysis in fact may be more valid than the multi-system, which itself eventually may gain wider and more general acceptance. *See generally* Budlowle & Murch, *Electrophoresis Reliability II—Historical Use in the FBI Laboratory, Validity, and Scientific Scrutiny*, Crime Laboratory Digest (in press) (arguing that the multi-system has been validated); Note, *The Admissibility of Electrophoretic Methods of Genetic Marker Bloodstain Typing Under the* Frye *Standard*, 11 Okla. City U.L. Rev. 773 (1986) (advocating multi-system as well as other techniques); *cf.* People v. Reilly, 242 Cal. Rptr. 496, 511-13 (Cal. App. 1987) (holding electrophoretic testing of dried blood to be generally accepted; multi-system test was not at issue). The entire debate may be mooted as still better techniques such as "DNA Fingerprinting" are developed. *See* Lewis, *DNA Fingerprints: Witness for the Prosecution*, Discover, June 1988, at 44; Moss, *DNA—The New Fingerprints*, 74 A.B.A.J. 66 (May 1988) (discussing matching based on very specific tests of genetic structure). It is interesting that George Sensabaugh, a professor at the University of California at Berkeley's School of Public Health, and a leading proponent of DNA matching, has called for additional study. According to Moss, Prof. Sensabaugh "says the forensic science community is concerned because the research validating DNA typing has come primarily from . . . private sector companies. That research should be published quickly for critical evaluation by other scientists, he says." 74 A.B.A.J. at 69.

a conclusion reached by the Michigan Supreme Court in a case decided five years after *Washington*.[202]

The reasoning in *Washington* also raises fundamental questions that relate to the use of any "matching" evidence.[203] These problems are particularly well illustrated by the evidentiary use of neutron activation analysis ("NAA").[204] NAA can detect "with great sensitivity the presence and concentrations of most chemical elements in materials, without, in most cases, destroying the sample analyzed."[205] As a valid analytical technique, NAA passes the acceptance test with relative ease,[206] but this

202. *See* People v. Young, 425 Mich. 470, 391 N.W.2d 270 (1986). The court in *Young* noted that no comprehensive control test had been run with the specific technique at issue, which had been developed by police scientists for police work. *See id.* at 491-92, 391 N.W.2d at 279-80. There were also questions about how the deterioration of blood affects the test results, *see id.* at 487-90, 391 N.W.2d at 277-79, and about the effect of crime scene contamination of blood samples. *See id.* at 495-99, 391 N.W.2d at 281-83. Thus, testimony from prosecution witnesses about the reliability of the test amounted to "the type of self-verification considered inconclusive in the scientific community." *Id.* at 476, 391 N.W.2d at 272. The court also was aware of the need to consider the views of nonforensic scientists who use electrophoresis, the blood testing technique at issue, *see id.* at 474-76, 391 N.W.2d. at 271-72, and was sensitive to the fact that "[t]he only prosecution witness having substantial experience with electrophoresis of evidentiary bloodstains relied on his own unpublished observations and an unpublished reliability study." *Id.* at 475, 391 N.W.2d at 272. The court noted that forensic uses of electrophoresis to analyze blood stains had not been "subjected to the scrutiny of the scientific community.... The scientific tradition expects independent verification of new procedures. When other scientists analyze and repeat the tests, they counteract the dangers of biased reporting." *Id.* at 499, 391 N.W.2d at 283.

203. In addition to obscuring the question of precisely how the disputed test could determine a match in blood types, the decision in *Washington* fails to recognize the logical flaws in concluding, even from a perfect match, that the defendant had been at the scene of the murder. The state's expert testified that the defendant's blood type was found in less than one percent of the black population. *See* State v. Washington, 229 Kan. 47, 49, 622 P.2d 986, 989 (1981). This would be relevant, however, only if it were known that the murderer was black, and that blacks comprised only a very small percentage of the population in the area where the murder occurred. If there were even a thousand blacks in the area, there would be about ten suspects, based on the blood-typing.

204. The NAA technique is very well described in a 1971 Comment:

> The detection of elements by activation analysis is possible because most elements can be made artificially radioactive. The resulting radioactive products can be identified by observation of the ways in which they undergo radioactive decay. Radioactive decay is a fairly complicated process during which the nuclei of most radioactive atoms (these nuclei are often referred to as radionuclides) emit a pulse of high-energy electromagnetic radiation called a gamma-ray which has an energy characteristic of the emitting nucleus. The activation analyst usually tries to count the number of gamma-rays of each energy which are emitted from a sample of activated material. He also observes the half-life of each emitting radionuclide. These measurements tell him what kind of radionuclide produced the gamma-rays, and from the number of gamma-rays of each energy he deduces the number of radioactive nuclei of that particular type which are present in the activated sample. From this information he can determine the preactivation elemental composition of the sample.

Comment, *The Evidentiary Uses of Neutron Activation Analysis*, 59 Cal. L. Rev. 997, 999-1000 (1971) (footnotes omitted).

205. *Id.* at 997 (footnote omitted).

206. *See id.* at 998.

does not validate conclusions based on the results of NAA testing.

The biggest problem with NAA evidence, as with the kind of blood matching performed in *Washington*, is the way in which experts misinterpret the results. The fact that paint found on the body of a hit-and-run victim exactly matches the paint on a suspect's car does not necessarily implicate the car. If the same kind of paint were used on a large number of cars, the match could not establish a reliable connection. Courts, however, often overlook this important gap in an expert's reasoning. In *People v. Collins*,[207] for example, NAA was used to determine if hairs found on a murder victim's clothing matched hairs found on the floor of the basement room where the murder allegedly occurred.[208] The prosecution's witness claimed that there was "a distinct and high level of probability" that the samples he examined had come from a common source,[209] and the court admitted this testimony over objections about the way the tests had been done.[210] Neither the court nor, apparently, the defense, considered the question of how unique the hairs were.[211] This basic logical flaw arises with regard to a number of matching techniques. It requires that courts probe the validity of the reasoning that leads to the conclusion that an identification has been made, rather than confining themselves to consideration of the analytical technique used to make the comparison.[212]

3. Letting the Scientists Decide: The "Acceptance of Reliability" Test

A number of courts have adjusted to the fact that *Frye* does not work very well for reliability issues by reformulating the general acceptance test to require "a showing that [a] technique has been generally accepted

207. 43 Mich. App. 259, 204 N.W.2d 290 (1972), *cert. denied*, 419 U.S. 866 (1974).

208. *See id.* at 264, 204 N.W.2d at 293.

209. *Id.*

210. *See id.* at 265, 204 N.W.2d at 293.

211. *See* 43 Mich. App. 259, 204 N.W.2d 290 (1972) (passim), *cert. denied*, 419 U.S. 866 (1974).

212. A number of commentators have discussed this problem. *See generally* Ellman & Kaye, *supra* note 45; Giannelli, *supra* note 4, at 1226; Tribe, *Trial by Mathematics: Precision and Ritual in the Legal Process*, 84 Harv. L. Rev. 1329 (1971); Comment, *supra* note 204, at 1014-20. The most frequently cited case on this point is People v. Collins, 68 Cal. 2d 319, 438 P.2d 33, 66 Cal. Rptr. 497 (1968), which involved a purse snatching committed by a black man with a mustache and beard, and a blond woman, who drove away from the crime scene together in a yellow car. *See id.* at 320-21, 438 P.2d at 34, 66 Cal. Rptr. at 498. The state introduced expert testimony from a statistician who testified that there was only a 1 in 12 million chance that this particular combination of factors could occur. *See id.* at 325, 438 P.2d at 36-37, 66 Cal. Rptr. at 500-01. The California Supreme Court held that this testimony was inadmissible because there existed no foundational evidence about the data upon which the probability estimate was based, *see id.* at 327, 438 P.2d at 37, 66 Cal. Rptr. at 502, and because even given the assumption that the combination of factors was as rare as the witness had concluded, there would still be a 40 percent probability that two or more couples would fit the description of the couple that had committed the crime. *See id.* at 331, 438 P.2d at 40, 66 Cal. Rptr. at 504.

as reliable in the scientific community in which it developed."[213] This maneuver fixes on the right issue in most cases, but may do so at the expense of turning legal decisions over to scientists,[214] a problem that becomes especially severe when acceptance is established on the basis of recognition within a single, limited field. In *Robinson v. State*,[215] another blood typing case, the court held that evidence of a match between blood on the defendant's clothing and the blood of an assault victim was admissible because the technique used was accepted as reliable in the field of forensic chemistry.[216] Given the testimony at trial, the court's decision in effect allowed the police department to decide whether the prosecutor's evidence was reliable enough for legal purposes.[217]

The aura of unfairness that surrounds *Robinson* is made all the more unsettling by the fact that the court did not consider the purposes for which other crime labs had accepted the test. Nor does the court's decision give the slightest hint about how frequently the test might produce an incorrect result. None of these holes in the legal analysis would have occurred had the court followed the approach proposed in this Article.

4. Acceptance Properly Used

Though often misused, *Frye* has worked well in cases involving issues directly related to the validity of reasoning. *United States v. Brown*[218] illustrates this better side of general acceptance. The defendant in *Brown*

213. People v. Shirley, 31 Cal. 3d 18, 34, 641 P.2d 775, 784, 181 Cal. Rptr. 243, 252 (emphasis added), *cert. denied*, 459 U.S. 860 (1982); *see* State v. Washington, 229 Kan. 47, 55, 622 P.2d 986, 992 (1981); Robinson v. State, 47 Md. App. 558, 575-76, 425 A.2d 211, 220-21 (1981).

214. A good discussion of this problem is found in Morse, *Crazy Behavior, Morals, and Science: An Analysis of Mental Health Law*, 51 S. Cal. L. Rev. 527 (1978):

> A scientist, *as a scientist*, might be able to state the quantitative probability that precisely defined behavior such as suicide was going to occur. If a scientist says "harm is likely," however, he is behaving as an ordinary layperson, assigning his own private, unscientific meaning to the terms 'harm' and 'likely.' This determination, however, is not part of the scientist's proper role in the courtroom. Rather, the judge or jury should consider the empirical data and then decide the social, moral, and legal question of whether the legal standard is met.

Id. at 591 (emphasis in original). Also see Markey, *Jurisprudence or "Juriscience"?*, 25 Wm. & Mary L. Rev. 525, 529-30 (1984):

> When questions arising from scientific or technical activities are presented to the courts under the artificial regime of bipartisan adversarial litigation, value questions may seem indistinguishable from questions of scientific fact. Failure to make the distinction, however, would allow moral, philosophical, and political decisions to be based solely on the outcome of a purely technical debate between scientific experts. This result has already occurred in a variety of instances.

Id. at 529-30; *see also* Steinbock, Richman & Ray, *supra* note 81.

215. 47 Md. App. 558, 425 A.2d 211 (1981).

216. *See id.* at 574-76, 425 A.2d at 220-21.

217. The state's witness was a forensic chemist employed by the county police department, *see id.* at 573, 425 A.2d at 219, and her testimony gave no indication of acceptance in any other area of science. *See id.* at 574-76, 425 A.2d at 220-21.

218. 557 F.2d 541 (6th Cir. 1977).

was accused of firebombing a Planned Parenthood Clinic.[219] Part of the evidence against him consisted of three hairs on a broken bottle found at the site of the bombing.[220] To link him to the bottle, and thus to the crime, the government presented testimony about ion microprobic analyses that they had done on the three suspect hairs and on hair samples from the defendant.[221]

The *Brown* court applied a four-part test, which determined admissibility based on the qualifications of the expert, the appropriateness of the subject of the testimony, conformity to a generally accepted explanatory theory, and probative value balanced against prejudicial effect.[222] For the ion microprobe testimony, the validity of the underlying theory was the central issue, and the court held that it had not been adequately established.[223] The experts had no demonstrable objective procedure for reaching their opinion. Neither of them claimed to be an expert on human hair, and their expertise went no further than having performed experiments on 130 hair samples.[224] "Both concede[d] that their test results [had] not been duplicated elsewhere and neither was able to cite any authority in the field [of mass spectrometry] in support of their positions."[225] There were no standards with which to test the accuracy of the measurements, and, most troubling to the court, the witnesses had made no attempt "to match the test samples against a statistically valid test group."[226]

The legal analysis in *Brown* avoids the problem of turning reliability decisions over to experts. It also avoids the self-validation problem illustrated by *State v. Washington*[227] and the labeling problem illustrated by *United States v. Franks*.[228] It still falls short of the proposed framework, however, because the four-factor test does not address adequately the distinction between reliability and validity. The court groped in this direction, but, by not jettisoning the baggage of *Frye*, failed to take the necessary final step.

B. *Alternatives to* Frye*: The Other Wrong Half*

The confusion of acceptance with reliability, compounded by the fact that the latter is the crucial issue in most forensic science cases, has led a number of commentators to attack *Frye* on a variety of grounds,[229] and

219. *See id.* at 544-45.
220. *See id.* at 554.
221. *See id.* at 554-56.
222. *See id.* at 556.
223. *See id.* at 557.
224. *See id.*
225. *Id.*
226. *Id.* at 558; *see also* United States v. Amaral, 488 F.2d 1148, 1153 (9th Cir. 1973) (articulating the standard used in *Brown*).
227. *See supra* notes 194-202 and accompanying text.
228. *See supra* notes 186-90 and accompanying text.
229. *See* Boyce, *supra* note 4, at 325-27 (suggesting the replacement of *Frye* with a

some courts have begun to pay heed. They generally purport to employ the kind of relevancy analysis advocated by Professor McCormick,[230] and their focus usually shifts to the reliability aspect of relevance, sometimes in the context of a multiple factor test. For cases involving only reliability, such as *United States v. Williams*,[231] this trend can represent a great improvement. Sometimes, however, courts simply exchange the *Frye* acceptance label for a reliability label, or they reverse the *Frye* problem and squeeze validity issues through a reliability hoop.

1. Reliability as a New Label

Like any other standard, reliability does little good when used only as a conclusory label. This problem is illustrated by *State v. Kersting*,[232] an Oregon case in which the state's intermediate appellate court explicitly rejected *Frye* for a "reasonable reliability" test.[233] The evidence at issue consisted of the microscopic comparison of the defendant's hair with hairs found on the clothing and clutched in the hand of a murder victim.[234] The defendant challenged the admissibility of the evidence on the ground that the method used was not "generally accepted as reliable."[235]

After a lengthy discussion of what it perceived to be wrong with *Frye* and right with its new test, the *Kersting* court concluded, without any discussion of the technique, that the hair comparison method, in fact, was reasonably reliable.[236] Like the court in *People v. Collins*,[237] which also admitted identification evidence based on testing of hair samples, the *Kersting* court completely failed to consider the fact that even a perfect match does not necessarily provide a specific identification.[238] Other than the fact that the method at issue in *Collins* (NAA) may have determined the composition of the hair more accurately than the method in *Kersting*, the two cases are nearly identical, except that one court used *Frye* as its label and the other used reliability.[239]

"reasonable reliability" test); Giannelli, *supra* note 4, at 1246-47 (advocating adjusting the burden of proof to accommodate the concerns *Frye* intended to address); Imwinkelried II, *supra* note 31, at 272-73 (advocating approach based more on the weight of the evidence than on admissibility); McCormick, *supra* note 4, at 911-12 (advocating multi-factor balancing test); Moenssens, *supra* note 4, at 548-55 (discussing problems with *Frye*).

230. *See supra* note 158 and accompanying text.

231. *See supra* notes 52-63 and accompanying text.

232. 50 Or. App. 461, 623 P.2d 1095 (1981), *aff'd*, 292 Or. 350, 638 P.2d 1145 (1982).

233. *See id.* at 470-71, 623 P.2d at 1101.

234. *See id.* at 465, 623 P.2d at 1098.

235. *Id.*

236. *See id.* at 471, 623 P.2d at 1101-02.

237. *See supra* notes 207-12 and accompanying text.

238. *See supra* notes 203-12 and accompanying text.

239. Indeed, the Oregon Supreme Court, which had not directly ruled on the scientific evidence issue in *Kersting*, *see* 292 Or. 350, 352, 638 P.2d 1145, 1146 (1982), *aff'g* State v. Kersting, 50 Or. App. 461, 623 P.2d 1095 (1981), subsequently required a more thorough analysis. *See* State v. Brown, 297 Or. 404, 416-18, 687 P.2d 751, 759-60 (1984) (en banc).

2. Forcing a Reliability Analysis of Validity Issues: Throwing Away Acceptance When It Is Needed

Because the distinction between validity and reliability has never been fully appreciated, some courts inappropriately have discarded acceptance. For example, in *Whalen v. State*,[240] the defendant in a rape-murder case appealed his conviction, claiming that the semen test used to establish that the victim had been raped by the defendant was not generally accepted.[241] The court found that the record contained no evidence of acceptance yet affirmed the conviction because the expert witness had described how the test was performed, how it worked, and how it would be affected by the presence of certain chemicals.[242] According to the court, this "fulfilled the hallmarks of admissibility, relevance and reliability."[243] It is difficult on these facts, however, to see how the court could reach a conclusion about reliability without considering the validity of the expert's explanation, and even more difficult to see how it could determine validity without looking at evidence of acceptance.[244]

Whalen is unusual in the way it highlights the misuse of reliability to decide an issue of validity. Because reliability subsumes validity, its misapplication tends not to create the obvious distortions that accompany the misuse of acceptance. A court that speaks in terms of reliability may be referring to the reasoning that supports a conclusion. In *United States v. Fosher*,[245] for example, expert testimony about the problems surrounding eyewitness identification was rejected for want of reliability, even though the court explicitly recognized that the expert's "mode of scientific analysis" was at issue.[246] This is precisely the sort of question that demands the acceptance test proposed in this Article,[247] but stating the decision in terms of reliability made no difference because without validity, there is no reliability.

The flaw in both *Whalen* and *Fosher* is their failure to allow for scientific concerns in reaching a decision about admissibility. Analytically, this failure mirrors the problem of using the acceptance test to determine reliability. Where forcing a *Frye* analysis in a reliability case often results in leaving legal decisions to scientists, forcing the use of reliability in a validity case often leaves too much of the decision about scientific merit

240. 434 A.2d 1346 (Del. 1980), *cert. denied*, 455 U.S. 910 (1982).

241. *See id.* at 1354.

242. *See id.*

243. *Id.*

244. In some ways, the *Whalen* case represents another manifestation of the labeling problem. Because the court saw that the testimony could not satisfy *Frye*, it latched onto reliability as a way of sustaining the conviction. *Whalen*, however, differs from *Kersting* and *Collins*. In those cases, the label covered a failure even to consider what the issue was, while in *Whalen*, the court applied the reliability test to a situation that clearly called for an evaluation of validity.

245. 590 F.2d 381 (1st Cir. 1979).

246. *Id.* at 383.

247. *See supra* notes 79-82 and accompanying text.

to judges and juries ignorant of the views of the scientific community. *Watson v. State*,[248] still another hair matching case, illustrates how this occurs. Without any consideration of the details of the underlying theory or of the reliabilty or accuracy of the technique, the *Watson* court used the relevancy test and held that whatever problems there might be could be cured by cross-examination.[249] "If . . . the witness is qualified as an expert, the testimony may be believed by the jury *despite scientific evidence to the contrary*."[250]

The *Watson* standard allows laymen to credit testimony that flies in the face of science. All that is required is a qualified expert willing to testify. The effect is the same as the self-validation seen in some *Frye* cases,[251] there being no difference between not probing scientific validity at all and allowing an expert to decide on his own what is accepted. Recalling Paul Feyerabend's acknowledgement regarding the use of subjective criteria for evaluating theories and reasoning,[252] one wonders how the court in *Watson* or *State v. Washington*[253] would react to a qualified physician who chooses to base his conclusions on the principles of a witch doctor, but who maintains that such principles are generally accepted. These two decisions indicate that neither court would limit this kind of testimony, though one doubts that either would adhere to its prior holding if actually confronted with such facts.[254]

3. Multiple Factor Forms of the Relevancy Test

While the reliability test is, in practice, the principal alternative to *Frye*, it developed as a consequence of applying the traditional relevancy standard to scientific evidence.[255] A number of those who advocate using

248. 64 Wis. 2d 264, 219 N.W.2d 398 (1974).

249. *See id.* at 272-74, 219 N.W.2d at 403.

250. *Id.* at 274, 219 N.W.2d at 403 (emphasis added); *see also* Boldt v. Jostens, Inc., 261 N.W.2d 92, 94 (Minn. 1977) (stating that the "truth of the [scientific] opinion need not be capable of demonstration").

251. *See supra* notes 194-202 and accompanying text.

252. *See supra* notes 120-22 and accompanying text.

253. *See supra* notes 194-202 and accompanying text.

254. One legal commentator, however, has maintained that treating and diagnosing physicians should be allowed to testify that an exposure to a chemical caused a plaintiff's disease, even if it is conceded that there is no accepted scientific evidence that the chemical can cause the disease at all. *See* Nesson, *supra* note 27, at 530-31. This position makes explicit the consequences of abandoning any effort at judicial control of the validity of reasoning. According to Professor Nesson, a treating or diagnosing physician always should be allowed to testify about causation unless a large number of controlled studies show that no conceivable causal connection can exist. *See id.* This proposal, of course, completely reverses the usual burden of proof.

255. *See* Giannelli, *supra* note 4, at 1232-45. As discussed in note 80, *supra*, Professor Giannelli sees reliability as the principle factor to consider in evaluating scientific evidence. He would deviate from the standard relevancy analysis only in making it explicit that the burden of proving reliability rests with the party seeking to introduce expert scientific testimony. *See id.* at 1245-49. He points out that in weighing probative value against prejudice, the probative value depends on the reliability of the technique or device at issue. *See id.* at 1247 n.379; *see also* Boyce, *supra* note 4, at 327 (advocating a reason-

one basic rule in this manner nonetheless recognize that such an approach can result in the inadequate kind of analysis seen in *Kersting*,[256] and so they have proposed various lists of factors to guide the determination of the relevancy and probative value of scientific evidence.[257] Professors Weinstein and Berger, for example, have outlined seven factors:

(1) The technique's general acceptance in the field;
(2) The expert's qualifications and stature;
(3) The use which has been made of the technique;
(4) The potential rate of error;
(5) The existence of specialized literature;
(6) The novelty of the invention; and
(7) The extent to which the technique relies on the subjective interpretation of the expert.[258]

---

able reliability test); Lederer, *Resolving the* Frye *Dilemma—A Reliability Approach*, 115 F.R.D. 84, 84-85 (1987) (proposing amendment to Rule 702 of the Federal Rules of Evidence to require that scientific evidence be reliable); McCormick, *supra* note 4, at 904 (suggesting new standards based on concerns about accuracy and reliability); Moenssens, *supra* note 4, at 564-65 (suggesting a new procedure rather than a new rule for dealing with scientific evidence, and noting that "[t]he new procedure must emphasize reliability of the technique rather than its general acceptance"); Saltzburg, *supra* note 11, at 216 (noting that "[a] foundation is needed that establishes sufficent reliability for a test to remove fears that later developments will suggest that an inaccurate and unjust result was reached at trial because the test was used"); *cf.* Starrs, Frye v. United States *Restructured and Revitalized: A Proposal to Amend Federal Evidence Rule 702*, 115 F.R.D. 92, 98-99 (1987) (proposing amendment to Rule 702 to require that scientific evidence be scientifically valid and noting that "[r]eliability has often been stated to be the linchpin without which scientific evidence would send the law reeling," and that "the difference between accuracy, validity, and reliability may be such that each is distinct from the other by no more than a hen's kick").

256. *See supra* notes 232-39 and accompanying text.

257. *See infra* note 258 and accompanying text.

258. J. Weinstein & M. Berger, 3 Weinstein's Evidence ¶ 702[03], at 702-18 to 702-19 (1987). Other commentators also have formulated lists of factors to be considered in evaluating scientific evidence. For example, Professor McCormick advocates an eleven-factor list including:

> (1) the potential error rate in using the technique, (2) the existence and maintenance of standards governing its use, (3) presence of safeguards in the characteristics of the technique, (4) analogy to other scientific techniques whose results are admissible, (5) the extent to which the technique has been accepted by scientists in the field involved, (6) the nature and breadth of the inference adduced, (7) the clarity and simplicity with which the technique can be described and its results explained, (8) the extent to which the basic data are verifiable by the court and jury, (9) the availability of other experts to test and evaluate the technique, (10) the probative significance of the evidence in the circumstances of the case, and (11) the care with which the technique was employed in the case.

McCormick, *supra* note 4, at 911-12 (footnotes omitted). In 1983, the National Conference of Lawyers and Scientists proposed consideration of, among other factors:

> (1) The reputation of the expert within the scientific community, (2) The strengths of opposing views and the standing of the persons who express them, (3) Whether the expert is prepared to discuss uncertainties in the techniques used to prepare the evidence and in the conclusions, and (4) Whether both sides to the controversy have reasonably comparable access to scientific authorities.

These seven factors capture much of the substantive flavor of this Article's proposal, but they are skewed heavily toward techniques, rather than reasoning and conclusions. In addition, the Weinstein-Berger list provides no analytical framework indicating how the factors relate to each other or to the question of admissibility or sufficiency. Despite these shortcomings, the list represents an improvement, as illustrated by the Oregon Supreme Court's use of it in *State v. Brown*,[259] to modify *State v. Kersting*.[260] *Brown* involved a defendant who had taken several lie detector tests and wanted to introduce the results into evidence at trial.[261] The request was denied, and the defendant was convicted.[262] On appeal, the court held that neither the general acceptance of *Frye* nor the reasonable reliability of *Kersting* provided a comprehensive standard.[263] Instead, it applied the Weinstein-Berger guidelines.[264] Use of these guidelines, however, still resulted in exclusion of the polygraph results when probative value was weighed against prejudicial effect.[265] The fact that general acceptance, reliability-relevance, and checklist-relevance all tend to the same outcome for polygraph evidence should not come as any surprise.[266] Courts have shown a remarkable ability to manipulate *Frye*

Symposium on Science and the Rules of Evidence, 99 F.R.D. 187, 231 (1983) (statement by Margaret Berger).

259. 297 Or. 404, 416-17, 687 P.2d 751, 759 (1984) (en banc). Similarly, in State v. Hall, 297 N.W.2d 80 (Iowa 1980), *cert. denied*, 450 U.S. 927 (1981), the court considered a variety of factors in determining that blood stain analysis evidence was reliable enough to permit admissibility and held that the determination of admissibility of scientific evidence necessarily must be made on an ad hoc basis, and that it would be impossible to establish rules applicable in every case. *See Hall*, 297 N.W.2d at 85.

260. 50 Or. App. 461, 623 P.2d 1095 (1981), *aff'd on other grounds*, 292 Or. 350, 638 P.2d 1145 (1982) (en banc). *See supra* notes 232-36 and accompanying text.

261. *See* 297 Or. at 407, 687 P.2d at 753.

262. *See id.*

263. *See id.* at 438, 687 P.2d at 772.

264. *See id.* at 416-17, 687 P.2d at 759.

265. *See id.* at 439, 445, 687 P.2d at 773, 776-77. This, of course, is the almost universal decision in polygraph cases, no matter what test is used. *See* Annotation, *Physiological or Psychological Truth and Deception Tests*, 23 A.L.R. 2d 1306, 1308 (1952 & Supps. 1982 & 1987) ("It would appear, at least in the absence of stipulation, that the courts almost uniformly reject the results of lie detector tests when offered in evidence for the purpose of establishing the guilt or innocence of one accused of a crime.").

266. *See* State v. Free, 493 So. 2d 781, 787 n.9 (La. Ct. App. 1986) (the court questioned whether, in applying a balancing test, it was in reality simply applying the general acceptance standard in a more explicit fashion), *cert. denied*, 499 So. 2d 83 (1987). A number of commentators have made the same point. See, for example, Saltzburg, *supra* note 11, in which the author points out that

> [c]ontentions that *Frye* should be replaced with a relevance analysis are calls for unnecessary reforms. *Frye* . . . is consistent with the approach that common law courts have traditionally taken toward all evidence, an approach that is carried forward in modern evidence codifications such as the Federal Rules of Evidence. . . . [I]t is not very helpful to debate the question whether *Frye* or a relevance approach to scientific evidence is preferable.

Id. at 209. The two approaches are essentially the same, despite the frequency with which they are assumed to differ. The more significant question is how much success a scientific claim must have before courts will rely on it.

to reach desired results, and lists such as the seven-factor checklist do little more than collect the various approaches. The result has been a similar skein of analytical problems for both the general acceptance test and the reliability test. Both often become labels, and both often lead to having decisions about admissibility made by the wrong people and in the wrong way.

C. *Acceptance and Reliability Combined: Two Wrong Halves Make One Right Whole*

A number of courts have recognized that acceptance relates to the theory or principle underlying an expert's conclusion,[267] and a number have differentiated reliability from acceptance.[268] Only a handful of decisions,[269] however, approximate the kind of analysis proposed in this Article, and none explicitly takes the view of validity and reliability

267. *See, e.g.*, United States v. Amaral, 488 F.2d 1148, 1153 (9th Cir. 1973) (admissibility to be determined by reference to four criteria: "1. qualified expert; 2. proper subject; 3. *conformity to a generally accepted explanatory theory*; and 4. probative value compared to prejudicial effect") (emphasis added); United States v. Brown, 557 F.2d 541, 556 (6th Cir. 1977) (citing *Amaral*); United States v. Green, 548 F.2d 1261, 1268 (6th Cir. 1977) (same); *see also* United States v. Downing, 753 F.2d 1224, 1235-37 (3d Cir. 1985) (*Frye*'s general acceptance test guards against unproven hypotheses, and is therefore one factor in a balancing test); United States v. Alexander, 526 F.2d 161, 163 (8th Cir. 1975) (assuming that *Frye* refers to scientific principles); United States v. Addison, 498 F.2d 741, 743 (D.C. Cir. 1974) (assuming without explanation that *Frye* refers to the theory from which a deduction is made); Ibn-Tamas v. United States, 407 A.2d 626, 638 (D.C. 1979) (stating that *Frye*'s general acceptance test properly focuses on methodology, not subject matter), *later appeal*, 455 A.2d 893 (D.C. App. 1983); State v. Smith, 50 Ohio App. 2d 183, 193-94, 362 N.E.2d 1239, 1246 (1976) (holding that the law of scientific evidence requires that testimony be based on a dependable principle that has gained general acceptance; derived from a proper apparatus and materials; and presented by a qualified expert; and pointing out the problem of self-validation that can result when the relevant field is defined too narrowly).

268. Most cases that reject *Frye* for a relevancy analysis in effect distinguish acceptance from reliability. *See, e.g.*, United States v. Gwaltney, 790 F.2d 1378, 1382 (9th Cir. 1986) (test used to match semen not generally accepted, but evidence shows it to be reliable), *cert. denied*, 107 S. Ct. 1337 (1987); United States v. Ferri, 778 F.2d 985, 988-90 (3d Cir. 1985) (footprint comparison admitted based on reliability), *cert. denied*, 476 U.S. 1172 (1986); United States v. Williams, 583 F.2d 1194, 1198 (2d Cir. 1978) (court determines reliability, regardless of the scientific "voting" pattern), *cert. denied*, 439 U.S. 1117 (1979); Jackson v. Garrison, 495 F. Supp. 9, 11 (W.D.N.C. 1979) (rejecting general acceptance as test for polygraph because it is at least as reliable as other forms of evidence that are admissible), *rev'd*, 677 F.2d 371, 372-73 (4th Cir.) (reversal based on ground that admissibility of polygraph was properly a state law issue), *cert. denied*, 454 U.S. 1036 (1981); Whalen v. State, 434 A.2d 1346, 1354 (Del. 1981) (evidence of reliability but no evidence of acceptance), *cert. denied*, 455 U.S. 910 (1982); State v. Kersting, 50 Or. App. 461, 471, 623 P.2d 1095, 1101-02 (1981) (holding that evidence did not support general acceptance, but appropriate standard is reasonable reliability), *aff'd on other grounds*, 292 Or. 350, 638 P.2d 1145 (1982) (en banc); *cf.* United States v. Addison, 498 F.2d 741, 744 (D.C. Cir. 1974) (applying *Frye* but distinguishing reliability from acceptance).

269. *See, e.g.*, United States v. Brown, 557 F.2d 541 (6th Cir. 1977) (discussed *supra* notes 218-26 and accompanying text); People v. Young, 425 Mich. 470, 391 N.W.2d 270 (1986) (discussed *supra* note 193); People v. Collins, 94 Misc. 2d 704, 405 N.Y.S. 2d 365

advocated here. *People v. Collins*,[270] a New York trial court decision, perhaps comes the closest.

In *Collins*, the defendants allegedly had made certain phone calls in furtherance of an extortion scheme, and some of these calls had been recorded.[271] The State obtained voice exemplars and made a pretrial motion to admit evidence based on a voiceprint comparison.[272] In deciding this motion, the *Collins* court retained the *Frye* test, but held that "the standard which must be applied to the admissibility . . . of any scientific test, is the twofold test of reliability and general scientific acceptance."[273] The court denied the motion, finding that the technique failed both prongs of the test.[274]

Although the court in *Collins* did not explicitly connect acceptance to the validity of the underlying reasoning, its acceptance analysis closely parallels this Article's proposed validity test. Citing a book about the

---

(1978) (discussed *infra* notes 270-82); Phillips *ex rel.* Utah State Dep't of Social Servs. v. Jackson, 615 P.2d 1228 (Utah 1980).

Phillips involved the use of HLA blood test results to establish paternity in an action to compel child support. The plaintiff mother had introduced the test results at trial, and the putative father appealed after a verdict against him. The Utah Supreme Court reversed, finding that there was insufficient evidence that the HLA test was reliable. *See Phillips*, 615 P.2d at 1238. The court's true focus, however, was on validity, and it articulated criteria that covered both issues. As the court explained, a sufficent foundation for admissibility of genetic evidence requires that six elements be addressed:

> (1) the correctness of the genetic principles underlying the test for determining paternity; (2) the accuracy and reliability of the methods utilized in application of the principles to determine paternity; (3) the effect of variables such as occur in persons of different nationalities or ethnic origins that would influence the accuracy of the test; (4) other factors that might tend to invalidate the test or significantly change the probability of accuracy; (5) establishing that the actual method employed and the particular test used in a given case were performed in accordance with proper procedures and with proper materials and equipment; and (6) the qualifications of the necessary witnesses.

Id. at 1235. The first criterion equates directly with validity at the theoretical level, and the second looks to the validity of the application. The four remaining criteria all relate to reliability or to foundational issues outside of the principal focus of this Article.

The real issue in *Phillips*, as in *Collins*, concerned the validity of the application. *See id.* at 1233. The evidence showed that the HLA test had been in use for about 15 years, but not necessarily for the determination of paternity. *See id.* at 1236. The plaintiff's expert apparently had testified about medical uses such as testing for tissue compatibility for purposes of organ transplantation rather than about its use for determining paternity. *See id.*

The *Phillips* court sought to distance itself from the general acceptance test of *Frye* with a rule that "[t]ests that have passed from the experimental stage may be admissible if their reliabilty is reasonably demonstrable," even if the test is not generally accepted. *Id.* at 1234. This does not really do away with acceptance, however, because passage from the stage of experiment could not be determined without consideration of acceptance.

270. 94 Misc. 2d 704, 405 N.Y.S.2d 365 (1978). It should be noted that this case has no connection to the two other *Collins* cases discussed *supra* notes 207-12 and accompanying text.

271. *See id.* at 705, 405 N.Y.S.2d at 366.

272. *See id.*

273. *Id.* at 706, 405 N.Y.S.2d at 367.

274. *See id.* at 720, 405 N.Y.S.2d at 375.

scientific method, the court found that "[i]t is certainly reasonable to expect science to withhold judgment on a new theory until it has been well tested in the crucible of controlled experimentation and study. Such a procedure would require replication of original experiments, and scrutiny of the results in various scientific journals."[275] The court also addressed the problem of defining the relevant scientific community, recognizing "that expertise in disciplines tangential to the one under consideration could have significant bearing on the issue under discussion."[276] The court noted that voiceprints were used by experts in linguistics and acoustic phonetics, and by speech scientists, psychologists and engineers,[277] and that the sound spectrograph had been developed as a linguistic tool to study language.[278] The court clearly recognized that the technique was accepted, but for purposes different from voice identification. There was only one study that supported use for identification, and it had not been replicated.[279] Thus, the court held that "[i]n this case . . . it is not surprising that so few qualified scientists accept this technique, which acceptance would necessarily be predicated on preliminary and incomplete experimentation."[280]

According to this Article's analytical framework, the *Collins* decision could have ended at this point. The issue then would have been viewed as a question of whether the identification represented a valid application of the voiceprint technique. The answer, in accordance with the court's analysis, would have been no, and this would have resolved any possible question about reliability. The court, however, undertook a separate evaluation of reliability, which became largely a repetition of the acceptance analysis, precisely because the unreliability derived from invalid reasoning. The fact that there was "no experimentation to show that two

275. *Id.* at 709-10, 405 N.Y.S.2d at 369 (citing T. Huxley, *The Method of Scientific Investigation*, in Science: Method and Meaning 6 (S. Rappaport & H. Wright eds. 1974)). Similarly, United States v. Baller, 519 F.2d 463 (4th Cir.), *cert. denied*, 423 U.S. 1019 (1975), held that

> [i]n order to prevent deception or mistake and to allow the possibility of effective response, there must be a demonstrable, objective procedure for reaching the opinion and qualified persons who can either duplicate the result or criticize the means by which it was reached, drawing their own conclusions from the underlying facts.

Id. at 466. In contrast to *Collins*, however, the court ultimately admitted the voiceprint testimony. *See id.* at 466-67.

276. *Collins*, 94 Misc. 2d at 708, 405 N.Y.S.2d at 368; *see also* United States v. Alexander, 526 F.2d 161, 164 n.6 (8th Cir. 1975) (noting that polygraph technique is based on a number of concepts and that "[e]xperts in neurology, psychiatry and physiology may offer needed enlightment upon the basic premises of polygraphy. Polygraphists often lack extensive training in these specialized sciences."); People v. Young, 425 Mich. 470, 474, 391 N.W.2d 270, 271 (1986) (noting that nonforensic scientists can evaluate forensic use of electrophoresis).

277. *See Collins*, 94 Misc. 2d at 708, 405 N.Y.S.2d at 368.

278. *See id.* at 708-09, 405 N.Y.S.2d at 368.

279. *See id.* at 709, 405 N.Y.S.2d at 368.

280. *Id.* at 710, 405 N.Y.S.2d at 369.

different voices will always appear different spectrographically,"[281] and the fact that there appeared "to be no objective criteria to determine when there is a sufficient sample"[282] go to the heart of validity. The application of objective criteria constitutes the very foundation of science.

D. *Psychiatric and Psychological Evidence: The Need for an Understanding of the Nature of Scientific Knowledge in Applying the Validity-Reliability Analysis*

The *Frye* debate has almost completely bypassed disputes about the admissibility of medical testimony,[283] but it has reached some forensic applications of psychiatry and psychology,[284] two fields closely related to medicine.[285] The case law that has developed around these applications merits separate review because it provides a bridge between medical evidence and what traditionally has been regarded as scientific evidence, and because psychiatry and psychology exemplify the kind of controversial science that most troubles the law.[286] The cases illustrate how the

---

281. *Id.* at 713, 405 N.Y.S.2d at 371.

282. *Id.* at 714, 405 N.Y.S.2d at 371. The court's conceptual confusion led to the use of the term "acceptance of reliability," but this had no real bearing on the substance of its analysis.

283. For example, in People v. McDonald, 37 Cal. 3d 351, 690 P.2d 709, 208 Cal. Rptr. 236 (1984), the court observed:

> We have never applied the *Kelly-Frye* rule to expert medical testimony, even when the witness is a psychiatrist and the subject matter is as esoteric as the reconstitution of a past state of mind or the prediction of future dangerousness, or even the diagnosis of an unusual form of mental illness not listed in the diagnostic manual of the American Psychiatric Association.

Id. at 373, 690 P.2d at 724, 208 Cal. Rptr. at 251; *see also* Boyce, *supra* note 4, at 325-27; Note, *Did Your Eyes Deceive You? Expert Psychological Testimony on the Unreliability of Eyewitness Identification*, 29 Stan. L. Rev. 969, 1022 (1977). *But see* Puhl v. Milwaukee Auto. Ins. Co., 8 Wis. 2d 343, 353-54, 99 N.W.2d 163, 168-69 (1960) (plaintiff's evidence rejected because of lack of consensus in the medical field), *overruled on other grounds*, Stromsted v. St. Michael Hosp. of Franciscan Sisters, 99 Wis. 2d 136, 142-44, 299 N.W.2d 226, 229-30 (1980).

284. *See, e.g.*, People v. Shirley, 31 Cal. 3d 18, 641 P.2d 775, 181 Cal. Rptr. 243 (discussed *infra* notes 294-303) (rejecting hypnotically enhanced recall testimony, based on *Frye*), *cert. denied*, 459 U.S. 860 (1982); People v. Bledsoe, 36 Cal. 3d 236, 681 P.2d 291, 203 Cal. Rptr. 450 (1984) (discussed *infra* note 317) (rejecting rape trauma syndrome); State v. Marks, 231 Kan. 645, 647 P.2d 1292 (1982) (discussed *infra* note 318) (accepting rape trauma syndrome); *cf.* United States v. Downing, 753 F.2d 1224 (3d Cir. 1985), *on remand*, 609 F. Supp. 784 (E.D. Pa.), *aff'd mem.*, 780 F.2d 1017 (3d Cir. 1985) (discussed *infra* notes 350-63 and accompanying text) (addressing overall legal approach to scientific evidence in context of case involving evaluation of studies of eyewitness reliability). *But see* McCord, *The Admissibility of Expert Testimony Regarding Rape Trauma Syndrome in Rape Prosecutions*, 26 B.C.L. Rev. 1143, 1182 (1985) (noting that *Frye* often is not applied to expert testimony based on the behavioral sciences).

285. *See In re* Freiburger, 153 Mich. App. 251, 257, 395 N.W.2d 300, 302 (1986) (noting that psychiatry is a branch of medicine).

286. A recent article in the Washington Post Health Supplement discusses at some length the extensive use of, and extensive problems with, psychiatric evidence. *See* Specter, *Diagnosis or Verdict? Psychiatrists on the Witness Stand*, Washington Post, July 28,

proposed validity-reliability mode of analysis is flexible enough to fit the complexity of the reasoning behind an expert's testimony, and they demonstrate the extent to which even a rudimentary grasp of the nature of scientific knowledge can facilitate an understanding of scientific issues.

At one extreme, psychological evidence based directly on an overarching explanatory theory requires a close validity analysis. Hypnotically-enhanced recall, for example, derives almost exclusively from the application of a theory about how the human mind works,[287] and the validity of that theory (and reasoning based upon it) should be the primary focus of legal inquiry. At the other extreme, testimony about psychological profiles involves almost no explanation or reasoning. A profile usually consists of little more than observed patterns of behavior,[288] and validity is a concern only when questions about the method of observation arise.

Studies of the accuracy of eyewitness identification fall in between hypnosis and profiles. These studies begin with tests of how various factors affect the way one person recalls the appearance of another. Generaliza-

---

1987, (Health Supplement) at 10, col. 1. The article notes the increasing use of expert witnesses of all kinds, and points out that none plays a more prominent role than does the psychiatrist:

> For decades, psychiatrists have participated as expert witnesses in American trials, testifying in a wide range of cases. . . . Legal and medical authorities agree that the testimony of psychiatrists can often determine the outcome of even the most public and controversial trials. But they also acknowledge that the credibility of psychiatrists on the witness stand is frequently in doubt.

Id. For a withering critique of Freudian analysis, see A. Grunbaum, The Foundations of Psychoanalysis (1984).

The trial of John Hinckley in 1982 highlighted many of the problems with psychiatric evidence. *See* Note, *Federal and Local Jurisdiction in the District of Columbia*, 92 Yale L.J. 292, 292-93 (1982). Largely in response to dissatisfaction with the finding that Hinckley was legally insane, Rule 704 of the Federal Rules of Evidence was amended in 1984. This rule generally allows experts to give testimony that embraces an ultimate issue to be decided by the trier of fact, but the 1984 amendment creates an exception under which

> [n]o expert witness testifying with respect to the mental state or condition of a defendant in a criminal case may state an opinion or inference as to whether the defendant did or did not have the mental state or condition constituting an element of the crime charged or of a defense thereto. Such ultimate issues are matters for the trier of fact alone.

Fed. R. Evid. 704(b).

Psychiatry and psychology have also been the subject of intense criticism from legal commentators. *See, e.g.*, Almy, *Psychiatric Testimony: Controlling the "Ultimate Wizardry" in Personal Injury Actions*, 19 Forum 233, 234 (Winter 1986) (pointing out that courts have not controlled the use of psychiatry in personal injury cases); Ennis & Litwack, *supra* note 80, at 719 (criticizing wide variations in psychiatric judgment and use of imprecise diagnostic categories); Morse, *supra* note 214, at 654 (advocating a reduced role for mental health professionals in judging the legal status of individuals); Comment, *The Psychologist as Expert Witness: Science in the Courtroom?*, 38 Md. L. Rev. 539, 543 (1979) ("Although the scientific bases of psychology may be adequate to sustain its validity as a social science, it is argued that in a legal context the courts should impose on psychology *qua* science the same standards for probative value that it demands of every other scientific process or technique.").

287. *See infra* Part II.D.1.

288. *See* McCord, *supra* note 284, at 1187.

tions are then derived from the test results. Both the validity of the tests and the probative value of the resulting pattern relative to a specific case therefore must be considered.

1. Hypnotically-enhanced Recall

The admissibility of hypnotically refreshed testimony has been a legal issue since at least 1904, when a New York appellate court held against a plaintiff who could not recall being seduced by the defendant until hypnosis "unlocked" her memory.[289] Exclusion remained the general rule until 1968, when the Maryland Court of Special Appeals held in *Harding v. State*[290] that a complaining witness in an assault case could testify despite the fact that she told a different story after hypnosis.[291] Discussing neither *Frye* nor the scientific basis for this use of hypnosis, the court in *Harding* simply relied on the statement of a psychologist who testified that "there was no reason to doubt the accuracy of the witness' recollections."[292]

The error of the Maryland decision became apparent as other courts undertook the scientific review from which it had abstained.[293] In *People v. Shirley*,[294] the California Supreme Court conducted a particularly searching analysis. Having considered the psychological literature on human memory in detail, the *Shirley* court cited several reasons for excluding hypnotically-induced recall.[295] Hypnosis, it found, is by its very nature "a process of suggestion," and the hypnotized subject has a "compelling desire to please the hypnotist."[296] There is also usually no way to verify the refreshed version of the facts, which may seem inordinately plausible because of the coherence of the testimony and the witness's recall of detail.[297] Finally, the increased confidence of a hypnotized wit-

289. *See* Austin v. Barker, 90 A.D. 351, 355, 85 N.Y.S. 465, 467 (1904). The case involved double hypnosis in that the plaintiff claimed the defendant had hypnotized her at the time of the seduction to make her forget it. *See id.*

290. 5 Md. App. 230, 246 A.2d 302 (1968), *cert. denied*, 395 U.S. 949 (1969).

291. *See id.* at 236, 246 A.2d at 306.

292. *Id.* at 246, 246 A.2d at 311.

293. Indeed, in a subsequent case, the Maryland Court of Appeals applied the *Frye* test and rejected *Harding*. *See* State v. Collins, 296 Md. 670, 702, 464 A.2d 1028, 1044 (1983). Legal commentary on the problems associated with hypnosis includes: Diamond, *Inherent Problems in the Use of Pretrial Hypnosis on a Prospective Witness*, 68 Calif. L. Rev. 313, 348-49 (1980) (advocating exclusion); Falk, *Posthypnotic Testimony—Witness Competency and the Fulcrum of Procedural Safeguards*, 57 St. John's L. Rev. 30, 59-60 (1982) (advocating admissibility if safeguards are used); Comment, *Hypnosis in Our Legal System: The Status of Its Acceptance in the Trial Setting*, 16 Akron L. Rev. 517, 535-36 (1983) (same); Note, *The Admissibility of Testimony Influenced by Hypnosis*, 67 Va. L. Rev. 1203, 1232-33 (1981) (same); and Note, *The Use of Hypnosis to Refresh Memory: Invaluable Tool or Dangerous Device?*, 60 Wash. U.L.Q. 1059, 1084-85 (1982) (same).

294. 31 Cal. 3d 18, 641 P.2d 775, 181 Cal. Rptr. 243, *cert. denied*, 459 U.S. 860 (1982).

295. *See id.* at 62-66, 641 P.2d at 801-04, 181 Cal. Rptr. at 270-72.

296. *Id.* at 63-64, 641 P.2d at 802, 181 Cal. Rptr. at 271.

297. *See id.* at 65-66, 641 P.2d at 803-04, 181 Cal. Rptr. at 271-72.

ness does not relate to increased accuracy or veracity.[298] Thus, the court found the foundational validity of the method wanting,[299] and testimony tainted by its use had to be excluded.[300]

The general acceptance test of *Frye* worked well in *Shirley* because hypnosis provides a paradigmatic example of the need to distinguish the validity of the reasoning underlying a method from its reliability. Proponents of hypnotically-aided recall view human memory as something like a videotape machine that may require some tinkering to get a full playback of pictures locked in the subconscious mind,[301] but there is no way of knowing what is there before the hypnotically-induced unlocking. There also is no way to distinguish between facts added or enhanced by the process of hypnosis and those genuinely uncovered by the process,[302]

298. *See id.* at 65, 641 P.2d at 803, 181 Cal. Rptr. at 272.

299. The Court wrote in terms of the unreliability of hypnosis, *see id.* at 39-40, 641 P.2d at 787, 181 Cal. Rptr. at 255, but clearly meant invalidity as that term is defined in this Article.

300. The *Shirley* court held that even testimony about facts recalled prior to hypnosis would be barred if they were covered while the witness was hypnotized, although it also held that prehypnotic testimony taken at a pretrial hearing might be admissible in lieu thereof. *See id.* at 70-71, 641 P.2d at 807, 181 Cal. Rptr. at 275. The court found that testimony given after hypnosis could not help but be colored by the hypnosis, whether or not recalled beforehand. *See id.* at 68-69, 641 P.2d at 806, 181 Cal. Rptr. at 274. Other jurisdictions, while rejecting hypnosis, have admitted such prior facts. *See* United States v. Waksal, 539 F. Supp. 834, 838 (S.D. Fla. 1982), *rev'd on other grounds*, 709 F.2d 653 (11th Cir. 1983); State *ex rel.* Collins v. Superior Court, 132 Ariz. 180, 188, 644 P.2d 1266, 1274 (1982); Merrifield v. State, 272 Ind. 579, 582-83, 400 N.E.2d 146, 149-50 (1980); State v. Koehler, 312 N.W.2d 108, 110 (Minn. 1981); People v. Hughes, 59 N.Y.2d 523, 545, 453 N.E.2d 484, 495, 466 N.Y.S.2d 255, 266 (1983). Still others allow posthypnotic testimony if the hypnotic session was conducted with certain safeguards. *See* State v. Hurd, 86 N.J. 525, 547, 432 A.2d 86, 97 (1981); State v. Armstrong, 110 Wis. 2d 555, 575, 329 N.W.2d 386, 396, *cert. denied*, 461 U.S. 946 (1983). Finally, some courts simply admit testimony after hypnosis, taking the position that any problems go to weight, not admissibility. *See* United States v. Awkard, 597 F.2d 667, 669 (9th Cir.), *cert. denied*, 444 U.S. 885 (1979); Clark v. State, 379 So. 2d 372, 374-75 (Fla. Dist. Ct. App. 1979). Counting the *per se* inadmissibility rule of *Shirley*, this comes to four options, all of which are discussed quite thoroughly in State v. Collins, 296 Md. 670, 681-700, 464 A.2d 1028, 1034-43 (1983); *see generally* Annotation, *Admissibility of Hypnotic Evidence at Criminal Trial*, 92 A.L.R. 3d 442 (1979). Given the problems outlined in *Shirley* and *Collins*, the *per se* rule seems the most appropriate. *But see* Rock v. Arkansas, 107 S. Ct. 2704, 2714-15 (1987) ("Arkansas' *per se* rule excluding all posthypnosis testimony infringes impermissibly on the right of a defendant to testify on his or her own behalf.").

301. *See Shirley*, 31 Cal. 3d at 57, 641 P.2d at 798, 181 Cal. Rptr. at 266; *see also* Buckhout, *supra* note 106, at 23 (pointing out the fallacy of viewing memory as a tape recorder, and that this version of how memory works "reflects a 19th Century view of man as perceiver, which asserted a parallel between mechanisms of the physical world and those of the brain"); Note, *The Admissibility of Testimony Influenced by Hypnosis*, 67 Va. L. Rev. 1203, 1208-09 (1981) (noting that some writers have erroneously concluded that "all sensory impressions are recorded and stored permanently in the memory, and that hypnosis facilitates the retrieval of these impressions").

302. *See* Diamond, *supra* note 293, at 314 ("[O]nce a potential witness has been hypnotized for the purpose of enhancing memory his recollections have been so contaminated that he is rendered effectively incompetent to testify.").

nor is there any way directly to determine probability of correctness.[303] If the theory about memory as a videotape were valid, hypnotically-enhanced recall would meet at least the first criterion of the proposed analytical framework, but based on an absence of scientific acceptance for any purpose,[304] it fails even this test.

When courts in hypnosis cases begin to think in terms of reliability, they quickly get into trouble. In *State v. Armstrong*,[305] the Wisconsin Supreme Court strayed from the question of reasoning and looked instead to a "reasonable degree of scientific certainty,"[306] as evidenced by the testimony of the hypnotist. Thus misdirected, the court held that, with safeguards, recall induced by hypnosis could be admissible.[307] Other courts also have focused on assurances of reliability without considering foundational validity.[308]

2. Profile or Syndrome Evidence

For over twenty years, courts have struggled with the admissibility of "profile" or "syndrome" evidence,[309] which is based on the idea that behavior sometimes follows predictable patterns in response to certain stresses or circumstances.[310] Rape trauma syndrome provides an excellent example of profile evidence, and of the problems it has posed for the law. This syndrome, which develops in certain rape victims, is a type of posttraumatic stress disorder, meaning that its central feature "is the development of characteristic symptoms after a psychologically traumatic incident . . . beyond the range of ordinary human experience."[311] The syndrome is thus little more than a systematized collection of observations.[312]

Because of the myths that cloud the average person's understanding of how rape affects its victims,[313] courts generally have admitted testimony

303. *See id.* ("After hypnosis the subject cannot differentiate between a true recollection and a fantasy or a suggested detail. Neither can any expert [n]or the trier of fact.").

304. *See* Buckhout, *supra* note 106, at 23 (indicating that modern researchers in the area of eyewitness identification reject the mechanistic videotape version of memory).

305. 110 Wis. 2d 555, 329 N.W.2d 386, *cert. denied*, 461 U.S. 946 (1983).

306. *Id.* at 565, 329 N.W.2d at 391.

307. *Id.* at 570, 329 N.W.2d at 394.

308. *See supra* note 300.

309. *See* Note, *The Unreliability of Expert Testimony on the Typical Characteristics of Sexual Abuse Victims*, 74 Geo. L.J. 429, 448 (1985).

310. *See* E. Cleary, *supra* note 158, § 206, at 634 (profiles are based on studies that "show a correlation between certain traits or characteristics and certain forms of behavior").

311. Annotation, *Admissibility, at Criminal Prosecution, of Expert Testimony on Rape Trauma Syndrome*, 42 A.L.R. 4th 879, 883 (1985).

312. *See* McCord, *supra* note 284, at 1187 (noting that "[a]ll of behavioral science consists of studying human reactions, attempting to find patterns in those reactions, giving names to those patterns, and thereafter examining particular individuals to see if their symptoms coincide with the symptoms of other people in the past whose symptoms have given rise to a category of behavior.").

313. *See* People v. Bledsoe, 36 Cal. 3d 236, 247-48, 681 P.2d 291, 298, 203 Cal. Rptr.

about rape trauma syndrome when offered to explain post-rape behavior that otherwise might be thought inconsistent with a victim's story. In *Delia S. v. Torres*,[314] for example, a psychologist was allowed to testify that victims often delay reporting attacks.[315] The defendant had cited the plaintiff's delay in reporting the incident as evidence that an alleged sexual assault had not really taken place.[316]

The cases split, however, when an expert has gone beyond such general observations about rape victim behavior and testified that the alleged victim's conduct evidences facts related to the occurrence of a specific rape, such as a lack of consent. Some courts have applied the *Frye* test to exclude specific comparisons between a victim's behavior and the rape trauma syndrome pattern,[317] while at least one court, also applying *Frye*, has held such comparisons admissible.[318] Maryland's highest court,

---

450, 457 (1984); Massaro, *Experts, Psychology, Credibility, and Rape: The Rape Trauma Syndrome Issue and Its Implications for Expert Psychological Testimony*, 69 Minn. L. Rev. 395, 402-03 (1985).

314. 134 Cal. App. 3d 471, 184 Cal. Rptr. 787 (1982).

315. *See id.* at 479, 184 Cal. Rptr. at 792.

316. *Torres* was a civil action for battery, rather than a criminal action for rape. The victim's embarrassment and shame had kept her from reporting the attack. *See id.* at 476, 184 Cal. Rptr. at 790. In addition to testimony about the typical reactions of rape victims, the court admitted testimony about the typical characteristics of rapists, which seems a poor decision in view of the imprecision of profiles and the potential prejudicial effect of such evidence. *See id.* at 479-80, 184 Cal. Rptr. at 792.

Other courts also have admitted testimony concerning the typical reactions of rape victims. *See* Terrio v. McDonough, 16 Mass. App. Ct. 163, 175-76, 450 N.E.2d 190, 198 (in civil action for sexual assault, court admitted testimony that it was not remarkable for a victim to return to the scene of the attack, or to feel safe with her attacker after the attack had occurred offered to explain victim's return to her attacker after rape when she discovered she had left purse and keys in his apartment), *review denied*, 390 Mass. 1102, 453 N.E.2d 1231 (1983); State v. Myers, 359 N.W.2d 604, 606 (Minn. 1984) (involving a sexually abused child); State v. Middleton, 294 Or. 427, 432-35, 657 P.2d 1215, 1217-19 (1983) (involving a minor who gave inconsistent post-rape stories); *see also* People v. Bledsoe, 36 Cal. 3d 236, 247-48, 681 P.2d 291, 298, 203 Cal. Rptr. 450, 457 (1984) (citing *Torres*, *Middleton* and *Terrio*, with approval).

317. *See, e.g.*, *Bledsoe*, 36 Cal. 3d at 251, 681 P.2d at 301, 203 Cal. Rptr. at 460 (finding that rape trauma syndrome had been developed for purposes of treatment, not for determining if rape took place, court concluded "expert testimony that a complaining witness suffers from rape trauma syndrome is not admissible to prove that the witness was raped"); State v. Saldana, 324 N.W.2d 227, 230 (Minn. 1982) ("Rape trauma syndrome is not a fact-finding tool, but a therapeutic tool useful in counseling. . . ."); State v. Taylor, 663 S.W.2d 235, 240 (Mo. 1984) (en banc) (expert's "statements that prosecutrix suffered from rape trauma syndrome and that she had been raped are not sufficiently based on a scientific technique, which is either parochially accepted or rationally sound").

318. *See, e.g.*, State v. Marks, 231 Kan. 645, 654, 647 P.2d 1292, 1299 (1982). *Marks* states:

> An examination of the literature clearly demonstrates that the so-called 'rape trauma syndrome' is generally accepted to be a common reaction to sexual assault. . . . As such, qualified expert psychiatric testimony regarding the existence of rape trauma syndrome is relevant and admissible in a case such as this where the defense is consent.

Id. (citations omitted); *cf.* State v. Kim, 64 Haw. 598, 601, 645 P.2d 1330, 1334 (1982) (allowing psychiatrist to testify about truthfulness of 13-year-old complaining witness; *Frye* issue not addressed).

viewing the issue as medical rather than scientific, also has allowed testimony based on comparison.[319] In overruling the state's intermediate appellate court, which had relied on *Frye*, the Maryland Court of Appeals held that general acceptance constituted "an unreasonably high standard for the admissibility of medical opinion evidence,"[320] and that "causes of emotional disturbances are complicated medical questions."[321]

In fact, all of the specific comparison cases miss or obscure the real issue, which is not complex, not medical and not related to acceptance or validity. The real concern is whether the rape trauma syndrome pattern is unique enough to be a reliable indicator of consent or occurrence for legal purposes. There is little doubt that the syndrome reflects the actual behavior of many rape victims, and little disagreement about the reasoning that relates the pattern to a specific case. As one commentator has put it:

> The theory of relevance for rape trauma syndrome evidence on the issue of consent is simple and straightforward. Women who engage in consensual intercourse do not experience rape trauma syndrome. Most women who are forced to engage in nonconsensual intercourse do experience rape trauma syndrome. Thus, if the complainant is experiencing rape trauma syndrome, it is because she was raped.[322]

For the commentator, a proponent of admissibility, this elementary logic means that "[t]he probative value of the evidence seems patently obvious,"[323] but his conclusion follows only if the syndrome comprises a precise enough symptomology to mark the victim clearly. Recognizing this requirement, another advocate of admissibility refers to the syndrome symptoms as "psychological 'bruises.' "[324] If she is correct that the symptoms "are similar but *not* identical to other posttraumatic stress-disorder symptoms," and distinguishable from reactions to other life stress,[325] her argument for allowing relatively broad use of rape trauma syndrome evidence would be well taken. To the extent that courts have addressed this question, however, they have not found that the symptoms differentiate precisely enough to provide legally reliable evidence.

In *People v. Bledsoe*,[326] the California Supreme Court pointed out that "rape trauma syndrome . . . does not consist of a relatively narrow set of criteria or symptoms whose presence demonstrates that the client or patient has been raped; rather, . . . it is an 'umbrella' concept, reflecting the

319. *See* State v. Allewalt, 308 Md. 89, 517 A.2d 741 (1986).

320. *Id.* at 98, 517 A.2d at 745.

321. *Id.* at 101, 517 A.2d at 747 (quoting Johnson v. Zerivitz, 234 Md. 113, 116, 198 A.2d 254, 255 (1964)); *see also* McCord, *supra* note 284, at 1194 (arguing that a reasonable degree of medical certainty should suffice to make testimony about rape trauma syndrome admissible).

322. McCord, *supra* note 284, at 1197.

323. *Id.*

324. *See* Massaro, *supra* note 313, at 440.

325. *Id.* at 447.

326. 36 Cal. 3d 236, 681 P.2d 291, 203 Cal. Rptr. 450 (1984) (en banc).

broad range of emotional trauma experienced by clients of rape counselors."[327] The Kansas Supreme Court, on the other hand, did not address the issue at all in holding similar evidence to be admissible,[328] and the Maryland Court of Appeals obscured it by resorting to a discussion of medical opinion evidence.[329] These mixed precedents, all decided by courts that recognize *Frye*, parallel the confusion that has resulted from using the acceptance test of *Frye* in other situations where reliability is the proper focus.

3. Eyewitness Identification

Wrongful convictions resulting from misidentifications have long been recognized as a serious problem,[330] and a number of commentators have responded by suggesting that expert psychological testimony about the inaccuracy and unreliability of eyewitness identification might serve to reduce jurors' misplaced reliance on it.[331] Psychologists have devised a number of experiments to test the accuracy of recall and to determine how various factors affect it. For example, the effect of preconceptions

327. *Id.* at 250, 681 P.2d at 300-01, 203 Cal. Rptr. at 460; *see also* State v. Saldana, 324 N.W.2d 227, 230 (Minn. 1982); *cf.* State v. Brodniak, 718 P.2d 322 (Mont. 1986). *Brodniak* raises the interesting question of whether the victim suffered posttraumatic stress disorder as a result of rape, or as a result of a physical beating inflicted after consensual intercourse. *See Brodniak*, 718 P.2d at 325. Though this fact pattern highlights the imprecision inherent in inferring a lack of consent from the rape trauma syndrome profile, the Supreme Court of Montana held that expert psychological testimony on the issue of consent was admissible. *See id.* at 327.

328. *See* State v. Marks, 231 Kan. 645, 654, 647 P.2d 1292, 1299 (1982).

329. *See* State v. Allewalt, 308 Md. 89, 98, 517 A.2d 741, 745 (1986).

330. *See* Note, *supra* note 283, at 969 (" 'The identification of strangers is proverbially untrustworthy. The hazards of such testimony are established by a formidable number of instances in the records of English and American trials.' ") (quoting F. Frankfurter, The Case of Sacco and Vanzetti 30 (1927)). Concern about the problem of eyewitness identification resulted in the Supreme Court's holding in United States v. Wade, 388 U.S. 218 (1967), that a postindictment lineup was a critical pretrial proceeding at which a defendant had the right to have counsel present: "The vagaries of eyewitness identification are well-known; the annals of criminal law are rife with instances of mistaken identification." *Id.* at 228.

331. *See* Buckhout, *supra* note 106, at 23; Johnson, *Cross-Racial Identification Errors in Criminal Cases*, 69 Cornell L. Rev. 934, 959 (1984) (citing a number of prior commentators); Loftus, *Reconstructing Memory: The Incredible Eyewitness*, 15 Jurimetrics J. 188, 190 (1975) ("Since eyewitness testimony carries so much weight, it is important to find out why disorientation occurs in a witness' memory."); Note, *Helping the Jury Evaluate Eyewitness Testimony: The Need for Additional Safeguards*, 12 Am. J. Crim. L. 189, 192 (1984) ("Although eyewitness misidentification is the single most frequent cause of wrongful conviction, the American judiciary has traditionally refused to give or allow juries special guidance on the factors affecting human memory. . . .") (footnote omitted); Note, *supra* note 283, at 970-71 ("[U]nless some steps are taken to ensure that the unreliability of eyewitness evidence is brought to the attention of the trier of fact, mistaken identifications will continue to be a major source of wrongful convictions."); *cf.* Levine & Tapp, *The Psychology of Criminal Identification: The Gap from* Wade *to* Kirby, 121 U. Pa. L. Rev. 1079, 1081-82 (1973) (discussing police practices in obtaining identifications, pointing out that eyewitness unreliability "only underscores the necessity of improving the quality of pretrial identification proceedings").

has been tested by using playing cards on which the spades are red instead of black. When people are shown such cards they tend to see hearts instead of spades because they do not expect the reversal of the colors.[332] To test the propensity to fill in specious details, subjects are shown simple but incomplete geometric shapes. Then they are asked to draw what they have seen. The longer the delay between viewing and drawing, the more likely a person is to fill in a complete shape.[333]

Laboratory experiments with cards and simple shapes are, of course, a far cry from real life—a "nagging gap"[334] that researchers have sought to close with more realistic and sophisticated experiments. In one of these, a mock assault was committed on a California campus. The event occurred before one-hundred forty-one unsuspecting witnesses and was videotaped to check their recollections against reality. The resulting descriptions were quite inaccurate. Without biased prompting, far less than half of the witnesses could pick out a photograph of the attacker.[335] Based on such work, many psychologists have concluded that eyewitness identification is far less reliable than most people think, and that informing jurors of its weaknesses would be "a needed contribution to the judicial system."[336]

The law, however, has not hastened to take up this offer of psychological help. Until recently, all appellate courts that had ruled on the question had excluded expert testimony about eyewitness reliability, often without addressing the scientific issues involved.[337] In the first major decision on the admissibility of such evidence, *United States v. Amaral*,[338] the court explicitly declined to reach the question whether such testimony "was in accordance with a generally accepted theory explaining the mechanism of perception."[339] Instead, it rejected expert testimony about eyewitnesses, holding that cross-examination would reveal any uncertainty or doubt about the accuracy of an identification.[340]

The logic of *Amaral* is difficult to follow, considering that one of the

---

332. *See* Buckhout, *supra* note 106, at 25.

333. *See id.* at 27.

334. *Id.* at 29.

335. *See id.* at 29-30.

336. *Id.* at 31.

337. *See* McCord, *Syndromes, Profiles and Other Mental Exotica: A New Approach to the Admissibility of Nontraditional Psychological Evidence in Criminal Cases*, 66 Or. L. Rev. 19, 26 (1987) (noting that many courts faced with such nontraditional psychological evidence resort to "legal bromides [that] have diverted the legal system from a complete and correct analysis of such evidence").

338. 488 F.2d 1148 (9th Cir. 1973).

339. *Id.* at 1153-54.

340. *See id.* at 1153. Other courts have also taken this approach. *See, e.g.*, United States v. Brewer, 783 F.2d 841, 843 (9th Cir.), *cert. denied*, 107 S. Ct. 118 (1986); United States v. Thevis, 665 F.2d 616, 641-42 (5th Cir. Unit B), *cert. denied*, 456 U.S. 1008 (1982); United States v. Brown, 540 F.2d 1048, 1054 (10th Cir. 1976), *cert. denied*, 429 U.S. 1100 (1977). In the analogous case of United States v. Fosher, 590 F.2d 381 (1st Cir. 1979), the court held that expert testimony must present a system of analysis that can add to common understanding. *See id.* at 383. The court rejected psychological

principal psychological findings shows that a witness' assurance, confidence, and certainty do *not* correlate with accuracy or reliability.[341] Not until the Arizona Supreme Court's 1983 decision in *State v. Chapple*,[342] however, was a trial court reversed for excluding testimony about these findings. Against a particularly egregious factual background,[343] the *Chapple* trial court's exclusion of psychological testimony was held to be an abuse of discretion because the expert would have addressed "many specific variables which affect the accuracy of identification and which apply to the facts of this case."[344] The validity of the study was assumed.[345]

Taking its cue from *Chapple*, the California Supreme Court in 1984, also reversed a trial court's decision to exclude psychological evidence regarding the problems with eyewitness recall. In *People v. McDonald*,[346] which, like *Chapple*, involved facts that virtually demanded reversal,[347] the California court reviewed the scientific literature in the area, and noted that "[t]he consistency of the results of these studies is impressive, and the courts can no longer remain oblivious to their implications for

testimony on witness reliability, however, without offering any explanation of why it had failed this test; the circuit court simply deferred to the district court. *See id.* at 383-84.

341. *See* State v. Chapple, 135 Ariz. 281, 294, 660 P.2d 1208, 1221 (1983) (en banc) (noting that past event reinforcement "will often tend to heighten the certainty of identification"); People v. McDonald, 37 Cal. 3d 351, 369, 690 P.2d 709, 721, 208 Cal. Rptr. 236, 248 (1984) (en banc) (noting that one of the foremost counterintuitive aspects of eyewitness identification is "the lack of correlation between the degree of confidence an eyewitness expresses in his identification and the accuracy of that identification"); Buckhout, *supra* note 106, at 30 (noting that good witnesses tend to express less certainty than poor witnesses); Note, *supra* note 283, at 983 ("If a witness' description of a suspect becomes more detailed as the investigation proceeds . . . the witness in this situation probably unconsciously has changed the image in memory to include details subsequently acquired from, for example, newspaper reports of the event or a mug shot of the defendant.").

342. 135 Ariz. 281, 660 P.2d 1208 (1983) (en banc).

343. The appellant in *Chapple* was convicted of murder based on the testimony of two witnesses who picked his photograph out of a group of pictures more than one year after the crime. *See id.* at 285, 660 P.2d at 1212. The witnesses had seen the murderer only on the day of the killing, had smoked marijuana that day, and had no other knowledge of him. *See id.* at 290-91, 660 P.2d at 1217-18. Further, the witnesses had selected the appellant's picture only after some difficulty and hesitation. *See id.* at 291, 660 P.2d at 1218.

344. *Id.* at 293, 660 P.2d at 1220.

345. *See id.* at 292, 660 P.2d at 1219.

346. 37 Cal. 3d 351, 690 P.2d 709, 208 Cal. Rptr. 236 (1984) (en banc).

347. The defendant in *McDonald* had been convicted of murder. Seven eyewitnesses said that they had seen him commit the crime, and one said that the murderer was definitely not the defendant. *See id.* at 355, 690 P.2d at 711, 208 Cal. Rptr. at 238. The defendant was black, and the one witness who gave testimony favorable to him was black, *see id.* at 362, 690 P.2d at 716, 208 Cal. Rptr. at 243, which raised the issue of the problems associated with cross-racial identification. *See generally* Johnson, *supra* note 331. On top of this, the witnesses who had identified the defendant had not had a very good opportunity to observe him, and the defendant had seemingly strong alibi testimony that placed him in Alabama, rather than California, on the day of the crime. *See* 37 Cal. 3d at 360, 690 P.2d at 714-15, 208 Cal. Rptr. at 241-42. Nonetheless, he was convicted and sentenced to death. *See id.* at 355, 690 P.2d at 711, 208 Cal. Rptr. at 238.

the administration of justice."[348] Oddly, the court went on to hold that California's version of the *Frye* rule did not apply because testimony based on such studies is not scientific in the sense that it derives from a mechanism, instrument or procedure.[349]

Despite its disclaimer, the *McDonald* decision evidences a clear concern about the kind of validity analysis advocated in this Article. Where *Chapple* reflects recognition of the need to consider the reliability of an expert's conclusion, *McDonald* reflects concern about the validity of the expert's reasoning, yet neither case displays a thorough analysis. Both parts of the proposed framework are required, as illustrated by Judge Becker's decision in *United States v. Downing*.[350]

The appellant in *Downing*, like the appellants in *Chapple* and *McDonald*, was convicted solely on the basis of eyewitness testimony, and he also sought to introduce evidence about its unreliability. In ruling on the admissibility of such testimony, Judge Becker applied a standard that included consideration of acceptance and the "relationship to more established modes of scientific analysis . . . [and] [t]he existence of a specialized literature"—factors that "bear on the likelihood that the scientific basis of the new technique has been exposed to critical scientific scrutiny."[351] This portion of the standard is, in effect, a test for validity. Another component of the standard, "[t]he frequency with which a technique leads to erroneous results,"[352] equates with reliability as the term is used in this Article. The analysis in *Downing* led the court to find, based on the record before it, that psychological evidence about eyewitness identification satisfied Rule 702.[353] The court remanded the case, however, to allow the government to present more evidence on the issue and to permit the trial court to determine whether or not the defendant could provide an adequate "explanation of precisely how the expert's testimony is relevant to the eyewitness identifications under consideration."[354]

On remand, the *Downing* trial court conducted an evidentiary hearing at which both the defendant's expert and a psychologist retained by the government testified.[355] The court then reaffirmed its doubts about the contested evidence and reinstated the defendant's conviction.[356] Its analysis illustrates how the focus of a validity analysis may have to shift in order to follow the reasoning of an expert.

The expert testimony at the *Downing* hearing on remand left the court with serious questions about the methodology used to generate the basic

348. *Id.* at 365, 690 P.2d at 718, 208 Cal. Rptr. at 245.
349. *Id.* at 372-73, 690 P.2d at 723-24, 208 Cal. Rptr. at 250-51.
350. 753 F.2d 1224 (3d Cir. 1985).
351. *Id.* at 1238-39.
352. *Id.* at 1239.
353. *See id.* at 1241.
354. *Id.* at 1242.
355. *See* United States v. Downing, 609 F. Supp. 784, 785-86 (E.D. Pa.), *aff'd mem.*, 780 F.2d 1017 (3d Cir. 1985).
356. *See id.* at 792.

data,[357] but even more telling was the court's finding that the next step of the reasoning was wrong because extrapolating from the data to the fact pattern at issue was inappropriate.[358] The witnesses who had identified the defendant had seen him for five to forty-five minute periods in the course of business dealings,[359] which meant they had good reason to pay attention to him. The studies, on the other hand, involved viewing periods of one minute or less,[360] with observers who were more uninterested.[361] The defendant's expert even conceded that there were "all the ingredients of a good memory of a person's face."[362] Thus, reasoning from the data to a conclusion about the unreliability of identifications made under the circumstances of the case was neither scientifically appropriate nor valid.[363]

4. The Need for an Understanding of the Philosophy of Science

The foregoing examples from psychiatry and psychology demonstrate how an analysis based on distinguishing reasoning from conclusions and validity from reliability can focus and clarify the legal evaluation of complex and controversial scientific evidence. They also illustrate how the problems that science poses for the law parallel and reflect the philosophical problem of defining science. As a practical matter, fixing on the proper focus for legal analysis depends on understanding how theory, observation, and confirmation form an interwoven scientific fabric.[364] Theories guide research and experimentation, and are then refined, reshaped and ultimately accepted in light of experimental results.[365] Sometimes, as in the case of hypnotically-enhanced recall, the reasoning or theoretical thread stands out as the source of difficulty.[366] For rape trauma syndrome, the observational thread is the real issue.[367] Testimony about eyewitness identification demonstrates how the law must often look at both strands.[368]

The two *Downing* decisions[369] highlight the need to consider validity apart from reliability and the need to shift the legal analysis of scientific evidence to conform to the interwoven pattern of scientific reasoning. Penetrating questions about the reasoning forced its flaws to surface. When courts shy away from such questions, they lose control of science,

357. *See id.* at 790.
358. *See id.* at 791-92.
359. *See id.* at 785.
360. *See id.* at 787.
361. *See id.* at 788.
362. *Id.*
363. *See id.* at 792. The same problem often appears in cases involving matching evidence. *See supra* notes 203-12 and accompanying text.
364. *See supra* notes 105-07 and accompanying text.
365. *See supra* notes 131-41 and accompanying text.
366. *See supra* Part II.D.1.
367. *See supra* Part II.D.2.
368. *See supra* Part II.D.3.
369. *See supra* notes 350-63 and accompanying text.

and verdicts completely at odds with reality may result. Nowhere is this more apparent than in toxic tort cases involving controversial medical testimony.

III. THE PROBLEM OF MEDICAL EVIDENCE

In most cases medical testimony relates to the legal issue of causation, and the usual test for admissibility or sufficiency is "reasonable medical certainty,"[370] rather than acceptance or the reliability aspect of relevancy. The certainty test involves the qualifications of the witness and the factual basis for his or her testimony, in addition to the degree of assurance he or she expresses.[371] The way courts apply the test reflects the fact that medical testimony often derives from educated judgments about specific patients rather than from the explicit development or direct application of scientific principles or theories.[372] Although they use

370. The reasonable certainty test derives from the common law. Usually it is considered the standard against which to measure the sufficiency of medical testimony, though some courts view it as a substantive fact in its own right, *see* Carpenter v. Nelson, 257 Minn. 424, 427-28, 101 N.W.2d 918, 921 (1960), or merely as a test to insure the reliability of expert testimony. *See* Noblesville Casting Div. of TRW, Inc. v. Prince, 438 N.E.2d 722, 729 (Ind. 1982). It is this kind of confusion that makes the test such a poor standard. *See infra* Part III.A.3.

For courts using reasonable certainty, or reasonable probability, to determine sufficiency, see Johnston v. United States, 597 F. Supp. 374, 412 (D. Kan. 1984), *cert. denied*, 108 S. Ct. 694 (1988); Maddocks v. Bennett, 456 P.2d 453, 457 (Alaska 1969); Murray v. Industrial Comm'n, 87 Ariz. 190, 199, 349 P.2d 627, 633 (1960); Harris v. Josephs of Greater Miami, Inc., 122 So. 2d 561, 562-63 (Fla. 1960); Bachran v. Morishige, 52 Haw. 61, 68-69, 469 P.2d 808, 813 (1970); Poweshiek County Nat'l Bank v. Nationwide Mut. Ins. Co., 261 Iowa 844, 857, 156 N.W.2d 671, 678 (1968); Hoard v. Shawnee Mission Med. Center, 233 Kan. 267, 277, 662 P.2d 1214, 1221-22 (1983); Millard v. State, 8 Md. App. 419, 426, 261 A.2d 227, 231 (Ct. Spec. App. 1970); Larson v. Johns-Manville Sales Corp., 427 Mich. 301, 317, 399 N.W.2d 1, 8 (1986); Rael v. F & S Co., 94 N.M. 507, 511, 612 P.2d 1318, 1322 (1979); *cf.* Zerr v. Trenkle, 454 F.2d 1103, 1106 (10th Cir. 1972) (trial court preferred "reasonable medical certainty" over "reasonable medical probability"); Pucci v. Rausch, 51 Wis. 2d 513, 518-19, 187 N.W.2d 138, 141 (1971) (distinguishing medical certainty from medical probability, holding that the latter "more accurately expresses the standard").

371. The qualification and factual basis aspects of the traditional approach are reflected in Rules 702 and 703 of the Federal Rules of Evidence. *See* Fed. R. Evid. 702 & 703. The reasonable certainty test itself is most often applied to questions related to the issue of sufficiency, while qualification is viewed as more related to admissibility. *Compare* Bertram v. Wunning, 417 S.W.2d 120, 123-24 (Mo. Ct. App. 1967) (testimony admitted but held insufficient because not expressed in terms of certainty) *with* Dengler v. State Farm Mut. Ins. Co., 135 Mich. App. 645, 649-50, 354 N.W.2d 294, 296 (1984) (testimony excluded because expert not qualified). The factual basis falls in between. For the judicial control of scientific evidence, however, these differences are not always important. *See infra* Part IV.D.

372. See, for example, Prokop v. Houser, 245 Iowa 480, 62 N.W.2d 781 (1954), in which the defendant appealed the admission of a doctor's testimony concerning the extent to which the plaintiff might be able to perform normal functions with an artificial arm. *See id.* at 481, 62 N.W.2d at 782. The defendant claimed that the doctor was not qualified to testify on this point because this was not a matter of medical knowledge, and because the testimony did not show that he had any general knowledge or experience about the functioning or efficiency of artificial arms. *See id.* at 482, 62 N.W.2d at 782.

a scientific approach to reach their conclusions,[373] doctors who testify about most injuries have no explicitly articulated theory of causation.[374] The doctor who testifies that a hernia very likely resulted from an automobile accident bases his or her conclusion on a knowledge of anatomy and experience gained from treating and observing patients,[375] as does a doctor who describes an injury as resulting from a blunt force.[376]

Medical reasoning about the cause of a traumatic injury essentially depends on the doctor's familiarity with a number of factors, such as the appearance of injuries caused by differently shaped objects. The underlying reasoning is scientific,[377] but it remains submerged, as do most of the scientific concerns that courts have attempted to address with the acceptance or reliability standard in the context of forensic science. Because

---

The court, however, held that the testimony was admissible because it was based on the doctor's learning and training. *See id.* at 483, 62 N.W.2d at 783. The court reached this holding after considering the doctor's knowledge of artificial limbs, and the fact that he had handled many amputation cases. *See id.* at 482, 62 N.W.2d at 782.

373. Although the holding in *Prokop* seems sound, it should be emphasized that modern medicine does take an explicitly scientific approach. *See* L. King, Medical Thinking (1982), where it is noted that the scientific aspect of medicine

> lies not in formidable apparatus nor the myriads of available tests, nor in overflowing libraries, but in that still small voice that I call critical judgment. This voice asks the important questions: 'Do you see a pattern clearly? How good is your evidence? *How sound is your reasoning*? Can you support your inferences with the means at your disposal? What are the alternatives? What hangs on your decision?' This voice, I believe, goes to the heart of scientific medicine.

Id. at 309 (emphasis added). Also, in Schwartz, Patil & Szolovits, *Artificial Intelligence in Medicine: Where Do We Stand?*, 316 New England J. of Medicine 685 (1987), it is pointed out that

> [p]rograms based on causal, pathophysiologic reasoning also have the great virtue of leaving a trail that can be converted into an English-language explanation of their diagnostic activities. Without such explanations, it is obviously unreasonable for the physician to rely on such programs; ultimately, a program, *like any consultant*, must justify its conclusions to the physician responsible for the patient's care.

Id. at 687 (emphasis added) (footnotes omitted).

374. *See* Rheingold, *The Basis of Medical Testimony*, 15 Vand. L. Rev. 473, 486 (1962) (a doctor's reasoning "most commonly serve[s] as a [more or less assumed] substratum").

375. *See, e.g.*, Bertram v. Wunning, 417 S.W.2d 120, 123 (Mo. Ct. App. 1967).

376. *See* Commonwealth v. Hart, 348 Pa. Super. 117, 501 A.2d 675 (1985). This case is particularly interesting because it shows not only the way in which doctors exercise judgment, but also the way in which they sometimes are allowed to reach conclusions well beyond the realm of medical practice. The case involved two abused children, one of whom died and one of whom suffered serious injuries. *See id.* at 119-20, 501 A.2d at 675-76. As to the child who died, expert medical testimony was offered that she had died from injuries inflicted by some blunt force. *See id.* at 121, 501 A.2d at 677. The doctor went on to testify, however, that to a reasonable degree of medical certainty the death was the result of homicide, *see id.*, 501 A.2d at 677, a conclusion clearly beyond the scope of medical practice.

377. *See* Engle & Davis, *Medical Diagnosis: Present, Past, and Future*, 112 Archives of Internal Med. 512, 516 (1963) ("Behind each diagnosis and classification is a complicated set of attributes and theory as to their relationships involving structural, physiological, and biochemical schemes which are continually being modified and which constitute the science of medicine. Diagnosis is the foundation and skeletal framework of medical science.").

courts historically have dealt mostly with medical cases that are simple and straightforward,[378] they tend to look to a doctor's qualifications rather than to the validity of his or her reasoning,[379] and they often accept expressed certainty as an adequate indication of reliability.[380] Such surrogate tests often break down, however, even for testimony regarding run-of-the-mill traumatic injuries.[381]

For more explicitly scientific medical questions, such as whether or not exposure to a chemical caused cancer or birth defects, the breakdown becomes complete. Causation involves fundamentally scientific questions.[382] As one writer puts it, "the patient seeks relief, the physician tries to provide it, and the scientist seeks understanding."[383] The traditional reasonable certainty analysis, however, fails to address the scientific aspect of causation, especially for the kind of evidence that now frequently confronts the law in toxic tort cases. No coherent and principled scheme of analysis has emerged, and, as a consequence, the cases sometimes point in opposite directions on virtually the same facts. Toxic tort litigation, more than any other area of the law, needs the order and rationality that the proposed approach to scientific evidence would bring.

A. *The Traditional Approach to Medical Evidence*

As usually applied, the three parts of the reasonable certainty analysis generally become separate and discrete mechanisms for avoiding consideration of either the validity of an expert's reasoning or the reliability of his or her conclusions. Some courts hold testimony both admissible and sufficient based solely on a doctor's qualifications.[384] Others consider the

378. Testimony about what caused a person's death, for example, is usually relatively simple and straightforward, though there may be questions about its legal significance. For example, it is generally recognized that excessive exertion can trigger a heart attack in a person who suffers from a heart condition. It is also generally recognized that arteriosclerosis predisposes a person to having a heart attack. If a person with heart disease exerts himself at work and collapses from a heart attack, there is no real scientific question about the cause of death. Both the predisposing condition and the exertion would very likely be "but for" contributing factors. The fact that the law might try to distinguish which factor was *the* cause of death has little, if anything, to do with science. *See, e.g.*, Barksdale Lumber Co. v. McAnally, 262 Ark. 379, 383, 557 S.W.2d 868, 871 (1977) (conflicting testimony on whether heart disease or work caused heart attack); Kostamo v. Marquette Iron Mining Co., 405 Mich. 105, 119-20, 274 N.W.2d 411, 417 (1979) (noting that stress will more likely trigger heart attack in person with arteriosclerosis than in other people, but that it is not possible to determine what particular stress caused what particular injury).

379. *See infra* Part III.A.1.

380. *See infra* Part III.A.3.

381. *See, e.g.*, Pike County Highway Dep't v. Fowler, 180 Ind. App. 438, 388 N.E.2d 630 (1979) (discussed *infra* notes 421-28 and accompanying text); Bertram v. Wunning, 385 S.W.2d 803 (Mo. Ct. App. 1965), *later appeal*, 417 S.W.2d 120 (Mo. Ct. App. 1967) (discussed *infra* notes 431-36 and accompanying text).

382. *See* L. King, *supra* note 373, at 191 (the search for causes is the "essence of science").

383. *Id.* at 131.

384. *See infra* notes 389-402 and accompanying text.

factual basis for a conclusion, but ignore the connective reasoning that links facts to conclusions.[385] Still others use the degree of expressed certainty as the sole criterion.[386] This piecemeal approach distorts legal analysis, especially for the testimony of witnesses with experience in tailoring their testimony to legal requirements.[387]

1. Qualification

The qualification of any expert witness includes both his or her field of expertise and the level of training and experience he or she has within that field. For the kind of forensic science around which the *Frye* debate over acceptance versus reliability has developed, the field itself tends to be a central issue, while for medical testimony, the question usually is the training of the individual expert. No matter what the education or experience of a polygraph expert, his or her testimony generally would be inadmissible under any rule.[388] Implicit in almost every case involving human injury or disease, however, is an assumption that medicine, as a field of knowledge, must have relevant evidence to offer.

The deference accorded physicians by courts and lawyers doubtless derives from the rigorous education and licensing of medical doctors. The law presumes that licensed physicians are qualified experts and, in practice, looks for little more by way of qualification.[389] Most courts take the view that a doctor need not be a specialist in the particular branch of his profession involved in a case in order to present testimony.[390] Thus, testimony from a cardiologist about a plaintiff's neurosis (albeit about cardiac problems) has been admitted,[391] as has testimony from a dentist about the difference between the pain of neuralgia and the pain from an infection,[392] and testimony from a general practitioner on eye problems.[393] Even in medical malpractice cases, where the issue frequently is the standard of practice within a specialty rather than causation, doctors usually are allowed to testify about areas outside their own fields of specialization.[394] Indeed, for a court to find a doctor unquali-

385. *See infra* notes 413-28 and accompanying text.

386. *See infra* notes 431-36 and accompanying text.

387. *See supra* note 3.

388. *See supra* note 265 and accompanying text.

389. *See, e.g.*, Kelly v. Carroll, 36 Wash. 2d 482, 490-91, 219 P.2d 79, 84-85, *cert. denied*, 340 U.S. 892 (1950).

390. *See generally* Annotation, *Competency of Physician or Surgeon as an Expert Witness as Affected by the Fact That He Is Not a Specialist*, 54 A.L.R. 860, 861 (1928) ("[B]y the great weight of authority, a physician or surgeon is not incompetent to testify, as an expert, merely because he is not a specialist in the particular branch of his profession involved in the case; although this fact may be considered as affecting the weight of his testimony."); *see also* Rose, *A Pragmatic Approach to Medical Evidence and the Lawsuit*, 5 U. Tol. L. Rev. 237, 249 (1974) ("[I]n most jurisdictions, the possession of the status of physician is the sole prerequisite to giving expert medical testimony.").

391. *See* Heinze v. Heckler, 581 F. Supp. 13, 14 (E.D. Pa. 1983).

392. *See* Sinclair v. Haven, 198 Wash. 651, 662, 89 P.2d 820, 825 (1939).

393. *See* Valmas Drug Co. v. Smoots, 269 F. 356, 359 (6th Cir. 1920).

394. *See generally* Annotation, *Competency of General Practitioner to Testify as Expert*

fied, it must as a rule be confronted with overwhelming evidence, such as a virtual admission from the witness.[395]

Some courts have accepted testimony explicitly acknowledged to be weak because of the trust they place in doctors. In *Boldt v. Jostens, Inc.*,[396] for example, a doctor testifying about a plaintiff's lung and kidney disease acknowledged that its etiology[397] was unknown.[398] He said that the disease was immunologic, but "that the antigen to which a victim reacts 'can probably be many different things and different for different people' and that it is not known whether the reaction results from one exposure to an antigen or from multiple exposures."[399] Nonetheless, he opined that exposure to glue fumes at the defendant's factory "had a great deal to do with [the plaintiff's] illness, and certainly caused aggravation."[400] In sustaining a workers' compensation award based on this evidence, the Supreme Court of Minnesota simply held that "the truth of the opinion need not be capable of demonstration."[401] Even more extreme are cases in which doctors are allowed to testify about matters completely outside their profession. For example, forensic pathologists, who examine gunshot wounds, often have been allowed to testify about

---

Witness in Action Against Specialist for Medical Malpractice, 31 A.L.R. 3d 1163, 1166 (1970) (noting that "while there is authority to the contrary, it may be reasonably concluded . . . that a general practitioner will not be held incompetent to testify in an action against a medical specialist solely on the ground that he is not also a specialist in that field"). Thus, in Baerman v. Reisinger, 363 F.2d 309 (D.C. Cir. 1966), the trial court was reversed for having refused to allow a general practitioner to testify about the standard of care for a cardiologist. The plaintiff had tried to introduce the testimony of the general practitioner to establish that the cardiologist should have diagnosed hypothyroidism. *See id.* at 310. The Court of Appeals for the District of Columbia noted that "it is settled law that '[a] physician is not incompetent to testify as an expert merely because he is not a specialist. . . .' " *Id.* (quoting Sher v. De Haven, 199 F.2d 777, 782 (D.C. Cir. 1952), *cert. denied*, 345 U.S. 936 (1953)).

395. *See, e.g.*, United States v. Zink, 612 F.2d 511, 514 (10th Cir. 1980) (defendant attempted to introduce testimony of a psychiatrist who had not completed his training); Arnold v. Loose, 352 F.2d 959, 962-63 (3d Cir. 1965) (testimony that diabetes caused automobile driver's death given by a doctor who, by his own admission, had never even read a text on diabetes or diabetic comas, was excluded in wrongful death action); State v. Adams, 481 A.2d 718, 728 (R.I. 1984) (trial court improperly permitted witness, who was not trained in forensic dentistry, to give forensic testimony about bite mark evidence in a murder trial when he conceded that such training was required).

396. 261 N.W.2d 92 (Minn. 1977) (per curiam).

397. Etiology relates to all the factors that contribute to a disease or abnormal condition. *See* Webster's Third New International Dictionary of the English Language Unabridged 782 (1961).

398. *See* 261 N.W.2d at 93.

399. *Id.*

400. *Id.*

401. *Id.* at 94; *see also* McGrath v. Irving, 24 A.D.2d 236, 238, 265 N.Y.S.2d 376, 378 (1965) (holding that, though expert doctor's knowledge of the cause of cancer was limited, "the medical qualifications of the plaintiff's expert were such as to entitle him to give an opinion"); *cf.* City of Seymour v. Industrial Comm'n, 25 Wis. 2d 482, 491-92, 131 N.W.2d 323, 328 (1964) (expressing doubt about testimony of doctor, but concluding that "[w]e cannot . . . hold as a matter of law that [the] expert medical testimony is incredible because contrary to scientific facts or knowledge.").

how guns work and offer opinions about the caliber of deformed bullets.[402]

Most courts do recognize that experts other than physicians can provide highly relevant testimony about medical questions and that sometimes they may be the most qualified witnesses.[403] In such cases, however, admissibility almost always is predicated upon the match between the field of expertise and the facts of the case. In evaluating this match, courts may engage in an analysis strikingly similar to that sometimes undertaken when general acceptance is considered in forensic science cases, thereby highlighting the fact that the qualification test is often a substitute for reviewing the validity of reasoning. In *Sandow v. Weyerhaeuser Co.*,[404] for example, the court reviewed the literature on clinical psychology to determine what the field covered before holding that a clinical psychologist was qualified to testify about depression allegedly caused by a head injury.[405]

Dengler v. State Farm Mutual Insurance Co.[406] illustrates how using a qualification test instead of a validity test can obscure and confuse the real issue in a case. In *Dengler*, the court rejected the attempt by an internist to establish that a fatal brain hemorrhage had been caused by a minor head injury that had occurred a year before death.[407] The witness had used the technique of differential diagnosis, which involves listing all possible causes and then eliminating those not present in the case in question.[408] Though the reasoning behind assembling the list and striking items from it was obviously the central issue, the court evaded it by finding that it "would have required an expert in neurology to take the next step and rule out all but hydrocephalus."[409] The witness had admitted his lack of qualification in this area,[410] which facilitated the court's decision, but what if he had not been so candid, or if a neurologist had given the same testimony? Deciding the relevant field would then provide only the context for a review of validity, a step that would never be reached if the issue were viewed purely as a matter of qualification.

Merely looking at an expert's qualifications, therefore, proves a poor surrogate for reviewing his or her reasoning, especially in view of the judicial reluctance to require that medical doctors confine their testimony to their specialty. Courts should regard medical fields of specialization as sources of standards by which to judge validity. Just as in *Frye-*

402. *See* Moenssens, *supra* note 4, at 558 n.60.

403. *See, e.g.*, Backes v. Valspar Corp., 783 F.2d 77, 78-79 (7th Cir. 1986) (chemist might be more competent to give opinion on causation of disease than doctor who examined the plaintiff).

404. 252 Or. 377, 449 P.2d 426 (1969).

405. *See id.* at 381, 449 P.2d at 428.

406. 135 Mich. App. 645, 354 N.W.2d 294 (1984) (per curiam).

407. *See id.* at 647, 649, 354 N.W.2d at 295, 296.

408. *See id.* at 649, 354 N.W.2d at 296.

409. *Id.*

410. *See id.*

type forensic cases, when validity is the issue, the best evidence is acceptance. For medical testimony, the application of this test requires reference to the literature and practice of the appropriate area of medical science. An opinion from a neurologist that does not conform to the accepted practices of neurology should not be admissible solely because of the expert's qualification in that field. Courts that use the *Frye* acceptance test intelligently recognize that several fields may contain useful insights into the validity of an expert's reasoning, and the same rationale may underlie the readiness to qualify medical experts without much regard to their field of specialization. Unfortunately, in medical cases, the analysis often ends with a determination of qualification.

2. Basis

When decisions about the admissibility or sufficiency of medical testimony move beyond qualification, the next consideration is often the underlying factual basis, which usually becomes another analytically inappropriate end point. For example, courts prefer testimony based on the personal observations of a treating physician.[411] Subjective information obtained from a patient is more suspect, especially when the patient provides the information to the doctor for purposes of litigation.[412]

The legal distinction between a doctor who examines a patient and takes his or her history in order to prepare for trial and a doctor who testifies after providing treatment results from a common sense concern about bias, but when the preference for treatment and observation becomes a substitute for further analysis, it can fail miserably. In *Baltimore Transit Co. v. Smith*,[413] a doctor who testified at trial first saw the plaintiff almost a year after he had been injured in a collision with a bus belonging to the defendant.[414] The doctor, who admittedly knew nothing of the plaintiff's condition before the collision, testified at trial that a long list of ailments had resulted from the accident, including headaches, neck pains, hearing loss, arm pains, a cramped hand, a sore chest, and difficulty in breathing.[415] He was unaware (because the plaintiff had not told him)[416] that the plaintiff had been discharged from the military for psychoneurosis, and while in the Army had "suffered from 'shortness of

411. *See, e.g.*, McGuire v. Davis, 437 F.2d 570, 572 (5th Cir. 1971) (noting "well-settled proposition that a physician who has examined an injured party may describe what he has seen and give his expert inferences therefrom," and holding that doctor could testify, based on his observation of plaintiff that she had felt pain during his examination of her) (footnote omitted).

412. *See, e.g.*, Fox v. Raftery, 45 A.D.2d 723, 723-24, 356 N.Y.S.2d 341, 342-43 (1974) (holding nontreating physician's testimony based upon plaintiff's past history and other medical reports inadmissible because the history and reports were not part of the trial record).

413. 252 Md. 430, 250 A.2d 228 (1969).

414. *See id.* at 431-32, 250 A.2d at 229.

415. *See id.* at 433, 250 A.2d at 229-30.

416. *See id.* at 436, 250 A.2d at 231.

breath' and 'a nervous heart.' "[417] The doctor apparently gave little explanation of his reasoning, except to say that the plaintiff's complaints and unspecified physical findings were "consistent" with the kind of injury sustained in the accident.[418]

After a plaintiff's verdict at trial, the defendant in *Baltimore Transit* appealed, contending that no "proper foundation was laid which would enable the doctor to establish a causal connection."[419] The Maryland Court of Appeals affirmed, noting that the doctor had been approached for treatment, and that "[t]hrough personal examination he had knowledge of [the plaintiff's] injuries. Had he failed to consider any essential facts in forming his opinion . . . then the weight of his testimony might be weakened, but its admissibility would not have been destroyed."[420] By looking only at the fact that the witness had examined the plaintiff, the court overlooked the gaps in the reasoning connecting his observations to his conclusions. Because nonspecific symptoms like headaches are consistent with many causal explanations, the expert's conclusion amounted to nothing more than a preliminary speculative hypothesis.

The predilection for observation and treatment sometimes leads courts to base decisions on these factors even when an expert does explicitly lay out his reasoning and the issue clearly concerns its validity. In *Pike County Highway Department v. Fowler*,[421] the plaintiff suffered a work-related injury to his foot, which eventually had to be amputated.[422] When he was awarded workers' compensation, the defendant employer appealed, claiming that an adequate foundation had not been laid for the testimony of the osteopath who had initially treated the plaintiff.[423] Though the general surgeon who performed the amputation thought the loss of the foot had resulted from arteriosclerotic obstructive disease,[424] the osteopath thought otherwise, expressing the opinion that trauma had caused the plaintiff's circulatory problems.[425] He also testified that he based his opinion on the temporal relationship between the trauma and the plaintiff's circulatory problems, and on his observations and examinations of the plaintiff over the years.[426]

417. *Id.* at 432, 250 A.2d at 229.

418. *Id.* at 435, 250 A.2d at 231.

419. *Id.* at 436, 250 A.2d at 231.

420. *Id.*; *see also* Commonwealth v. Vaughn, 329 Mass. 333, 335, 108 N.E.2d 559, 561 (1952) (in murder case in which medical examiner's testimony as to time of death was challenged, court held that "[i]t was not an essential requirement that the question call for the grounds or reasons upon which the opinion was based").

421. 180 Ind. App. 438, 388 N.E.2d 630 (1979).

422. *See id.* at 440, 388 N.E.2d at 632.

423. *See id.* at 444, 388 N.E.2d at 634.

424. *See id.* at 446, 388 N.E.2d at 635.

425. *See id.*, 388 N.E.2d at 636. Osteopathy is one of several forms of drugless therapy, and is considered far less scientifically valid than mainstream medicine. On the legal implications of osteopathy, see Annotation, *Liability of Drugless Practitioner or Healer for Malpractice*, 19 A.L.R.2d 1188 (1951).

426. *See* 180 Ind. App. at 443-44, 388 N.E.2d at 634.

The defendant in *Pike County* directly challenged the osteopath's reasoning, pointing out that for it to be correct, an embolism would have had to travel from the foot, through the heart and lungs, and back to the leg, and that this would have killed the plaintiff.[427] The real issue obviously was the validity of the osteopath's explanation, and it should have been argued and decided in terms of accepted medical principles. Instead, the court dismissed the reasoning as unimportant. "*Except for his subsequent explanation. . . .* [the osteopath] had already laid out a thorough factual foundation for his opinion."[428] Although *Pike County* is unusual in explicitly abstaining from an evaluation of an expert's reasoning, it serves to highlight the consequences of misplacing emphasis on the factual basis for an expert opinion.

3. Expressed Certainty

In many cases, disputes about medical testimony center on the degree of certainty expressed by an expert, yet this shift of attention from qualifications and factual basis brings no improvement.[429] "Reasonable medical certainty" usually serves as nothing more than an undefined label.[430] In *Bertram v. Wunning*,[431] for example, the issue was the admissibility of a doctor's testimony about the cause of the plaintiff's hernia, which had developed three months after she was injured in an automobile accident.[432] The doctor said there was a ninety percent chance that the accident had caused the hernia, but he would not use the magic words "reasonable medical certainty."[433] On appeal from a plaintiff's verdict, the appellate court held that this constituted mere speculation and remanded for a new trial.[434] The second time around, the witness was asked if ninety percent meant reasonable certainty to him, and he re-

427. *See id.* at 446, 388 N.E.2d at 635-36.

428. *Id.* at 445, 388 N.E.2d at 635 (emphasis added).

429. *See, e.g.*, Sears, Roebuck & Co. v. Workmen's Compensation Appeal Bd., 48 Pa. Commw. 161, 169, 409 A.2d 486, 491 (1979) (in holding medical testimony admissible because expressed with sufficient positiveness, court noted that lack of possible scientific verification does not detract from positiveness).

430. *See* Almy, *supra* note 286, at 238-39 (noting that traditional rules require that expert testimony be sufficiently certain, and that certainty standard varies with jurisdiction). By way of contrast, see Zatz & Sherwood, *Defending Speculative Injury Claims*, Toxics Law Reporter at 76 (June 17, 1987), which argues that Judge Enslen's decision to grant partial summary judgment in Stites v. Sundstrand Heat Transfer, Inc., 660 F. Supp. 1516 (W.D. Mich. 1987) demonstrates that "reasonable certainty" is not just "a pair of 'magic words.'" Zatz & Sherwood, *supra*, at 78. The summary judgment in the *Stites* case went to claims for increased risk and increased fear of cancer. Under these circumstances, the court held that a fifteen percent probability that cancer would develop was insufficient to survive a motion for summary judgment. The plaintiffs' experts could give *no* quantitative estimate of probability. *See Stites*, 660 F. Supp. at 1524-25.

431. 385 S.W.2d 803 (Mo. Ct. App. 1965).

432. *See id.* at 804 (accident occurred on April 23, 1960; hernia discovered on July 19, 1960).

433. *Id.* at 805.

434. *See id.* at 807-08.

sponded that it did.[435] For other reasons, the appellate court ordered still another trial, but on the expert opinion issue, the court made it clear that "testimony in the instant case materially differs from that [the doctor] gave at the first trial. . . . [because there was] a definite affirmance [of] 'reasonable medical certainty.' "[436]

Courts that have endeavored to avoid *Wunning*'s magic words approach to medical certainty generally have failed. The wide variation in the way they apply the concept of reasonable medical certainty makes it clear that the standard has no analytical value. Some decisions, including the two in *Wunning*, indicate that probabilistic evidence by itself cannot be reasonably certain, even if the probability is very high,[437] while others indicate that a statement of mere possibility will pass the test if properly phrased.[438] In *Maddocks v. Bennett*,[439] for example, the plaintiff claimed that improper application of hair dye had caused her to go bald. The court accepted as sufficient an expert's medical testimony that "there is a reasonable probability that the application *might* have been a cause."[440] By way of further contrast, at least one court has viewed reasonable certainty as a substantive fact that must be proved by a preponderance of evidence, though it conceded that "[c]onfusion understandably arises because the fact to be proved is, in itself, a matter of probability."[441] Still other courts have held that reasonable certainty is *not* a fact to be proved, but rather "a formulation designed to guarantee the trustworthiness or reliability of the opinion offered."[442]

Given these mixed precedents, many courts have abjured any require-

435. Bertram v. Wunning, 417 S.W.2d 120, 123 (Mo. Ct. App. 1967).

436. *Id.* at 125; *see also* McMahon v. Young, 442 Pa. 484, 485-86, 276 A.2d 534, 535 (1971) (statement that defendant's conduct probably caused plaintiff's condition is not enough; medical expert must form opinion with sufficient certainty).

437. *See, e.g.*, Johnston v. United States, 597 F. Supp. 374, 412 (D. Kan. 1984) ("A statistical method which shows a greater than 50% probability does not rise to the required level of proof."), *cert. denied*, 108 S. Ct. 694 (1988).

438. *See infra* notes 439-41 and accompanying text.

439. 456 P.2d 453 (Alaska 1969).

440. *Id.* at 458 (emphasis added); *see also* Exxon Corp. v. Fleming, 253 Ark. 798, 801, 489 S.W.2d 766, 769 (1973) (both of plaintiff's medical experts' statements that there was a possibility that an accident had caused death by triggering a pre-existing condition satisfied reasonable certainty requirement); *cf.* deMars v. Equitable Life Assurance Soc., 610 F.2d 55, 58 (1st Cir. 1979) (witness testified that "there is a reasonable certainty that [death] could have happened independently of his lung disease," but on cross-examination, stated that this meant that death had *in fact* happened independent of the underlying condition); Sears, Roebuck & Co. v. Workmen's Comp. Appeal Bd., 48 Pa. Commw. 161, 166-67, 409 A.2d 486, 488-89 (1979) (strength and positiveness of doctor's belief made testimony sufficient).

441. Carpenter v. Nelson, 257 Minn. 424, 428, 101 N.W.2d 918, 921 (1960).

442. Noblesville Casting Div. of TRW, Inc. v. Prince, 438 N.E.2d 722, 729 (Ind. 1982); *see also* Boose v. Digate, 107 Ill. App. 2d 418, 423, 246 N.E.2d 50, 53 (1969) ("[R]easonable [medical] certainty refers to the general consensus of recognized medical thought and opinion concerning the probabilities of conditions in the future based on present conditions.").

ment that a medical expert use particular phraseology.[443] Nonetheless, most courts persist in the assumption that some rule other than the traditional tort law preponderance of the evidence standard applies when judging the sufficiency of medical testimony.[444] In practice, the "other rule" may just be preponderance in disguise,[445] yet if this is the case, it only emphasizes that reasonable certainty adds nothing to the law except the opportunity for confusion.[446] It elevates form over substance, and it leads courts to deviate from established tort law principles by requiring either more than the preponderance test or less.[447]

4. The Failure of the Traditional Approach

Because the traditional approach to medical evidence avoids review of the reasoning of medical experts, courts cannot consistently and rationally resolve disputes about expert medical testimony. When equally qualified experts review the same data and arrive at opposite conclusions that they express with equal certainty, the traditional approach provides no guidance. Some cases represent a legitimate battle of the experts,[448] but in many, the testimony on one or both sides is patently unscientific.[449] Only by evaluating the validity of the reasoning process through which the experts connect the data to their conclusions can courts decide such cases in a consistent way and avoid the confusion and injustice that

443. *See, e.g.*, Sentilles v. Inter-Caribbean Shipping Corp., 361 U.S. 107, 109 (1959) ("The matter does not turn on the use of a particular form of words by the physicians in giving their testimony."); Noblesville Casting Div. of TRW, Inc. v. Prince, 438 N.E.2d 722, 729 (Ind. 1982) (citing several cases that hold that no particular phraseology is required for admissibility); Ernest v. Boggs Lake Estates, Inc., 12 N.Y.2d 414, 416, 190 N.E.2d 528, 529, 240 N.Y.S.2d 153, 155 (1963) (affirming holding that record as a whole, rather than specific words, can be sufficient evidence of required degree of certainty).

444. *See supra* note 370.

445. *See* Noblesville Casting Div. of TRW, Inc. v. Prince, 438 N.E.2d 722, 731 (Ind. 1982) ("The rule is simply a counterpart to the standard and burden of proof."); *cf.* Kostamo v. Marquette Iron Mining Co., 405 Mich. 105, 135, 274 N.W.2d 411, 425 (1979) (a workers' compensation claimant "need establish causality only by a preponderance of the evidence").

446. For example, see Hamil v. Bashline, 481 Pa. 256, 392 A.2d 1280 (1978), in which the court required the plaintiff to prove causation by a preponderance of the evidence. *See id.* at 265, 392 A.2d at 1284. The court somehow concluded, however, that testimony that there was a 75 percent chance that defendant's conduct had caused death of plaintiff's husband would be insufficient because it did not constitute reasonable medical certainty. *See id.* at 263, 273, 392 A.2d at 1283, 1288. The court relaxed what it viewed as the usual certainty test to accommodate the facts of the case, which involved a negligent failure to treat a heart attack victim. *Cf.* Poweshiek County Nat'l Bank v. Nationwide Mut. Ins. Co., 261 Iowa 844, 857, 156 N.W.2d 671, 678 (1968) (preponderance and reasonable medical certainty apparently treated as equivalent; reasonable certainty seen as recognition that absolute certainty is not required).

447. For a case requiring a higher standard, see Johnston v. United States, 597 F. Supp. 374, 412 (D. Kan. 1984), *cert. denied*, 108 S. Ct. 694 (1988). For a case setting a lower standard, see Maddocks v. Bennett, 456 P.2d 453, 458 (Alaska 1969).

448. *See supra* notes 372-78 and accompanying text.

449. *See supra* notes 396-401 and accompanying text.

result from reliance on the unfounded personal opinions of individual doctors.

In addition to obscuring issues related to the validity of reasoning, the traditional approach also tends to prevent thorough analysis of the reliability of medical conclusions. Beyond the fact that invalid reasoning cannot lead to reliable conclusions, reliance upon an expressed certainty test has the effect of turning over to doctors most of the decisions about legal sufficiency that courts should make. In most cases, it is unclear what the witness' opinion really means, and it is equally unclear how he or she arrived at it.[450]

B. *The Special Problems Created by Non-traditional Medical Evidence*

Scientific research has increased greatly our knowledge of latent diseases, such as cancer, and public concern and litigation have followed in short order.[451] Many cases have multiple plaintiffs and widespread publicity, and almost all involve complex and controversial scientific questions.[452] The validity of reasoning about causation, more than ever, often becomes the central issue, but courts continue to avoid it. Some essentially have stuck with the reasonable certainty approach,[453] which does not work very well, even for traditional medical evidence. Other courts have undertaken more searching analyses, but only a few have directly

450. This problem rarely becomes explicit. Leibowitz v. Ortho Pharmaceutical Corp., 224 Pa. Super. 418, 307 A.2d 449 (1973) (per curiam), provides an example, however. *Leibowitz* involved thrombophlebitis allegedly caused by oral contraceptives. At trial, the testimony of one of the plaintiff's doctors became so confused that the judge intervened to ask if the doctor meant to say there was "reasonable medical certainty" about causation. *Id.* at 425, 307 A.2d at 454. Thus pressed, the witness said he thought taking the contraceptives had been a "significant factor." *Id.* On cross-examination, he stated, "I refuse to use the term 'cause' in any part of my practice." *Id.* at 426, 307 A.2d at 454. He never made clear what he really thought about medical certainty.

451. This does not mean that science has established specific causal connections, as Professors Schroeder and Shapiro have noted in discussing occupational diseases:

> Articles and books about occupational disease routinely report dramatic evidence of its epidemic proportions. Data revealing that thousands die or are disabled each year are regularly cited. Documentation of specific instances of diseases that have taken terrible tolls is readily available. The usefulness of these studies is limited, however, by the nature of the data upon which they rely. There are tens of thousands of chemicals used in the workplace, thousands of which may be toxic or carcinogenic. In order to establish the specific dangers posed by these chemicals, researchers must use animal experimentation and epidemiological studies. Using these methods to correlate exposure to vapors or dusts with harmful effects is difficult. As a result, estimates of the total incidence of occupational disease are little more than a statistical house of cards.

Schroeder & Shapiro, *Responses to Occupational Disease: The Role of Markets, Regulation, and Information*, 72 Geo. L.J. 1231, 1231-32 (1984) (footnotes omitted).

452. *See* Wagner, *The New Elite Plaintiffs' Bar*, 72 A.B.A. J. 44, 44 (Feb. 1986) (noting that cases such as the Agent Orange litigation, Dalkon Shield litigation, asbestos litigation, tampon litigation, Bendectin litigation, and the Bhopal litigation "have thousands of clients, a national scope and involve complex scientific causation questions").

453. *See infra* Part III.B.1.

addressed the reasoning of the experts upon whose opinions the cases turn.[454]

1. Use of the Traditional Approach

a. Ferebee v. Chevron Chemical Co.

Two recent cases exemplify the passive acceptance of medical testimony in toxic tort litigation and the untoward consequences that can result. The first case, *Ferebee v. Chevron Chemical Co.*,[455] was decided by the United States Court of Appeals for the District of Columbia Circuit in 1984. *Ferebee* involved an agricultural worker who allegedly contracted a fatal lung disease as a result of long-term skin exposure to paraquat, an herbicide distributed by the defendant, Chevron.[456] After losing at trial, Chevron appealed, claiming that the plaintiff's evidence of causation did not support the jury's verdict.[457]

Chevron argued that all recognized ill-effects of paraquat occur within a short time of exposure, and cease when exposure ends, and that the plaintiff's illness did not come close to this pattern.[458] The plaintiff had not experienced any symptoms until ten months after he had last used the herbicide.[459] Nonetheless, the plaintiff's experts, two pulmonary specialists who had treated him, opined that paraquat had caused his illness.[460] Both admitted that cases like the plaintiff's were rare,[461] but one of them identified three other cases that he felt were "similar."[462]

Refusing to probe the validity of the experts' reasoning, the court held that

> [j]udges, both trial and appellate, have no special competence to resolve the complex and refractory causal issues raised by the attempt to link low-level exposure to toxic chemicals with human disease. On questions such as these, which stand at the frontier of current medical and epidemiological inquiry, *if experts are willing to testify* that such a link exists, it is for the jury to decide whether to credit such testimony.[463]

The case thus came down to a "classic battle of the experts, a battle in which the jury must decide the victor."[464]

---

454. *See infra* Part III.B.2. The separation of cases according to the level of judicial inquiry into the scientific evidence follows the analysis of Rothstein and Crew. *See* Rothstein & Crew, *When Should the Judge Keep Expert Testimony From the Jury?*, 1 Inside Litigation at 19 (Apr. 1987).

455. 736 F.2d 1529 (D.C. Cir.), *cert. denied*, 469 U.S. 1062 (1984).

456. *See id.* at 1531-32.

457. *See id.* at 1532.

458. *See id.* at 1535.

459. *See id.*

460. *See id.*

461. *See id.*

462. *Id.*

463. *Id.* at 1534 (emphasis added).

464. *Id.* at 1535.

The *Ferebee* court did add that the methodology used by an expert to reach a conclusion had to be sound,[465] but by this it only meant that the "use of tissue samples, standard tests, and patient examination"[466] constituted an appropriate method of diagnosis, a point that was not in dispute.[467] Though the trial court explicitly recognized that the crucial question was how the experts had concluded that causation followed from the diagnosis,[468] neither the appellate nor the trial court examined the reasoning connecting the diagnosis to this conclusion. Unlike the district court in *United States v. Downing*,[469] both *Ferebee* courts failed to shift their analysis to follow the expert's reasoning through to its final conclusion.

Ferebee reveals not only the usual legal deference to the testimony of medical doctors,[470] but also represents an open retreat from scientific standards. The court of appeals in *Ferebee* held that "[i]n a courtroom, the test for allowing a plaintiff to recover in a tort suit of this type is not scientific certainty but legal sufficiency . . . . [T]he fact that . . . science would require more evidence before conclusively considering the causation question resolved is irrelevant."[471] In effect, the court miscast science in terms of absolute certainty only to cast it aside. Instead, the court relied on each of the three elements of the traditional test. The experts were well-qualified, they based their opinions on the results of accepted diagnostic techniques, and they expressed sufficient certainty.[472] Unfortunately, this mode of analysis can produce legal verdicts completely out of keeping with scientifically established facts or scientific reasoning.

b. Wells v. Ortho Pharmaceutical Corp.

Wells v. Ortho Pharmaceutical Corp.,[473] the second recent case illustrating passive acceptance of medical testimony, shows the full consequence of stretching the *Ferebee* standard to its limits. In *Ferebee*, the court merely failed to address the scientific merit of the plaintiff's evidence, but in *Wells*, ignoring science meant the rejection of valid evi-

---

465. *See id.* at 1535-36.

466. *Id.* at 1536.

467. *See* Ferebee v. Chevron Chem. Co., 552 F. Supp. 1293, 1301 (D.D.C. 1982), *aff'd*, 736 F.2d 1529 (D.C. Cir.), *cert. denied*, 469 U.S. 1062 (1984).

468. *See id.*

469. 609 F. Supp. 784 (E.D. Pa.), *aff'd mem.*, 780 F.2d 1017 (3d Cir. 1985) (discussed *supra* notes 355-63 and accompanying text).

470. *See Ferebee*, 736 F.2d at 1535 (court referred to the plaintiff's experts as "eminent specialists," and was impressed by the fact that they had been plaintiff's "treating physicians").

471. *Id.* at 1536.

472. *See Ferebee*, 552 F. Supp. at 1296 (plaintiff's experts testified at trial that "diagnosis of paraquat poisoning was made, to a reasonable degree of medical certainty, long before [plaintiff's] death").

473. 788 F.2d 741 (11th Cir.), *reh'g denied en banc*, 795 F.2d 89 (11th Cir.), *cert. denied*, 107 S. Ct. 437 (1986).

dence favorable to the defendant. The issue in *Wells*, which was tried without a jury, was whether the plaintiff's birth defects had been caused by a spermicidal jelly used by her mother for approximately four weeks between conception and the time she discovered she was pregnant.[474] Although some preliminary epidemiological studies had suggested such a causal relationship, further research had failed to bear out this hypothesis.[475] Indeed, the trial court had found the studies inconclusive, and therefore had based its decision on the demeanor, tone, motives, biases, and interests that might have influenced each expert's opinion.[476] Finding that the plaintiff had established causation to a reasonable degree of medical certainty, the court awarded $5.1 million in damages.[477] On appeal, Ortho renewed its arguments that the scientific evidence was in its favor rather than the plaintiff's, but the Eleventh Circuit, following *Ferebee*, held that "it does not matter in terms of deciding the case that the medical community might require more research and evidence before conclusively resolving the question. What matters is that [the] factfinder found sufficient evidence of causation in a legal sense in this particular case . . . ."[478]

The impact of the decision in *Wells* is not, however, limited to the particular litigants in the case. The court's decision ran counter to the widely held scientific view that spermicides cause *no* birth defects. The case therefore has prompted the medical community to take the two *Wells* courts to task for ignoring accepted, well-established (not certain) scientific facts. Writing in the New England Journal of Medicine, two physicians from the National Institute of Child Health and Human Development noted that the *Wells* "decision took the medical community by surprise, because the overwhelming body of evidence indicates that spermicides are not teratogenic."[479] They further wrote that the plaintiff

474. *See id.* at 742-43.

475. *See id.* at 744 n.4.

476. *See* Wells v. Ortho Pharmaceutical Corp., 615 F. Supp. 262, 266-67 (N.D. Ga. 1985).

477. *See id.* at 298.

478. *Wells*, 788 F.2d at 745; *see also* Osburn v. Anchor Laboratories, Inc., 825 F.2d 908, 915 (5th Cir. 1987) (expert's opinion "need not be generally accepted in the scientific community before it can be sufficiently reliable and probative to support a jury finding") (citing Wells v. Ortho Pharmaceutical Corp., 788 F.2d 741, 744-45 (11th Cir.), *reh'g denied en banc*, 795 F.2d 89 (11th Cir.), *cert. denied*, 107 S. Ct. 437 (1986) and Ferebee v. Chevron Chem. Co., 736 F.2d 1529, 1535 (D.C. Cir.), *cert. denied*, 469 U.S. 1062 (1984)), *cert. denied*, 56 U.S.L.W. 3718 (April 19, 1988).

479. Mills & Alexander, *Occasional Notes: Teratogens and "Litogens"*, 315 The New England J. of Med. 1234, 1235 (1986). Also see 256 J.A.M.A. 3095-96 (1986), in which three letters relative to vaginal spermicides and congenital disorders are published. The first, by one of the co-authors who originally had suggested that exposure to spermicide might cause certain birth defects, noted that "our article was not corroborated by subsequent studies." *Id.* at 3095.

In the second, others of the co-authors replied to defend the study, but even they cautioned that "[o]ur original publication . . . listed a number of caveats against a causal interpretation of the findings . . . . We concluded that 'the results should be considered tentative until confirmed by other data.' " *Id.* at 3096. They went on to add that "[n]ow,

had won "despite testimony citing the considerable medical evidence that spermicides do not cause birth defects," and despite the United States Food and Drug Administration's decision that warnings about birth defects were not warranted.[480]

This kind of scientific outcry is not common, and should not be taken lightly. The two courts that decided *Wells* reached a conclusion purportedly based on science, but at odds with the scientific community. This is akin to deciding that the earth is flat, despite significant scientific opinion to the contrary. Rejecting science and passively accepting the opinions of "willing testifiers" leads to such results. When the scientific merit of a scientific conclusion is disputed, courts cannot rationally and consistently resolve the dispute without at least some consideration of scientific criteria.

2. Active Review

a. *Agent Orange Litigation*

In contrast to the passive approach of *Ferebee* and *Wells*, a number of courts have endeavored to hold medical experts to the standards of scientific practice. The best known examples come from the Agent Orange litigation, in which Judge Weinstein approved the settlement of class claims[481] and granted summary judgment against opt-out plaintiffs, largely because of inadequate evidence of causation.[482] In approving the settlement, Judge Weinstein reviewed most of the available literature on the health effects of low-level dioxin exposure, and concluded, notwithstanding expert deposition testimony to the contrary, that adequate proof was lacking.[483]

To avoid summary judgment, the opt-out plaintiffs had to overcome

---

six years after our study, the possibility of a teratogenic effect of spermicides in relation to limb reduction anomalies remains in substantial doubt." *Id.*

Still another of the original co-authors wrote in a third letter:

> No articles published subsequently have shown unequivocal evidence that spermicides are teratogenic. In retrospect, I believe our article should never have been published. In our present litigious environment, the reservations and qualifications written into a published report are often ignored, and the article is used as 'proof' of a causal relationship. It would have been much more appropriate for our findings to have gone through the more traditional process of being discussed among colleagues at scientific meetings . . . this peer review process could have altered the way the manuscript was written or might have delayed its submission until more information was available.

Id.

480. Mills & Alexander, *supra* note 479, at 1235.

481. *In re* "Agent Orange" Prod. Liab. Litig., 597 F. Supp. 740, 749 (E.D.N.Y. 1984), *aff'd on other grounds*, 818 F.2d 145 (2d Cir. 1987).

482. *In re* "Agent Orange" Prod. Liab. Litig., 611 F. Supp. 1223, 1260-63 (E.D.N.Y. 1985).

483. *Agent Orange*, 597 F. Supp. at 782-95; *see also* Lilley v. Dow Chem. Co., 611 F. Supp. 1267, 1275 (E.D.N.Y. 1985) (in rejecting expert testimony offered by plaintiff who had opted out of Agent Orange settlement, the court noted that the expert had relied improperly upon certain studies: "The authors of these studies acknowledge that more

strong epidemiologic evidence favorable to the defendants and had to offer sufficient evidence in support of their claims. They failed on both counts. Judge Weinstein recognized at the outset that sound epidemiologic studies were "the only useful studies having any bearing on causation," and that "all reliable studies of the effect of Agent Orange on members of the class so far published provide no support for plaintiffs' claims of causation."[484] He then moved on to an evaluation of the affidavits of the plaintiffs' two principal experts, Drs. Singer and Epstein.

Cutting to the core of Dr. Singer's reasoning, he found that it "amounts to this: the [plaintiffs] complain of various medical problems; animals and workers exposed to extensive dosages of [dioxin, the toxic impurity in Agent Orange] have suffered from related difficulties; therefore *assuming nothing else caused the [plaintiffs'] afflictions*, Agent Orange caused them."[485] The assumption circumvented the causation problem, but only at the expense of making the analysis "so guarded as to be worthless."[486] Dr. Epstein's reasoning was similarly flawed, because it contained "no showing that the incidence of the diseases [in question] are [sic] greater in the Agent Orange-exposed population than in the population generally."[487] In fact, his deposition testimony revealed that he "failed to apply what he himself recognized as appropriate criteria."[488]

Despite these devastating findings, Judge Weinstein held that Dr. Singer's general scientific technique had been accepted by a sufficient number of *courts* to allow judicial notice of its general acceptance.[489] He further held that the doctor's method of drawing inferences "could withstand the flexible approach to Rule 702 admissibility followed in the Second and Third Circuits,"[490] and that Dr. Epstein's testimony would also pass Rule 702 muster.[491] Thus, Judge Weinstein did not base his holding directly on the defects in their reasoning. In rejecting their testimony, he instead invoked Rule 703 of the Federal Rules of Evidence,[492] which allows expert reliance on otherwise inadmissible facts or data only if "of a type reasonably relied upon by experts in the particular field in forming

research is necessary and that no more than a suggestion or vague association may be hypothesized at present.").

484. *In re* "Agent Orange" Prod. Liab. Litig., 611 F. Supp. 1223, 1231 (E.D.N.Y. 1985) (Weinstein, J.); *see also In re* Swine Flu Immunization Prod. Liab. Litig., 508 F. Supp. 897, 907 (D. Colo. 1981) ("Where . . . the exact organic cause of a disease cannot be scientifically isolated, epidemiologic data becomes highly persuasive."), *aff'd sub nom.* Lima v. United States, 708 F.2d 502 (10th Cir. 1983) (per curiam).

485. 611 F. Supp. at 1237-38 (emphasis added).

486. *Id.* at 1238.

487. *Id.* at 1239.

488. *Id.* at 1254.

489. *See id.* at 1243. Why another court's acceptance should influence a judgment about scientific validity is unclear.

490. *Id.* at 1243.

491. *See id.*

492. For a detailed discussion of Judge Weinstein's use of Rule 703, see Carlson, *Policing the Bases of Modern Expert Testimony*, 39 Vand. L. Rev. 577, 579-83 (1986).

opinions or inferences upon the subject [at issue]."[493] He also cited Rule 403, which requires a balancing of probative value against the danger of unfair prejudice, confusion of issues, or misleading the jury.[494] The experts had relied upon questionnaire checklists for information about the plaintiffs, and on animal studies for information about the health effects of dioxin,[495] yet had ignored available epidemiologic information.[496]

The court's reliance on Rules 703 and 403 in the *Agent Orange* decision reflects the absence of clear guidance about the role of reasoning in expert testimony. Though the court obviously understood the deficiencies in how the experts had reached their conclusions, it framed its decision in terms of the bases upon which the experts relied,[497] which harkens back to the traditional approach to medical testimony. *Agent Orange* demonstrates that even without clarification, the Rules can be used effectively to bar certain scientific expert testimony,[498] but its articulated rationale tends to confuse the facts relied upon with the conclusions based on those facts. A more direct evaluation of the validity of reasoning is also possible under the current Rules,[499] but a clearly separate requirement for valid reasoning would eliminate this confusion, which has sometimes led to poor results.[500]

b. Johnston v. United States

Johnston v. United States,[501] a case involving cancers allegedly caused by low-level radiation exposure, illustrates a more direct analysis of validity. *Johnston* was tried without a jury under the Federal Tort Claims Act,[502] but in rejecting the testimony of the plaintiffs' experts, the court took pains to make it clear that it found their reasoning and conclusions deficient as a matter of law.[503] The decision concentrated on the lack of

493. Fed. R. Evid. 703.

494. *See* Fed. R. Evid. 403.

495. *See Agent Orange*, 611 F. Supp. at 1246-48.

496. *See id.* at 1248.

497. *See supra* notes 492-96 and accompanying text.

498. Other cases have also used the existing Rules effectively to exclude certain scientific evidence. *See, e.g.*, Viterbo v. Dow Chem. Co., 646 F. Supp. 1420, 1424-26 (E.D. Tex. 1986) (citing Agent Orange decision), *aff'd*, 826 F.2d 420 (5th Cir. 1987); *see also* Bendectin cases cited *infra* notes 517-26; *cf.* Marder v. G.D. Searle & Co., 630 F. Supp. 1087, 1090 & 1094 (D. Md. 1986) (excluding expert testimony under Rule 703, court noted that foundation for one expert's testimony was lacking, and that his explanation of causation was speculative and based on an untested theory), *aff'd mem.*, Wheelahan v. G.D. Searle & Co., 814 F.2d 655 (4th Cir. 1987).

499. *See infra* notes 501-05 and accompanying text.

500. *See, e.g.*, Wells v. Ortho Pharmaceutical Corp., 788 F.2d 741 (11th Cir.) (discussed *supra* Part III.B.1.b), *reh'g denied en banc*, 795 F.2d 89 (11th Cir.), *cert. denied*, 107 S. Ct. 437 (1986).

501. 597 F. Supp. 374 (D. Kan. 1984).

502. 28 U.S.C. §§ 2671-80 (Supp. IV 1986).

503. *See Johnston*, 597 F. Supp. at 383 ("[s]urely, if the United States had timely objected to [the expert] testimony, these cases would probably have been over in the course of two days"). *But see* Allen v. United States, 588 F. Supp. 247, 419-23 (D. Utah 1984)

confirmation for the conclusions, which under the circumstances rendered them unscientific:

> Anyone who has been trained in the scientific method realizes that a hypothesis is a scientist's educated speculation. The scientist then designs experiments to test his hypothesis in order to determine whether or not his speculation is correct. . . . That is how scientists learn what is fact and what is not true.[504]

When their reasoning was measured against scientific practice in this way, the plaintiffs' experts failed miserably. The court found that their

> conclusions are not supported by any fact other than that the instruments are coated with a radioactive paint and each plaintiff has cancer. . . . This Court is disappointed with the apparent fact that these so-called experts can take such license from the witness stand; these witnesses say and conclude things which, in the Court's view, they would not dare report in a peer-reviewed format.[505]

Just as in *Agent Orange*, the connective reasoning in *Johnston* came down to the coincidence of the plaintiffs' exposure and their diseases, and, just as in *Agent Orange*, this was not enough. The *Johnston* court, however, looked directly to accepted scientific practice, as evidenced by peer review, and based its decision on the expert's failure to conform with this practice.

3. Impatience with Limited Knowledge

The difference between the level of review in the *Ferebee-Wells* cases and in the *Agent Orange-Johnston* cases may reflect different views of the competence of courts to assess scientific reasoning, but another important factor is judicial unwillingness to accept the limited knowledge that science sometimes provides. To recognize, as the *Ferebee* and *Wells* courts did, that a question stands at the frontier of scientific inquiry and then to accept the testimony of those willing to go beyond the frontier seems an obvious self-contradiction. Such decisions show that distaste for science has two sides. Ignorance leads not only to deference to experts, but also to arrogance toward them.

Some commentators have maintained that for legal purposes, science does not have to be scientific, and that the law can dictate how scientists should reach their conclusions. A recent student Note concludes that "[a]t the heart of the problem presently confronted by the courts in toxic tort suits is the inability to determine causation quantitatively when

(admitting testimony from the same experts rejected in *Johnston*), *rev'd on other grounds*, 816 F.2d 1417 (10th Cir. 1987), *cert. denied*, 108 S. Ct. 694 (1988).

504. *Id.* at 393. It should be noted how the vestiges of Logical Positivism still permeate the court's opinion. For a discussion of Logical Positivism, see *supra* Part I.C.1.

505. 597 F. Supp. at 415. On the importance of peer review, see Perry v. United States, 755 F.2d 888, 892 (11th Cir. 1985) (swine flu case); Kubs v. United States, 537 F. Supp. 560, 562 (E.D. Wis. 1982) (same); *supra* note 479.

trans-scientific issues are involved."[506] "Trans-science" relates to questions that science cannot answer.[507] Normally, no answer would mean no proof, and plaintiffs would lose. For unspecified reasons, however, the Note assumes that scientists somehow can come up with the needed evidence. It thus calls for a new standard of liability: "The standard must adjust to the inability of trans-science to quantify the effects of a substance. It must also resolve or *circumvent* the evidentiary and procedural problems resulting from the inherently hypothetical, rather than factual, nature of trans-science."[508]

An article by Professor Charles Nesson[509] goes even further. It not only advocates relaxed standards but actually presumes to explain how scientists should conduct their work. Although researchers who have studied the effects of dioxin on laboratory animals are very reluctant to extrapolate their results to humans,[510] Professor Nesson believes they should. Without citing any support for his position, he maintains that "[i]f a chemical shows carcinogenic effects in a population of healthy lab animals, a doctor might anticipate similar results . . . [in] human beings."[511] He assumes the "[l]ab rats are extremely healthy, bred to be strong, without a history of disease or defect."[512] In fact, the animals used to test for carcinogenicity usually are bred to be very sensitive to chemicals, because for the purpose of regulation, it is better to err on the side of caution.[513] For a crusader impatient with the limits of science,

---

506. Note, *supra* note 86, at 431.

507. *Id.* at 431 n.20 ("questions that seemingly are part of science yet in fact transcend science").

508. *Id.* at 443 (emphasis added).

509. *See* Nesson, *supra* note 27.

510. *See* Lilley v. Dow Chem. Co., 611 F. Supp. 1267, 1273 (E.D.N.Y. 1985); *infra* note 513.

511. Nesson, *supra* note 27, at 532.

512. *Id.*

513. The difficulty of extrapolating results of mouse and rat tests for carcinogenicity has been a particularly hot topic lately for a number of prominent scientists, particularly Bruce Ames of the Department of Biochemistry at the University of California, Berkeley. *See* Ames, Magaw & Gold, *Ranking Possible Carcinogenic Hazards*, 236 Science 271 (1987), which states:

> Extrapolation from the results of rodent cancer tests done at high doses to effects on humans exposed to low doses is routinely attempted by regulatory agencies when formulating policies attempting to prevent future cancer. There is little sound scientific basis for this type of extrapolation, in part due to our lack of knowledge about mechanisms of cancer induction, and it is viewed with great unease by many epidemiologists and toxicologists. Nevertheless, to be prudent in regulatory policy, and in the absence of good human data (almost always the case), some reliance on animal cancer tests is unavoidable.

Id. (citations omitted). *See also* Ames, Testimony Before Calif. State Senate Committee on Toxics and Public Safety Management. (Nov. 11, 1985):

> All risk calculations based on rat and mouse cancer tests, both from natural and man-made carcinogens, are hypothetical: thus, they should be taken with a great dose of skepticism . . . . Of the carcinogens tested in both rats and mice in our database on potency, 42% of the chemicals were positive in the mouse and negative in the rat, or vice versa. Thus, even two closely related short-lived

neither scientific facts nor the standards accepted by scientists seem to matter.

4. The Bendectin Example: The Larger Benefits of Active Review

When courts accept qualified, willing testifiers[514] they not only lose control of scientific evidence, but they also create the kind of litigation "crapshoot" for which the law is often criticized.[515] The Bendectin cases, more than any others, illustrate how this can occur. They also demonstrate that active review of scientific evidence can prevent it.

Bendectin was a drug designed to alleviate the "morning sickness" nausea that commonly accompanies pregnancy. The drug was prescribed routinely from 1957 until 1983,[516] when litigation about birth defects drove it from the market. Because it was so widely used and because about two to five percent of all children are born with defects, whether or not their mother used Bendectin during pregnancy,[517] the pool of potential plaintiffs is large and ubiquitous. Courts in forty-nine states and the District of Columbia have had to deal with Bendectin

---

creatures, the rat and the mouse, don't predict very well for each other whether a substance is a carcinogen.

Id. at 3. Similar views are expressed by Philip Abelson:

> Many of the experiments that have been cited as proving a potential carcinogenicity of a chemical for humans have been performed on inbred strains of mice that have a natural incidence of liver tumors. In humans, there are taboos against inbreeding, which often leads to genetic impairments. Thus the use of inbred mice, though convenient experimentally, is suspect. More important is the fact of high natural incidence of liver tumors in the test mice. The usual response of these animals to massive doses of a chemical is to develop an even higher incidence of liver tumors. When this happens, the chemical is labeled a potential carcinogen in humans. It so happens that in humans primary liver cancer is rare with the exception of alcoholics and those who have suffered from hepatitis. . . . Thus extra liver tumors in a naturally tumorigenic mouse is of dubious relevance to humans.

Abelson, *Cancer Phobia*, 237 Science 473 (1987); *see also* Schroeder & Shapiro, *supra* note 451, at 1232-33 (noting problems attending extrapolation from animal studies to man); Scroggin, *Cancer-Risk Assessments*, Trial, Oct. 1987 at 49, 50 (pointing out that many scientists question the validity of extrapolating high exposure data to predict the effects of very low exposures).

514. *See* Ferebee v. Chevron Chem. Co., 736 F.2d 1529, 1534 (D.C. Cir.) ("if experts are willing to testify . . . it is for the jury to decide whether to credit such testimony"), *cert. denied*, 469 U.S. 1062 (1984).

515. Jerome Skinner, one of the plaintiffs' lawyers in the Bendectin litigation, once referred to the trial of the consolidated cases in the Southern District of Ohio, *see infra* note 528 and accompanying text, as " 'the biggest crapshoot that ever occurred in the American legal system.' " Kaufman, *Jury Concludes Bendectin Caused No Birth Defects*, The Cincinnatti Enquirer, March 13, 1985, A-1, col. 5, at A-16, col. 3.

516. *See* Letter from Dr. Frank C. Woodside, III, of Dinsmore & Shohl, a Cinncinatti law firm, to the author (Sept. 28, 1987) (available in the files of the Fordham Law Review). Dr. Woodside, who also is an attorney, has represented Merrell-Dow in a number of Bendectin cases.

517. *See* Richardson v. Richardson-Merrell, Inc., 649 F. Supp. 799, 800 (D.D.C. 1986) (two to four percent); *In Re* Richardson-Merrell, Inc. "Bendectin" Prods. Liab. Litig., 624 F. Supp. 1212, 1229 (S.D. Ohio 1985) (two to five percent).

cases,[518] and in almost every instance, the litigation has boiled down to a fight over causation, often with the same expert witnesses doing battle. The results have been anything but consistent.

In *Oxendine v. Merrell Dow Pharmaceuticals, Inc.*,[519] the District of Columbia Court of Appeals reversed a trial court that had granted a judgment notwithstanding the verdict to the defendant manufacturer. The trial court had found the testimony of the plaintiff's expert unfounded and insufficient.[520] The appellate court relied on *Ferebee* and simply repeated the explanation of the plaintiff's sole causation witness, with little consideration given to whether or not his reasoning conformed to accepted scientific practice.[521]

Undeterred by *Oxendine*, a federal district court judge in the District of Columbia granted another judgment notwithstanding the verdict in *Richardson v. Richardson-Merrell, Inc.*[522] after a trial dubbed a "virtual reprise of *Oxendine*."[523] The judge in *Richardson* pointed out that although the *Oxendine* appellate court had accurately summarized the testimony, it had not addressed "the significance of certain evidence bearing upon the current state of scientific knowledge. In consequence, it judicially reopened an esoteric twenty-year-old controversy which is by now essentially settled within the scientific community."[524] Based on "the totality of the published scientific literature on the subject of Bendectin,"[525] the *Richardson* court held that

> the issue being a scientific one, reasonable jurors could not reject [the scientific] consensus without indulging in precisely the same speculation and conjecture which the multiple investigations undertook, but failed, to confirm. That Dr. Done [plaintiff's principal expert] remains an unbeliever and was willing to testify to his disbelief 'with reasonable medical certainty' does not mandate that this case be left as the jury decided it.[526]

In addition to the conflict between *Oxendine* and *Richardson*, various courts have returned four verdicts for plaintiffs and twelve for the defendant,[527] including a verdict for the defendant when causation was tried to a jury as a separate issue for 1,174 consolidated cases in the

518. *See* Letter from Dr. Frank C. Woodside, III to the author (Sept. 28, 1987) (available in the files of the Fordham Law Review). Vermont is the sole exception. *See id.*

519. 506 A.2d 1100 (D.C. 1986).

520. *See id.* at 1103.

521. *See id.* at 1104-08.

522. 649 F. Supp. 799 (D.D.C. 1986).

523. *Id.* at 799.

524. *Id.*

525. *Id.* at 802.

526. *Id.* at 803; *see also* Lynch v. Merrell-Nat'l Labs., 646 F. Supp. 856, 862-67 (D. Mass. 1986), *aff'd*, 830 F.2d 1190 (1st Cir. 1987); *cf.* Will v. Richardson-Merrell, Inc., 647 F. Supp. 544, 547-48 (S.D. Ga. 1986) (testimony of surgeon excluded because he admitted he was not an expert on teratogenicity).

527. *See* 1 Inside Litigation 44 (Sept. 1987).

Northern and Southern Districts of Ohio,[528] and a $95 million plaintiff's verdict in still a third District of Columbia case.[529] Further clouding the picture is the fact that the *Oxendine* case has been reopened recently, due to a finding that the plaintiff's most important expert witness had perjured himself.[530] Thus, the law on the issue of causation of birth defects by Bendectin is uncertain at best. Different courts treat the same evidence differently, and some juries credit expert testimony that others apparently do not.

This state of uncertainty has driven the drug from the market and must give pause to manufacturers contemplating the development of new drugs and other products.[531] Some uncertainty is of course unavoidable, but a manufacturer should not be subject to liability when valid scientific reasoning does not support the inference that its product has caused an injury. Conflicting legal mazes such as the Bendectin litigation result from the unwillingness of some courts to keep scientific testimony within the bounds of scientific practice.

Although scientific knowledge is neither certain nor conclusive, neither is it impossible to tap. The fundamental scientific commitment to objective reasoning and objective evidence and the process of criticism, testing, and review, through which that commitment is maintained, allow a court to separate valid scientific reasoning from the personal biases of individual scientists. This is what the judge in *Richardson* did,[532] and if others followed suit, the uncertainty about Bendectin litigation would not exist.

IV. APPLICATION OF THE UNIFIED THEORY

Distinguishing reasoning from conclusions, and validity from reliability leads to a unified theory that encompasses all forms of scientific evidence, but using this theory to achieve rationality and consistency raises

528. *See Lynch*, 646 F. Supp. at 858. Some uncertainty exists about the exact number of cases consolidated. *See In re* Richardson-Merrell, Inc. "Bendectin" Prods. Liab. Litig., 624 F. Supp. 1212, 1216 n.1 (S.D. Ohio 1985) (indicating total of 1,186 cases with which the court had some contact).

529. *See* Ealy v. Richardson-Merrell, No. 83-3504, slip op. at 1-2 (D.D.C. Oct. 1, 1987). The award subsequently was reduced by $75 million because plaintiff's evidence was insufficient to serve as the basis for punitive damages. *See* Order dated Oct. 1, 1987.

530. *See* Findings of Fact Relevant to Defendant's Motion for Relief Pursuant to Rule 60(b), and Memorandum Order, Oxendine v. Merrell Dow Pharmaceuticals, Inc., No. 1245-82, slip op. at 4 (D.C. Super. Ct. Feb. 12, 1988) ("The court finds that [Dr. Done's] testimony was so deliberately false that *all* his testimony on behalf of plaintiff is suspect.") (emphasis in original).

531. *See* Fladwell, *U.S. Firms Abandoning Birth Control Industry in Wake of Lawsuits*, Washington Post, May 1, 1988, at H1, col. 1 (noting that private companies are withdrawing not only from marketing birth control technology but also from developing it; also quoting Harold Nash, a senior scientist at the Population Council, a nonprofit group: "It is possible that pharmaceutical companies with their large stores of chemicals might have found entirely new leads to contraception if they had continued research"); *see generally* Huber, *supra* note 3.

532. *See supra* notes 522-26 and accompanying text.

a number of practical concerns. Lawyers would have to force scientific issues into the open, and courts would then have to base their evidentiary rulings upon accepted scientific practice, tasks for which members of both bar and bench traditionally have had little taste.[533] Nonetheless, when courts have to confront the question of scientific validity, they usually deal with it successfully.[534]

The problem lies more in a failure to confront the issue of valid versus invalid reasoning than in explicitly deciding it incorrectly. The solution is to require experts to explain the reasoning leading to their conclusions and then to have courts judge the reasoning against scientific practice. The reliability of the conclusions themselves should be judged against legal standards involving probativeness, potential prejudice, and, ultimately, the burden of proof.

Adoption of this Article's proposed clarification of Federal Rule of Evidence 702[535] would focus legal attention on these fundamental questions. In some cases, the result would be to facilitate the evidentiary use of new scientific developments. In others, unscientific reasoning would be excluded. Procedurally, the kind of changes envisioned would best be effected through the use of pretrial motions.

A. *The Ability of Courts to Deal with Science*

Most of the controversy about scientific evidence has arisen in the context of cases involving medical testimony or *Frye*-type forensic science. When courts have to deal with science that is neither forensic nor medical, however, their response to scientific issues has often been consistent, well-informed, and appropriate. In a variety of contexts, ranging from patent disputes to so-called creation science, the ability of courts to understand science is quite apparent.

1. Patent Cases

Patent litigation often requires courts to evaluate science at a very detailed level, and they generally prove more than equal to the task. For example, a patent case may turn on the question of whether an invention is an obvious application of "prior art."[536] A party mounting a challenge to a patent will endeavor to establish that its subject matter "would have been obvious at the time the invention was made to a person having ordinary skill in the art to which [the] subject matter pertains."[537] Under

533. *See supra* note 31.

534. *See infra* notes 536-59 and accompanying text.

535. *See supra* text accompanying note 79.

536. *See, e.g.*, Milgo Elec. Corp. v. United Business Comms., Inc., 623 F.2d 645, 654 (10th Cir.) (in affirming validity of plaintiff's patent for device facilitating computer telecommunication, court held that obviousness turns on whether the device would have been conceivable to a worker of ordinary skill in the particular field, noting that in each case the prior art must be considered), *cert. denied*, 449 U.S. 1066 (1980).

537. 35 U.S.C. § 103 (1982 & Supp. IV 1986).

this rule, courts frequently must determine how a scientist would reason to broaden existing knowledge into new areas. The inquiry underlying such decisions involves consideration of a variety of evidence, including scientific journals and studies, as well as expert testimony and prior patents. The detailed judicial understanding of science that results is quite remarkable.

In *Ashland Oil, Inc. v. Delta Resins & Refractories, Inc.*,[538] for example, the court had to consider the molecular structure of resins used to form molds for metal casting. Ashland held a patent for making the resins and claimed that Delta had infringed upon it.[539] The district court held the patent invalid,[540] and Ashland appealed. The Court of Appeals for the Federal Circuit reversed and remanded because the lower court had misconstrued the burden of proof regarding Ashland's patent claim.[541] To guide the lower court on remand, the circuit court also reviewed the evidence and explained in detail how differences in bridging mechanisms between phenol rings in various molecules affected the degree to which certain prepatent publications had pointed the way to the discovery of the resin at issue.[542] The court clearly had mastered several complex principles of organic chemistry, and, even more significantly, considered the literature more probative than the expert testimony offered by Delta.[543] This refusal to rely blindly upon experts accords with the usual practice in patent cases, and further emphasizes the depth of substantive scientific review often undertaken by courts.

2. Creation Science

While patent cases illustrate judicial ability to grapple with science at a detailed level, cases involving the first amendment and creation science demonstrate that courts are equally capable of addressing issues related to the philosophical nature of science and scientific knowledge. Creation science, or "scientific creationism," arose in the 1960's in response to an increased emphasis on the theory of evolution in the biology curriculum of many schools.[544] The idea that species have evolved and developed over time contradicts a literal interpretation of the Book of Genesis because evolution implies that complex organisms such as man were not created instantly by God.[545] For religious fundamentalists who believe

538. 776 F.2d 281 (Fed. Cir. 1985), *cert. denied*, 475 U.S. 1017 (1986).

539. *See id.* at 284.

540. *See id.* at 283; Ashland Oil, Inc. v. Delta Resins & Refractories, Inc., 587 F. Supp. 1406 (E.D. Mich. 1984), *rev'd*, 776 F.2d 281 (Fed. Cir. 1985), *cert. denied*, 475 U..S 1017 (1986).

541. *See* 776 F.2d at 293.

542. *See id.* at 296-97.

543. *See id.* at 294.

544. *See* McLean v. Arkansas Bd. of Educ., 529 F. Supp. 1255, 1259 (E.D. Ark. 1982).

545. In fact, evolution does not explain, or even address, the origin of life. It deals only with how life has developed since its origin. *See id.* at 1266 ("the scientific community does not consider origins of life a part of evolutionary theory").

in the literal truth of the Bible, this contradiction means that evolution must be wrong.

To prove the error of evolution, some fundamentalists have sought to put their creationist view on a scientific footing,[546] and, by about 1977, one fundamentalist group began to advocate that states require the teaching of creation science as well as evolution.[547] Arkansas passed a Balanced Treatment Act in 1981,[548] as did Louisiana.[549] Both laws quickly drew challenges and ultimately both were struck down as violative of the first amendment's establishment clause.[550] The decision in *McLean v. Arkansas Board of Education*[551] provides a particularly thoughtful and thorough analysis of how science works and how courts can distinguish it from efforts to clothe religious beliefs in scientific garb.[552]

In order to determine if the challenged statute had any legitimate educational value that might counterbalance its advancement of fundamentalist precepts, the *McLean* court undertook an analysis of the scientific validity of creationism. The court noted that science is what scientists do and what is accepted by the scientific community,[553] and listed five essential characteristics of science: "(1) It is guided by natural law; (2) It has to be explanatory by reference to natural law; (3) It is testable against the empirical world; (4) Its conclusions are tentative, i.e., are not necessarily the final word; and (5) It is falsifiable."[554]

The court found that creationism lacked these characteristics and that not a single recognized scientific journal had ever published an article espousing creation science.[555] In response to the argument that a conspiracy against new ideas might explain the paucity of publication, the court noted that the scientific community consists of individuals and groups who work independently in a variety of fields.[556] The court observed that a scientist's work "is published and subject to review and testing by [his] peers,"[557] and it is "inconceivable that such a loose knit

546. *See id.* at 1259.
547. *See id.* at 1261.
548. Ark. Stat. Ann. §§ 80-1663 to -1670. (Supp. 1985).
549. La. Rev. Stat. Ann. § 17:286.1-.7 (West 1982 & Supp. 1988).
550. U.S. Const. amend. I.
551. 529 F. Supp. 1255 (E.D. Ark. 1982).
552. Litigation over the Arkansas creationism statute ended with the district court decision, *see id.*, but a challenge to the almost identical Louisiana statute eventually reached the United States Supreme Court. In Aguillard v. Treen, 634 F. Supp. 426, 427 (E.D. La. 1985), the court granted summary judgment to plaintiffs who had challenged the law's constitutionality. The Court of Appeals for the Fifth Circuit affirmed. *See* Aguillard v. Edwards, 765 F.2d 1251 (5th Cir. 1985). Applying a more limited analysis than that used by the Eastern District of Arkansas in *McLean*, the United States Supreme Court affirmed the Fifth Circuit but it never explicitly plumbed the meaning of "science." *See* Edwards v. Aguillard, 107 S. Ct. 2573 (1987).
553. *See* McLean v. Arkansas Bd. of Educ., 529 F. Supp. 1255, 1267 (E.D. Ark. 1982).
554. *Id.*
555. *See id.* at 1268.
556. *See id.*
557. *Id.*

group of independent thinkers in all the varied fields of science could, or would, . . . effectively censor new scientific thought."[558] Guided by a sound and thorough understanding of what comprises the actual practice of science, the *McLean* court easily focused on the scientific deficiencies of creationism and rejected it.[559]

B. *The Need to Make Reasoning Explicit and to Judge It Against Scientific Practice*

The above-described patent and creation science cases indicate that courts have the ability to deal with scientific issues, which suggests that in the areas of forensic science and medical testimony, the real problem may be that these issues never get considered. Experts often use ill-defined terminology,[560] and give vague and ambiguous testimony that lawyers simply fail to challenge[561] and that courts often accept even when it is challenged. Eliminating these problems requires a clear evidentiary standard that makes explicit the need to examine reasoning as well as conclusions. Such a standard would focus attention on the validity of reasoning and would provide the basis for appellate review when trial courts refuse to address it.

If courts in cases involving forensic science and medical testimony followed the lead of the *McLean* court and based their decisions about scientific evidence on accepted scientific practice,[562] most of the evidentiary problems surrounding science in these areas would disappear. Courts that have effectively controlled medical and forensic scientific evidence generally have done so because they identify and examine the reasoning behind conclusions.[563] They compare that reasoning to the accepted

---

558. *Id.*

559. An excellent account of the *McLean* case is provided in Thomas, *Commentary: Science v. Creation Science*, 11 Science, Technology & Human Values 47 (Summer 1986). Interestingly, some philosophers of science think that creation science could pass *McLean's* five-part test, *see McLean*, 529 F. Supp. at 1267 (though not necessarily the "science is what scientists do" test). One philosopher has written that members of his profession should only testify about *generally accepted* philosophy. *See* Thomas, *supra*, at 48-49 (citing Quinn, *The Philosopher of Science as Expert Witness*, in Science and Reality 32 (J. Cushing, C. Delaney & G. Gutting eds. 1984)).

560. *See* Smith, *What's in a Word: Expert Jargon and the Reasoning Behind It*, Defense Research Institute Monograph (in press).

561. Professor Giannelli has warned against inferring that testimony is reliable simply because it is unopposed. *See* Giannelli, *supra* note 4, at 1243; *see also* United States v. Luschen, 614 F.2d 1164, 1170 (8th Cir.) (defense lawyer neither objected to nor cross-examined witness about alleged deficiencies in scientific evidence later raised on appeal), *cert. denied*, 446 U.S. 939 (1980); *supra* note 31; *cf.* Fensterer v. State, 493 A.2d 959, 964 (Del.) (noting that effective cross-examination requires that expert commit to the basis for his opinion), *rev'd*, 474 U.S. 15 (1985).

562. *See supra* notes 553-59 and accompanying text.

563. *See, e.g.*, *In re* "Agent Orange" Prod. Liab. Litig., 611 F. Supp. 1223, 1242-43 (E.D.N.Y. 1985) (discussed *supra* notes 482-96 and accompanying text); Johnston v. United States, 597 F. Supp. 374, 431 (D. Kan. 1984), *cert. denied*, 108 S. Ct. 694 (1988) (discussed *supra* notes 501-05 and accompanying text); People v. Young, 425 Mich. 470, 477-78, 490-91, 391 N.W.2d 270, 273, 279 (1986) (involving electrophoresis blood typ-

practices of the scientific community, often turning to its literature for guidance,[564] and determine if the reasoning is appropriate to the facts of a given case.[565] Their analysis focuses not simply on the legal sufficiency of what an expert concludes, but on the way the conclusion was reached.

The New York trial court that prohibited the use of voiceprints for identification in *People v. Collins*[566] took precisely this approach. The court's analysis centered on the crucial element in the reasoning behind the voiceprint technique—the premise that voice spectrographs for an individual are relatively invariable, and that there is significant variation between different individuals.[567] Although the tradition that encrusts the judicial evaluation of medical testimony tends to prevent the kind of review of reasoning seen in *Collins*, some courts have taken a similar ap-

---

ing); People v. Collins, 94 Misc. 2d 704, 709-10, 405 N.Y.S.2d 365, 368-69 (1978) (discussed *supra* notes 270-82 and accompanying text). The *Young* case is discussed in detail in Black, *supra* note 193, at 1509.

564. *See, e.g.*, Richardson v. Richardson-Merrell, Inc., 649 F. Supp. 799, 803 (D.D.C. 1986) (discussed *supra* notes 522-26 and accompanying text); *In re* "Agent Orange" Prod. Liab. Litig., 611 F. Supp. 1223, 1240 (E.D.N.Y. 1985) (discussed *supra* notes 482-496 and accompanying text); Johnston v. United States, 597 F. Supp. 374, 428-30 (D. Kan. 1984) *cert. denied*, 108 S. Ct. 694 (1988) (discussed *supra* notes 501-05 and accompanying text); People v. Young, 425 Mich. 470, 489-90, 391 N.W.2d 270, 279 (Mich. 1986) (discussed *supra* note 563); People v. Collins, 94 Misc. 2d 704, 709-11, 405 N.Y.S.2d 365, 368-70 (1978) (discussed *supra* notes 270-82 and accompanying text).

565. *See, e.g.*, United States v. Downing, 753 F.2d 1224, 1242 (3d Cir.), *on remand*, 609 F. Supp. 784, 791-92 (E.D. Pa.), *aff'd mem.*, 780 F.2d 1017 (3d Cir. 1985) (discussed *supra* notes 350-63 and accompanying text); People v. Young, 425 Mich. 470, 490, 391 N.W.2d 270, 279 (Mich. 1986) (discussed *supra* note 563); People v. Collins, 94 Misc. 2d 704, 712-13, 405 N.Y.S.2d 365, 370-71 (1978) (discussed *supra* notes 270-82 and accompanying text).

566. 94 Misc. 2d 704, 405 N.Y.S.2d 365 (1978) (discussed *supra* notes 270-82 and accompanying text).

567. *See id.* at 712-18, 405 N.Y.S.2d at 370-74. The technique's principal advocate had acknowledged that the parameters responsible for variation were not well determined, and that the premise had been inferred indirectly. *See id.* at 713-14, 405 N.Y.S.2d at 371. The court noted, moreover, that "there has been no experimentation to show that two different voices will always appear different spectrographically." *Id.* at 713, 405 N.Y.S.2d at 371. The inference had not been tested "in the crucible of controlled experimentation . . . and scrutiny of the results in various scientific journals." *Id.* at 709, 405 N.Y.S.2d at 369.

It was not necessary in *Collins* to consider the appropriateness of voiceprint identification to the specific facts of the case because the invalidity of the reasoning made any forensic application improper. Sometimes, however, the court must follow the reasoning further. United States v. Downing, 753 F.2d 1224 (3d Cir.), *on remand*, 609 F. Supp. 784 (E.D. Pa.), *aff'd mem.*, 780 F.2d 1017 (3d Cir. 1985) (discussed *supra* notes 350-63 and accompanying text), which involved the psychology of eyewitness identification, demonstrates how the relationship between reasoning and specific facts can become the main concern. On remand, the trial court in *Downing* noted a number of differences between the situations covered by psychological studies and the circumstances at issue in the case. *See* 609 F. Supp. at 792 ("It is perhaps in establishing the 'fit' between the scientific research presented by Dr. Buckhout and the disputed factual issues of this case that defendant's argument . . . is weakest."). The court held that testimony based on the studies was inadmissable because they bore too little relationship to the circumstances involved in the case at bar. *See id.* at 792.

proach to medical issues. Cases such as *Johnston v. United States*[568] and the *Agent Orange* opt-out cases[569] demonstrate that, even for medical evidence, courts do have the ability to isolate an expert's reasoning for analysis and to compare it with accepted practices.[570]

C. *The Unified Theory's Effect on Admissibility: Two Contrasting Examples*

Although application of the unified theory often would make it more difficult to introduce some kinds of scientific evidence, in other cases the theory would facilitate admissibility. Requiring an accepted explanation of the reasoning behind expert testimony leads not only to the exclusion of specious and unscientific opinions but also makes it easier for courts to admit new, well-founded conclusions. Two examples highlight these contrasting, though consistent, results.

The first example involves the recognition, about twenty years ago, that some heart attacks suffered by munitions workers occurred as a result of withdrawal from exposure to nitroglycerin.[571] The second involves the generally rejected idea that common environmental exposures can compromise a person's immune system, thereby causing a host of non-specific and otherwise unexplained symptoms.[572] Courts have allowed expert witnesses to testify about each of these conclusions,[573] even though neither would have survived a typical application of the *Frye* test. Under the unified theory approach, the nitroglycerin evidence would easily qualify, but the immune deficiency evidence would not.

---

568. 597 F. Supp. 374 (D. Kan. 1984) (discussed *supra* notes 501-05 and accompanying text).

569. 611 F. Supp. 1223 (E.D.N.Y. 1985) (discussed *supra* notes 482-96 and accompanying text).

570. In both of these cases, the experts essentially based their conclusions about causation on the fact that plaintiffs had been exposed, at very low levels, to substances that had been shown to be at least potentially harmful at higher levels. *See In re* "Agent Orange" Prod. Liab. Litig., 611 F. Supp. at 1237-39; *Johnston*, 597 F. Supp. at 389, 415. As the *Johnston* court put it, reasoning that plaintiffs contracted cancer from the mere fact that instruments used by them were coated with radioactive paint, without experimental testing, constituted nothing more than speculation, *see* 597 F. Supp. at 393, which the experts "would not dare report in a peer-reviewed format." *Id.* at 415.

571. *See* Lange, Reid, Tresch, Keelan, Bernhard & Coolidge, *Nonatheromatous Ischemic Heart Disease Following Withdrawal from Chronic Industrial Nitroglycerin Exposure*, 46 Circulation 666, 675-77 (1972) [hereinafter Lange]; Lund, Haggendal & Johnsson, *Withdrawal Symptoms in Workers Exposed to Nitroglycerin*, 25 Brit. J. Indus. Med. 136, 136 (1968) [hereinafter Lund].

572. *See* American Academy of Allergy and Immunology, *Position Statement on Clinical Ecology*, 78 J. of Allergy & Clinical Immunology 269, 270 (1986) [hereinafter Academy Statement]; California Medical Association Scientific Task Force on Clinical Ecology, *Clinical Ecology—A Critical Appraisal*, 144 W. J. of Med. 239, 239 (1986) [hereinafter Task Force Appraisal].

573. *See, e.g.*, Stoleson v. United States, 708 F.2d 1217, 1220 (7th Cir. 1983) (nitroglycerin); Sterling v. Velsicol Chem. Corp., 647 F. Supp. 303, 499-507 (W.D. Tenn. 1986) (chemically induced immune deficiency syndrome), *aff'd in part, rev'd in part*, No. 86-6087, slip op. (6th Cir. May 24, 1988).

1. Nitroglycerin and Heart Attacks: Clear Proof Based on Clear Reasoning

Doctors long have recognized that workers who handle nitroglycerin and nitroglycol in the course of making munitions and explosives may experience immediate reactions that include headaches, nausea, and heart palpitations,[574] but that, except in severe cases, a tolerance develops within approximately one week.[575] In the early 1950's, however, an additional concern developed when reports appeared that some workers were suffering chest pains, and even death, due to withdrawal from exposure.[576] These effects generally were attributed to nitroglycol.[577] Then, in 1967, three Swedish researchers pinpointed withdrawal from nitroglycerin as the more probable cause.[578] Their study, based on nine cases,[579] soon was confirmed by an American study first presented in 1971.[580] The American work also was based on only nine cases,[581] but through careful analysis of electrocardiograms, angiograms, and other data it revealed the mechanism responsible for the problems associated with withdrawal: nitroglycerin exposure causes arteries to dilate, making it easier for the heart to pump blood,[582] but when exposure ends, the arteries tend to contract, which has the opposite effect and can lead to spasms and heart attacks.[583]

The relationship established by the American study was thoroughly explained and documented and was verified by detailed testing. In addition, both the American and the Swedish studies appeared in well-established, peer-reviewed journals. Nonetheless, this relationship was neither widely known nor generally accepted in the early 1970's.[584] If *Frye* had been applied to the conclusion and not to the reasoning, testimony about the relationship would have been inadmissible. Under the approach proposed in this Article, however, such a clear and verified explanation, appearing in a peer-reviewed journal, would have indicated admissibility. The absence of dissenting views would also have favored admissibility.

574. *See* Lund, *supra* note 571, at 136.
575. *See id.*
576. *See id.*
577. *See id.*
578. *See id.*
579. *See id.* at 136-38.
580. *See* Lange, *supra* note 571.
581. *See id* at 667.
582. *See id.* Indeed, doctors often use nitroglycerin to treat patients who have heart conditions. *See* Jordan, Seth, Henry, Wilen & Franciosa, *Dose Requirements and Hemodynamic Effects of Transdermal Nitroglycerin Compared with Placebo in Patients with Congestive Heart Failure*, 71 Circulation 980, 980 (1985).
583. *See* Lange, *supra* note 571, at 677.
584. *See* Stoleson v. United States, 629 F.2d 1265, 1267 (7th Cir. 1980) (noting that the Lange study documented for the first time the relationship between angina and chronic exposure to nitroglycerin, and that this "marked the first medical identification of the causal relation; neither clinical cardiology texts nor medical journals had previously discussed or described the phenomenon of angina among workers regularly exposed to nitroglycerin.").

Of course, if the probability were too low that the relationship applied to a specific, individual plaintiff's case, the evidence would fail the reliability portion of the proposed test. In fact, when a Wisconsin munitions plant worker sued the government in 1972 because of heart problems, causation was not contested.[585]

2. Clinical Ecology: Defective Reasoning Leads to Defective Conclusions

In recent years, some plaintiffs have attempted to establish a causal link between chemical exposures and diseases by introducing testimony based on "clinical ecology."[586] Experts in this field have testified that exposure to common industrial chemicals damages the immune system, and they maintain that this damage can in turn cause cancer and other diseases.[587] This line of reasoning provides plaintiffs with a wonderful argument because it relates virtually any disease to chemicals.[588] The only problem is that the scientific community generally has condemned the theories of clinical ecologists as unfounded and unscientific.[589]

Clinical ecology provides a classic example of how a court can resolve evidentiary disputes about science by reference to scientific literature. It also illustrates the need to view the literature critically, with an eye toward why articles are published and in what journals. Science thrives on new ideas, and journals often accept articles to generate debate and to foster the criticism and review so necessary to maintaining scientific objectivity. Thus, proponents of the theory of clinical ecology, which goes back some fifty years,[590] occasionally have published in well-regarded journals.[591] They also have their own publications, though these are not cited often by other scientists.[592] A quick and uncritical review of the literature therefore might indicate that the reasoning of clinical ecologists passes the threshold test of publication, but further review reveals that this does not necessarily imply acceptance.

---

585. *See id.* at 1267-68.

586. Marshall, *Woburn Case May Spark Explosion of Lawsuits*, 234 Sci. 418, 418 (1986).

587. *See id.*; *see also* Marshall, *Immune System Theories on Trial*, 234 Sci. 1490, 1490 (1986).

588. *See* Marshall, *supra* note 586, at 418 (statement of Donald Evans, Deputy General Counsel of the Chemical Manufacturers Association).

589. *See* Marshall, *supra* note 587, at 1491 (pointing out that even a scientist on the staff of the Environmental Defense Fund, who is a strong advocate of chemical regulation, has referred to clinical ecologists as "unspeakable"); *see also* Brodsky, *'Allergic to Everything': A Medical Subculture*, 24 Psychosomatics 731, 731 (1983) (referring to clinical ecology as a movement "on the fringe of established medicine").

590. *See* Task Force Appraisal, *supra* note 572, at 243.

591. *See, e.g.*, Randolph, *Ecologic Orientation in Medicine: Comprehensive Environmental Control in Diagnosis and Therapy*, 23 Annals of Allergy 7 (1965).

592. For example, the Journal of Clinical Ecology cannot be found at either the Johns Hopkins Medical Library, or the University of Maryland Medical School Library, libraries at two of the nation's top medical schools.

Both the American Academy of Allergy and Immunology[593] and the California Medical Association[594] have published official positions in their journals denouncing clinical ecology. Their disapproval derives from the fact that clinical ecologists tend only to describe their hypotheses and to give anecdotal examples.[595] The California group has observed that the few studies that have "used scientifically sound methods have provided evidence that the effectiveness of certain treatment methods used by clinical ecologists is based principally on placebo response."[596] The group found no convincing evidence that unique, recognizable syndromes occur in clinical ecology patients, or that the diagnostic tests used by practitioners in the field are either effective or reliable.[597]

The American Academy of Allergy and Immunology similarly has concluded that:

> The theoretical basis for ecological illness . . . has not been established as factual, nor is there satisfactory evidence to support the actual existence of "immune system dysregulation" or maladaption. There is no clear evidence that many of the symptoms [cited by clinical ecologists] are related to allergy, sensitivity, toxicity, or any other type of reaction from foods, water, chemicals, pollutants, viruses, and bacteria . . . . Properly controlled studies defining objective parameters of illness, properly controlled evaluation of the treatment modalities, and appropriate patient assessment have not been done. Anecdotal articles do not constitute sufficient evidence of a cause-and-effect relationship between symptoms and environmental exposure. The major techniques used by the clinical ecologists are controversial and unproven.[598]

Given such clear expositions of the flawed reasoning behind clinical ecology, a court should have no trouble excluding evidence based upon its tenets. At least one court, however, has accepted the testimony of clinical ecologists. In *Sterling v. Velsicol Chemical Corp.*,[599] damages

593. *See* Academy Statement, *supra* note 572.

594. *See* Task Force Appraisal, *supra* note 572. California has a particular interest in clinical ecology because the current advocates of the theory come mostly from the west coast and Texas. *See* Marshall, *supra* note 587, at 1490-91. Prominent among such experts is Dr. Alan Levin, who is from San Francisco. *See id.* at 1490.

595. *See* Task Force Appraisal, *supra* note 572, at 243.

596. *Id.*

597. *See id.*

598. Academy Statement, *supra* note 572, at 270; *see also* American Academy of Allergy and Immunology, *Candidiasis Hypersensitivity Syndrome*, 78 J. of Allergy & Clinical Immunology 271, 272-73 (1986) (recommending that the concept of candidiasis [yeast germ] hypersensitivity syndrome not be accorded recognition); American Academy of Allergy and Immunology, *Unproven Procedures for Diagnosis and Treatment of Allergic and Immunologic Diseases*, 78 J. of Allergy & Clinical Immunology 275, 275-76 (1986) (editorial calling for appropriate evaluation and testing); Lowell, *Some Untested Diagnostic and Therapeutic Procedures in Clinical Allergy*, 56 J. of Allergy & Clinical Immunology 168, 168-69 (1975) (same); *cf.* Grieco, *Controversial Practices in Allergy*, 247 J.A.M.A. 3106, 3110 (1982) (discussing lack of objective evidence to support the use of certain diagnostic and therapeutic techniques).

599. 647 F. Supp. 303 (W.D. Tenn. 1986), *aff'd in part, rev'd in part*, No. 86-6087, slip

were awarded explicitly for harm to the plaintiffs' immune systems.[600] Another court, in dealing with one of the same experts who testified in *Sterling*, reached the opposite conclusion. In *Schickele v. Rhodes*,[601] the defendants were granted a motion *in limine* to exclude the expert's testimony insofar as it was based on clinical ecology.

3. The Contrast Between the Examples and Why Scientific Deficiencies Should Be Apparent to Courts

The foregoing examples demonstrate the ways in which valid scientific reasoning differs from invalid reasoning. The nitroglycerin/heart attack relationship was established through painstaking objective analysis that conformed to the scientific criteria of testability, objectivity and impartiality, and to the requirement that evidence and reasoning be connected.[602] Clinical ecology fails on all of these counts. Comparison of one of the basic clinical ecology publications[603] with the 1971 American nitroglycerin study[604] illustrates the difference between the two examples. Where the nitroglycerin article contains a wealth of data, the clinical ecology article contains virtually none. It merely sets forth the hypotheses of the field and describes treatment regimes based on those hypotheses. Clinical ecologists have, in effect, attempted to create their own niche, separate from the rest of science. Their maneuver may avoid the crucible of criticism, but it is precisely this avoidance that makes their field unscientific.[605]

A court having to decide the admissibility of evidence based upon clinical ecology would find an almost complete absence of supporting

---

op. (6th Cir. May 24, 1988). The Sixth Circuit held Velsicol liable for some damages, but rejected clinical ecology and reversed on the issue of immune system damage. *Cf.* Barth v. Firestone Tire & Rubber Co., 673 F. Supp. 1466, 1469 (N.D. Cal. 1987) (holding that allegations of injury to immune system can survive a motion to dismiss, although noting that current medical science "cannot supply to the legal system information concerning the nature of present injury or of causation"); Anderson v. W.R. Grace & Co., 628 F. Supp. 1219, 1227 (D. Mass. 1986) (expert's affidavit suffices to bar summary judgment despite the fact that the affidavit did not specifically identify the illnesses suffered by each plaintiff as a result of immune system changes, nor state that plaintiffs suffered more ailments than the average person would have over the same time span).

600. *See* 647 F. Supp. at 344-47. This decision has sparked the same kind of scientific outcry as that which followed *Wells*. *See* Miller, *Courtroom Science and Standards of Proof*, The Lancet, Nov. 28, 1987, at 1283 (noting that in the *Sterling* and *Wells* cases and others, "[s]ubjectivity and impression [have] superseded scientific judgement; packaging and presentation rate higher than the scientific validity of the testimony."); *supra* Part III.B.1.b.

601. No. C-451843, slip op. (Ariz. Superior Ct. Maricopa County July 30, 1986); *cf.* Higgins v. Aerojet-General Corp., 1986 Env't Rep. (BNA) 1183 (Cal. Super. Ct. 1986) (jury verdict for defendant in action alleging injuries, including immune system damage, resulting from defendant company's disposal of trichloroethylene and other solvents in unlined ditches on its property).

602. *See supra* notes 127-29 and accompanying text.

603. *See* Randolph, *supra* note 591.

604. *See* Lange, *supra* note 571.

605. *See supra* notes 142-54 and accompanying text.

data.[606] It also would have the benefit of the criticism that has appeared in the literature. To the extent that clinical ecology has been subjected to the process of testing and critical review that provides the basis for scientific objectivity, it has failed miserably. Courts should have no trouble distinguishing such shoddy work from the kind of thorough effort that led to the understanding of the effects of nitroglycerin withdrawal. Application of the unified theory, with its focus on reasoning, will not impose an impossible burden on courts. They can and should exercise control over the use of scientific evidence.

D. *Procedural Issues*

When the unified theory approach leads to the conclusion that scientific evidence does not conform to accepted scientific practice, a court should either exclude the evidence as inadmissible or find it insufficient to sustain a verdict.[607] Inadmissibility offers the more appropriate basis for control when a conclusion is based on invalid reasoning, but as the judgment notwithstanding the verdict in the *Richardson* Bendectin case illustrates, insufficiency can serve just as well once an invalid conclusion is admitted.[608] In practice, the distinction between admissibility and sufficiency may really be just a question of how central the disputed scientific evidence is to the litigation.

If, in addition to the challenged evidence, a party has undisputed evidence sufficient to meet its burden of proof, the challenge usually will be decided on the basis of admissiblity. On the other hand, if the evidence at issue is vital to the party's case, summary judgment based on insufficiency may result. In either case, however, the court's analysis of the evidence probably will be the same. For example, whether a pretrial dispute about scientific evidence is framed in terms of a motion for summary judgment based on an *insufficient* opinion expressed at a deposition, or on a motion *in limine* based on the *inadmissibility* of the same opinion, a court's decision about the evidence most likely will turn on the same set of facts.

Although a court has available a variety of mechanisms for exercising control either before, during, or after trial,[609] pretrial resolution of dis-

606. The clinical ecology testimony relied upon by the trial court in *Sterling* illustrates the paucity of supporting data. When asked to explain immunology and allergy, one of the plaintiffs' principal witnesses simply gave a history of who had developed the field and how its two component parts once had been pursued separately. *See* Sterling v. Velsicol Chem. Co., 647 F. Supp. 303, 499-500 (W.D. Tenn. 1986), *aff'd in part, rev'd in part*, No. 86-6087, slip op. (6th Cir. May 24, 1988).

607. Even if courts do not exercise control directly, forcing reasoning into the open, and making it clear to the jury that scientific acceptance is to be the basis for finding it valid, would vastly improve the use of scientific evidence.

608. *See* Richardson v. Richardson-Merrell, Inc., 649 F. Supp. 799, 799-800 (D.D.C. 1986); *see also supra* notes 522-26 and accompanying text.

609. Motions *in limine* or motions for summary judgment would provide the principal opportunities before trial. Objections also will be raised during testimony at trial. After

putes about scientific evidence through summary judgment or orders *in limine* should be particularly encouraged.[610] Existing discovery procedures can generate adequate information before trial upon which to base decisions.[611] Indeed, the Manual for Complex Litigation contemplates the use of court-appointed experts, special masters, and magistrates to resolve evidentiary issues.[612] This kind of collateral factfinding requires significant time and effort, however, and would not be appropriate during trial. Prompted by this concern, the Third Circuit, in *In re Japanese Electronic Products Antitrust Litigation*,[613] held that *in limine* consideration of the evidentiary issues raised by expert testimony is more efficient than delaying a ruling until trial, reasoning that this procedure would permit more thorough briefing and argument.[614]

The more conclusive procedure of summary judgment also can be used to handle questionable scientific evidence, as illustrated by the disposition of the *Agent Orange* opt-out cases.[615] Since the *Agent Orange* decision, Supreme Court rulings in three landmark cases have made summary judgment an even more viable alternative.[616] Under the holding in *An-*

---

both sides have presented their cases, a motion for a directed verdict would be appropriate, and finally, a motion for judgment notwithstanding the verdict.

610. *See* Imwinkelried I, *supra* note 16, at 577-81 (advocating preliminary screening of validity by judge under Rules 901, 611, and 403 of the Federal Rules of Evidence). Professor Imwinkelried's proposal essentially is limited to forensic techniques and is predicated on a Logical Positivist view of the scientific method that would not provide a workable test for the complex patterns of expert reasoning seen in toxic tort cases. Nonetheless, the procedural part of his proposal fully accords with the analytical framework proposed in this Article. *See also* Epstein, *Motions in Limine—A Primer*, 8 Litigation 34 (Spring 1982) (discussing utility of motions *in limine*); Saltzburg, *Tactics of the Motion in Limine*, 9 Litigation 17 (Summer 1983) (same).

611. *See* Note, *supra* note 1, at 695; *see also* Manual for Complex Litigation, Second § 21.5 (1985) (covering special referrals to court-appointed experts, special masters, magistrates, and others); *Symposium on Science and the Rules of Legal Procedure*, 101 F.R.D. 599 (1983).

612. *See* Manual for Complex Litigation, Second § 21.5 (1985). In addition, Rule 43(e) of the Federal Rules of Civil Procedure contemplates full evidentiary hearings, if necessary, on pretrial motions. *See* Fed. R. Civ. P. 43(e).

613. 723 F.2d 238, 260 (3d Cir. 1983), *rev'd on other grounds sub nom.* Matsushita Elec. Indus. Co. v. Zenith Radio Corp., 475 U.S. 574 (1986).

614. *See* 723 F.2d at 260; *see also* Argus Inc. v. Eastman Kodak Co., 612 F. Supp. 904, 923, 926 (S.D.N.Y. 1985) (court must exclude, *in limine*, exhibits and damage theories based on speculation and guesswork; *in limine* consideration of expert testimony is also deemed appropriate), *aff'd*, 801 F.2d 38 (2d Cir. 1986), *cert. denied*, 107 S. Ct. 1295 (1987); Fowler v. Firestone Tire & Rubber Co., 92 F.R.D. 1, 3-4 (N.D. Miss. 1980) (granting defendant tire manufacturer's motion to exclude documents related to tire designs other than design of tire actually involved in accident that injured plaintiff).

615. *See In re* "Agent Orange" Prod. Liab. Litig., 611 F. Supp. 1223 (E.D.N.Y. 1985); *In re* "Agent Orange" Prod. Liab. Litig., 611 F. Supp. 1267 (E.D.N.Y. 1985); *see also supra* notes 483-96 and accompanying text.

616. *See* Matsushita Elec. Indus. Co. v. Zenith Radio Corp., 475 U.S. 574 (1986); Anderson v. Liberty Lobby, Inc., 477 U.S. 242 (1986); Celotex Corp. v. Catrett, 477 U.S. 317 (1986). For a good discussion of these cases, see generally Childress, *A New Era for Summary Judgments: Recent Shifts at the Supreme Court*, 116 F.R.D. 183 (1987); Wallance, *Summary Judgment Ascending*, 14 Litigation 6 (Winter 1988).

derson v. Liberty Lobby, Inc.,[617] for a plaintiff to survive a motion for summary judgment, "there must be evidence on which the jury could reasonably find for the plaintiff. The judge's inquiry, therefore, unavoidably asks whether reasonable jurors could find by a preponderance of the evidence that the plaintiff is entitled to a verdict . . . ."[618]

Anderson was a libel case, but its interpretation of summary judgment applies to scientific as well as other testimony. Relying on *Agent Orange*, *Anderson*, and other recent Supreme Court cases, the Western District of Michigan recently granted partial summary judgment in a case involving claims that the defendant's conduct had increased the risk that the plaintiffs would contract cancer.[619] The court held that the plaintiffs' expert testimony on the issue of causation was insufficient.[620]

Even if courts take full advantage of available procedures, the commitment of judicial resources required for the rational and consistent control of scientific evidence should not be minimized. Ultimately, however, improvements in fairness and efficiency would result. Over the long run, decisions based on a principled review of the validity of disputed reasoning would reduce the number of cases in which issues related to a specific kind of testimony would have to be addressed.

CONCLUSION

Rational and consistent judicial control of scientific evidence requires an analytical framework based on a distinction between the scientific validity of the reasoning that connects facts to conclusions, and the reliability of the conclusions themselves in view of legal requirements. Within such a framework, a uniform and consistent standard can be applied to all kinds of science in all kinds of cases. To put this proposal into practice, the Federal Rules of Evidence should be clarified by establishing an explicit requirement for scientifically valid reasoning, and a separate requirement for legally sufficient reliability. This change would resolve the longstanding debate about the *Frye* general acceptance test, and it also would greatly improve the judicial evaluation of medical testimony, to which *Frye* traditionally has not been applied.

In recent years the need for a change in the rules governing scientific evidence has become especially acute. The emergence of toxic tort litigation has brought disputes about the reasoning of experts to the fore. In toxic tort cases the conclusions reached by experts very often are quite clear (even if not clearly sufficient), as are the facts upon which the experts rely. The real issue is whether scientifically valid reasoning connects facts to conclusions. Historically, cases involving traditional medical testimony or forensic science have turned on other issues. In-

617. 477 U.S. 242 (1986).

618. *Id.* at 252.

619. *See* Stites v. Sundstrand Heat Transfer, Inc., 660 F. Supp. 1516 (W.D. Mich. 1987).

620. *See id.* at 1524-25.

stead of reasoning and its validity, courts usually have focused on factors such as the probability of error or the degree of certainty expressed by a witness. This approach has led to aberrations even in the more traditional cases, but in toxic tort cases it has sometimes resulted in decisions that defy reality.

When confronted with disputed and questionable expert scientific testimony, courts need not and should not passively admit opinions that do not derive from accepted scientific practices, and lawyers need not and should not leave unscientific testimony unchallenged. With a rudimentary understanding of how science depends on a commitment to objectivity and on a process of criticism and review, both judges and lawyers are more than capable of distinguishing valid from invalid reasoning. Indeed, cases that have arisen in contexts that do not involve medical or forensic science clearly demonstrate that the judicial system can effectively evaluate scientific evidence if the effort is made.

The question of scientifically valid reasoning ultimately comes down to a question of acceptance, and the law can apply this test just as well as science applies it. Courts should not reopen questions that science has resolved, nor should they presume the existence of knowledge that science cannot provide.[621] A legal standard that looks to scientific acceptance as the best evidence of valid reasoning would produce decisions consistent both with science and with each other, and would avoid most of the problems that scientific evidence has so long posed for the law.

621. *See* Marshall, *supra* note 587, at 1491. In this article, Marshall discusses the inappropriateness of resolving scientific issues in the courtroom. He quotes Ellen Silbergeld, staff scientist with the Environmental Defense Fund, who has expressed fears that experts who advocate unsound theories will frighten away good scientists: " 'I hate to see a novel area of science being worked out in the courts. If it discourages good people from getting involved, it would be unfortunate.' " *Id.*