

**THE ADMISSIBILITY OF SCIENTIFIC EVIDENCE:
EXPLORING THE SIGNIFICANCE OF THE
DISTINCTION BETWEEN FOUNDATIONAL
VALIDITY AND VALIDITY AS APPLIED**

Edward J. Imwinkelried†

TABLE OF CONTENTS

I. FOUNDATIONAL VALIDITY	821
A. <i>Foundational Validity in Science</i>	821
B. <i>Foundational Validity in Evidence Law</i>	823
II. VALIDITY AS APPLIED	825
A. <i>Validity as Applied in Evidence Law</i>	825
B. <i>Validity as Applied in Science</i>	831
1. <i>The Technical Sense</i>	831
2. <i>The Deeper Scientific Sense</i>	832
C. <i>The Recognition of the General Notion of Range of Validation</i>	832
D. <i>Specific Examples of the Notion of Range of Validation</i>	834
E. <i>Incorporating the Range of Validation Concept into an Approach to Determining the Validity as if the Applied of Proffered Expert Testimony</i>	839
CONCLUSION	846

“[E]xperts commonly extrapolate from existing data. . . . [However, a] court may conclude that there is simply too great an analytical gap between the data and the opinion proffered.”¹

The use of expert testimony in American trials is widespread and accelerating. In a Rand Corporation study of California trials in courts of general jurisdiction, the researchers reported that experts appeared in eighty-six percent of the trials.² In the study, on average there were 3.3 experts per trial.³ A more recent study found that the average had risen to

† Edward L. Barrett, Jr. Professor of Law Emeritus, University of California Davis; former chair, Evidence Section, American Association of Law Schools; coauthor, Giannelli, Imwinkelried, Roth & Campbell Moriarty, *Scientific Evidence* (6th ed. forthcoming 2020). This article will be the basis for Professor Imwinkelried’s presentation at the Sino-Swiss Evidence Conference, University of Lausanne, Switzerland, September 2021.

1. Gen. Elect. Co. v. Joiner, 522 U.S. 136, 146 (1997).

2. Samuel R. Gross, *Expert Evidence*, 1991 Wis. L. Rev. 1113, 1119 (1991).

3. *Id.*

4.31 experts per trial.⁴ One commentator has asserted—with only slight hyperbole—that in the United States, trial by jury is evolving into trial by expert.⁵

Understandably, American courts and evidence commentators have devoted an enormous amount of attention to the legal standard governing the admissibility of scientific testimony. Until the mid 1970s, most American jurisdictions followed the standard announced in 1923 in *Frye v. United States*.⁶ Under that traditional standard, a scientific methodology, that is, a theory or technique, could serve as a basis for admissible testimony only if its proponent could show that the methodology had gained general acceptance within the relevant scientific circles.⁷ However, in 1975, the Federal Rules of Evidence took effect.⁸ The Rules made no mention of the general acceptance test.⁹ In a line of cases, including *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993),¹⁰ *General Electric Co. v. Joiner* (1997),¹¹ *Kumho Tire Co., Ltd. v. Carmichael* (1999),¹² and *Weisgram v. Marley Co.* (2000),¹³ the Supreme Court abandoned the *Frye* test and substituted an empirical validation/reliability standard derived from the text of Federal Rule 702.¹⁴ In pertinent part, the current version of Rule 702 reads:

A witness who is qualified as an expert by knowledge, skill, training, or education may testify in the form of an opinion or otherwise if:

the expert's scientific, technical, or other specialized knowledge will help the trier of fact to understand the evidence or to determine a fact in issue;

(b) the testimony is the product of reliable principles and methods; and

4. RONALD J. ALLEN, RICHARD. B. KUHN, ELEANOR SWIFT, DAVID S. SCHWARTZ, & MICHAEL S. PARDO, *EVIDENCE: TEXT, CASES AND PROBLEMS* 649 (5th ed. 2011).

5. William T. Pizzi, *Expert Testimony in the US*, 145 N.L.J. 82 (1995).

6. 293 F. 1013 (D.C. Ct. Apps. 1923).

7. *Id.*

8. See Public Law No. 93-595.

9. See Joseph R. Meaney, *From Frye to Daubert: Is a Pattern Unfolding?*, 35 *JURIMETRICS*, 191, 191 (1995) (“[T]he text of Federal (or Uniform) Rule of Evidence 702 on expert testimony does not explicitly mention *Frye*.”).

10. 509 U.S. 579 (1993).

11. 522 U.S. 136 (1997).

12. 526 U.S. 137 (1999).

13. 528 U.S. 440 (2000).

14. See generally 2 *FEDERAL EVIDENCE TACTICS* § 7.02 (2019) (explaining that the Court now views Federal Rule 702(a) as creating an empirical testing and validation method that applies to both controlled empirical testing and other primary methods of validating expert testimony after *Daubert* and *Kumho*). See *Daubert*, 509 U.S. at 582; see also *Gen. Elec. Co.*, 522 U.S. at 142.

2020] **The Admissibility of Scientific Evidence** 819

(c) the expert has reliably applied the principles and methods to the facts of the case.¹⁵

In *Daubert*, the Court explained that the statutory reference to “scientific knowledge” requires that the proponent present enough empirical data and reasoning to persuade the trial judge that the expert’s methodology is reliable in the sense that it is “supported by appropriate [scientific] validation.”¹⁶ A 2000 amendment to Rule 702 imposed the additional requirement that the proponent demonstrate that the expert “reliably applied” the methodology to the specific facts of the instant case.¹⁷

In 2016, the President’s Council of Advisors on Science and Technology (PCAST) released a highly publicized report entitled *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods*.¹⁸ Chapter three of the report is devoted to “The Role of Scientific Validity in the Courts.”¹⁹ The chapter distinguishes between “foundational validity” and “validity as applied.”²⁰ Foundational validity corresponds to Rule 702(a)’s requirement that the proponent establish the general reliability of the expert’s methodology, while validity as applied is equivalent to Rule 702(d)’s mandate that the proponent demonstrate that the expert has properly applied the methodology in the pending case.²¹ Like the 2000 amendment to Rule 702, the PCAST report highlighted the distinction between the question of the general reliability of an expert methodology and the propriety of its application in the pending case.²² That distinction is an important one; in many of the studies of forensic laboratory performance, researchers have found that although the expert employed a trustworthy methodology, the expert erred because he or she misapplied the methodology.²³

15. FED. R. EVID. 702.

16. *Daubert*, 509 U.S. at 590.

17. FED. R. EVID. 702, Adv. Comm. Note 2000 Amend.

18. PRESIDENT’S COUNCIL OF ADVISORS ON SCIENCE AND TECHNOLOGY, EXEC. OFFICE OF THE PRESIDENT, *FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING SCIENTIFIC VALIDITY OF FEATURE COMPARISON METHODS* (2016) [hereinafter PCAST].

19. *See id.* at 40.

20. *See id.* at 43.

21. Eric S. Lander, *Fixing Rule 702: The PCAST Report and Steps to Ensure the Reliability of Forensic Feature-Comparison Methods in the Criminal Court*, 86 *FORDHAM L. REV.* 1661, 1664–65 (2018). Professor Lander was the Co-Chair of PCAST at the time it prepared the 2016 report. *See* PCAST, *supra* note 18, at 43.

22. *See* Lander, *supra* note 21; *see also* PCAST, *supra* note 18, at 43.

23. Edward J. Imwinkelried, *The Debate in the DNA Cases Over the Foundation for the Admission of Scientific Evidence: The Importance of Human Error as a Cause of Forensic Misanalysis*, 69 *WASH. U. L. REV.* 19, 26, 32 (1991) [hereinafter Imwinkelried, *Debate*].

The thesis of this article is that in order to correctly enforce the distinct requirement for proof of validity as applied, courts need to more carefully examine the parameters of the validation studies used to establish the general foundational validity of an expert methodology. In particular, the courts must determine the methodology's extent or range of validation demonstrated in those studies and should find validity as applied lacking when the proponent's expert attempts to employ the methodology in a fact situation exceeding that range. To develop that thesis, the article proceeds in two parts. The first part of this article discusses foundational validity. Initially, this part describes the concept from a scientific perspective. The part then demonstrates that the courts have incorporated the concept of foundational validity into their admissibility analysis.

Part two turns to the principal focus of this article, namely, validity as applied. Just as part one examines foundational validity from both the scientific and legal viewpoints, part two adopts the same approach to validity as applied. To begin with, part two demonstrates that the cases and court rules differentiate between the concepts of foundational validity and validity as applied. Part two then demonstrates that the validity as applied concept is deeply embedded in scientific reasoning, especially in metrology, the science of measurement. Next, part two notes the striking analogy between a judicial determination of whether to extend a common-law precedent to a new fact situation and a judicial decision whether to permit an expert to apply a methodology to a fact situation beyond the precise parameters of the validation studies. Part two elaborates on the practical challenge facing a judge required to make the latter decision. Part two argues that if the judge lacks the information necessary to evaluate the propriety of an expert extrapolation, the outcome should be the exclusion of the testimony about the extrapolation. The judge should assign the proponent of the extrapolation the burden of proof on the defensibility of the extrapolation.

The conclusion argues that in order to properly enforce the validity as applied requirement in the future, the courts must scrutinize validation studies far more closely than most courts have done in the past. The courts must move beyond a fixation with the quantitative aspects of validation studies and expand their focus to include the qualitative aspects of the studies, that is, the conditions under which the methodology was validated. The courts can give the proponent of an extrapolation a powerful incentive to provide the trial judge with the information needed to make informed rulings on the validity as applied issue by making it crystal clear that the proponent has the burden of establishing an empirical justification warranting any application of the methodology that seemingly

exceeds the demonstrated range of validation. The Supreme Court's forceful language in *Joiner* and the explicit prescription in Federal Rule 702(d) demand nothing less.²⁴

I. FOUNDATIONAL VALIDITY

A. Foundational Validity in Science

Chapter 3 of the PCAST Report contains a lucid explanation of the concept of foundational validity.²⁵ Suppose that a researcher wants to investigate the hypothesis that a particular expert technique or theory is valid. In order to falsify or validate the hypothesis, the researcher can employ the “method . . . that has characterized the natural sciences since the 17th century, consisting in systematic observation, measurement, and experimentation, and the formulation, testing, and modification of hypotheses.”²⁶ In other words, the researcher conducts a particular type of experiment, namely, an empirical validation study.²⁷ In *Daubert*, Justice Blackmun commented that “a key question” in assessing the sufficiency of the proponent's showing of reliability is whether the proponent's hypothesis “has been . . . tested.”²⁸ When the hypothesis is the validity of an expert methodology, the test takes the form of a validation study.²⁹ The PCAST Report emphatically states that extensive experience with a technique by forensic practitioners is no substitute for such validation;³⁰ “[f]oundational validity is a *sine qua non*, which can only be shown through empirical studies.”³¹

In designing the validation study, the researcher controls certain variables and investigates to determine whether, by controlling those variables, he or she can make an accurate prediction of the outcome of the experiment.³² In the final analysis, the hypothesis is a conditional

24. See 522 U.S. 136 at 146–47; see also FED. R. EVID. 702(d).

25. PCAST, *supra* note 18, at 43. PCAST defines *scientific* standards under the legal standards in Rule 702(c) and 702(d). *Id.* PCAST defines “foundational validity” as “the *scientific* standard corresponding to the legal standard of evidence being based on ‘reliable principles and methods.’” *Id.*

26. *Id.* at 46 n.101 (quoting *Scientific method*, OXFORD DICTIONARY (2d ed. 2016)).

27. *Id.* at 46, 52.

28. *Daubert*, 509 U.S. at 593.

29. See *id.* at 590 (finding that in order to qualify as “scientific knowledge” an inference must be derived by the scientific method. The proposed testimony must be supported by appropriate validation of the expert's methodology to meet the requirement and establish a standard of evidentiary reliability).

30. PCAST, *supra* note 18, at 6, 55.

31. *Id.* at 66.

32. See *id.* at 65–66.

proposition: *If* conditions A, B, and C are controlled, *then* what is likely to be the nature of outcome D?³³ In evaluating the accuracy of the methodology,³⁴ the researcher attempts to assess both the specificity and the sensitivity of the methodology.³⁵

—How specific is the technique? In what percentage of cases in which the methodology predicts a positive outcome or conclusion (for example, that the sample is a specific contraband drug), does the methodology lead to an erroneous conclusion (a false positive or Type I error)?

—And how sensitive is the technique? In what percentage of cases in which the methodology predicts a negative outcome or conclusion (for instance, that the sample is not a specific drug), does the methodology yield the converse type of error (a false negative or Type II error)?

Of course, to make reliable assessments, the researcher must know the ground truth.³⁶ Thus, if the hypothesis relates to the accuracy of a measuring device, the researcher can use certified reference material (RM) supplied by a national or international metrological authority such as the National Institute of Standards and Technology (NIST).³⁷

Again, the hypothesis is a conditional proposition.³⁸ The researcher must try to validate the methodology for its intended use.³⁹ Thus, as a generalization if the researcher is endeavoring to validate the use of the methodology for forensic casework, he or she ought to control the variables by specifying conditions that are representative of real world cases.⁴⁰ In the words of a 1979 National Academy of Sciences report on the sound spectrography technique for identifying voices, the study must explore the validity of the methodology “over the range of conditions usually met in practice.”⁴¹ The PCAST Report uses fingerprint examination as a further illustration of the point.⁴² In real life casework, fingerprint examiners routinely encounter low quality latent prints, that is, prints that are both

33. See Edward J. Imwinkelried, *Coming to Grips with Scientific Research in Daubert's “Brave New World”: The Courts’ Need to Appreciate the Evidentiary Difference Between Validity and Proficiency Studies*, 61 BROOK. L. REV. 1247, 1258 (1995).

34. See *id.* at 47–48.

35. *Id.* at 50.

36. See William A. Woodruff, *Evidence of Lies and Rules of Evidence: The Admissibility of fMRI-Based Expert Opinion of Witness Truthfulness*, 16 N.C. J.L. & TECH. 105, 223 (2014).

37. TED VOSK, FORENSIC METROLOGY: SCIENTIFIC MEASUREMENT AND INFERENCE FOR LAWYERS, JUDGES, AND CRIMINALISTS 81–82 (1st ed. 2014).

38. See PCAST, *supra* note 18, at 60; see also *supra* text accompanying note 33.

39. *Id.* at 46.

40. *Id.* at 48, 52, 66.

41. NATIONAL ACADEMY OF SCIENCES, ON THE THEORY AND PRACTICE OF VOICE IDENTIFICATION 58 (1979).

42. See PCAST, *supra* note 18, at 52, 149.

partial and distorted.⁴³ Even if an experiment involving a fingerprint technique produced an impressive accuracy rating, the experiment would furnish little validation for the forensic application of the technique if the finger marks in the study were complete and high-quality, fully scanned prints.⁴⁴

B. Foundational Validity in Evidence Law

Prior to *Daubert*, the courts were often content to rely on proxies rather than directly addressing the question of the empirical validity of an expert technique.⁴⁵ By way of example, instead of reviewing the empirical data in the relevant validation studies, under *Frye* a court confined its inquiry to how well accepted or popular a methodology was within the relevant scientific circles.⁴⁶ The appellate courts precluded trial judges from scrutinizing the empirical data underlying the methodology in part on the assumption that trial judges lacked the competence to critically evaluate the validation studies.⁴⁷

In this respect, *Daubert* works a sea change in Evidence law.⁴⁸ Post *Daubert*, trial judges may no longer “[hide] from science.”⁴⁹ *Daubert* tasks trial judges to determine whether the methodology underlying proffered expert testimony is reliable in the classic scientific sense:

“Science is not an encyclopedic body of knowledge about the universe. Instead, it represents a process for proposing and refining theoretical explanations about the world that are subject to further testing and refinement.” [I]n order to qualify as “scientific knowledge” [within the intendment of that expression in Rule 702], an inference or assertion must be derived by the scientific method. Proposed testimony must be supported by appropriate validation—i.e., “good grounds,” based on what is known. In short, the requirement that an expert’s testimony

43. *Id.* at 52.

44. *Id.*

45. Bert Black, Francisco J. Ayala & Carol Saffran-Brinks, *Science and the Law in the Wake of Daubert: A New Search for Scientific Knowledge*, 72 TEX. L. REV. 715, 723–24 (1994).

46. *Id.* at 725.

47. 1 PAUL C. GIANNELLI, EDWARD J. IMWINKELRIED, ANDREA ROTH & JANE CAMPBELL MORIARTY, SCIENTIFIC EVIDENCE § 1.06(a) (5th ed. 2012).

48. See Black, Ayala & Saffran-Brinks, *supra* note 45, at 722 (“The analysis used by pre-*Daubert* courts that applied the Rules in lieu of *Frye* typically involves balancing various enumerated factors, albeit without any guidance on how the factors relate to each other or how they fit into a coherent picture of the way science actually works.”); *but see* FED. R. EVID. 702, Adv. Comm. Note 2000 Amend. (commenting that *Daubert* did not work a “seachange over federal evidence law,” and that caselaw shows that courts do not commonly reject expert testimony despite the courts’ gatekeeper role post-*Daubert*).

49. Black, Ayala & Saffran-Brinks, *supra* note 45, at 722.

pertain to “scientific knowledge” establishes a standard of evidentiary reliability.⁵⁰

Justice Blackmun continued:

Faced with a proffer of expert scientific testimony, the trial judge must . . . [make] a preliminary assessment of whether the reasoning or methodology underlying the testimony is scientifically valid. “Scientific methodology today is based on generating hypotheses and testing them to see if they can be falsified”⁵¹

In the macrocosm, society places faith in science because that is an “immense body” of empirical data demonstrating the successful “results” of applying scientific methodology.⁵² *Daubert* prescribed that in the microcosm of deciding whether to admit testimony about a specific scientific methodology, the judge should similarly focus on the results in the empirical data.⁵³

Perhaps the best synthesis of *Daubert* line of authority’s pronouncements on foundational validity is that the expert must marshal enough empirical data and reasoning to convince the trial judge that by employing the particular methodology that he or she proposes relying on, the expert can accurately draw the specific type of inference that the expert contemplates testifying to.⁵⁴ The judge’s analytic focus should be on the expert’s particular methodology, not the global validity of the expert’s discipline.⁵⁵ Furthermore, as the Advisory Committee Note to the 2000 amendment to Rule 702 emphasizes, out of respect for the jury’s role, the judge does not pass on the question of whether the specific opinion drawn

50. *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 590 (1993).

51. *Id.* at 592–93.

52. See JOHN ZIMAN, RELIABLE KNOWLEDGE: AN EXPLORATION OF THE GROUNDS FOR BELIEF IN SCIENCE 6–7, 127 (1978).

53. See *Daubert*, 509 U.S. at 593.

54. Edward J. Imwinkelried, *The Best Insurance Against Miscarriages of Justice Caused by Junk Science: An Admissibility Test that Is Scientifically and Legally Sound*, 81 ALB. L. REV. 851, 857 (2017/2018) [hereinafter Imwinkelried, *Insurance*].

55. D. Michael Risinger, *Defining the “Task at Hand”: Non-Science Forensic Science After Kumho Tire Co. v. Carmichael*, 57 WASH. & LEE L. REV. 767, 769–70, 772, 774, 798 (2000). In *Kumho*, the Court made it clear that the trial judge must conduct a “very particular analysis” of the expert’s ability to perform the specific task at hand. *Id.* at 774 (citing *Kumho*, 526 U.S. at 141). In *United States v. Fujii*, the court balked at issuing a sweeping rule on the global validity of questioned document examination; however, the court refused to admit a QD examiner’s testimony purportedly identifying the author of a document composed in Japanese handprinting; the testimony in the record indicated that when persons are taught that style of handprinting, they are encouraged to suppress individual characteristics and precisely reproduce the figure. See 152 F. Supp. 2d 939, 940, 941 (N.D. Ill. 2000).

by the expert is correct,⁵⁶ rather, the judge's limited role is to review the empirical data to determine whether they establish the expert's ability to draw the type or kind of inference he or she proposes testifying to, such as an inference as to a person's credibility or a disputed fact on the historical merits.⁵⁷

II. VALIDITY AS APPLIED

In the typical case, the litigants are not concerned only about the general or foundational validity of an expert methodology.⁵⁸ The litigants are also concerned about the procedures actually applied in the case and whether the application was proper.⁵⁹ As previously stated, the 2000 amendment to Rule 702 added an express requirement that the proponent show that the expert properly applied the methodology to the facts of the instant case.⁶⁰ The amendment reflects the elementary insight that as a matter of logic, validity as applied is just as essential to the reliability of the testimony proffered as foundational validity.⁶¹

A. Validity as Applied in Evidence Law

Although most of the early commentary on the 1993 *Daubert* decision focused on the new empirical standard for foundational validity announced in that case, the *Daubert* opinion itself included references to the concept of validity as applied.⁶² Justice Blackmun wrote that the trial judge must determine whether the "methodology properly can be applied

56. FED. R. EVID. 702, Adv. Comm. Note 2000 Amend. (the trial judge does not apply "the merits standard of [the] correctness" of the opinion).

57. Imwinkelried, *Insurance*, *supra* note 54, at 866–69. In the original *Daubert* opinion, Justice Blackmun had cautioned that "[t]he focus . . . must be solely on principles and methodology, not on the conclusions that they generate." 509 U.S. at 595.

58. PCAST, *supra* note 18, at 56 (describing foundational validity as a method that can be reliable in principle, compared to validity as applied as a method that has been reliably applied in practice).

59. *Id.* at 56, 66.

60. FED. R. EVID. 702, Adv. Comm. Note 2000 Amend. The amendment specifically provides that the trial court must scrutinize not only the principles and methods used by the expert, but also whether those principles and methods have been properly applied to the facts of the case. See *United States v. Gomez-Paz et al.*, 2011 U.S. Dist. LEXIS 105442, 2011 WL 4345891 (D. Colo. Sept. 16, 2011) (although the court found that the prosecution's foundation "squeaks by" under Rule 702, the court stressed that the proponent must show the "reliable application" of the expert's methodology).

61. See FED. R. EVID. 702, Adv. Comm. Note 2000 Amend. (stating that trial courts must use *Daubert* factors to assess the reliability and helpfulness of proffered expert testimony).

62. See *Daubert*, 509 U.S. at 593.

to the facts in issue.”⁶³ He stated that the methodology must “fit” the “facts of the case”⁶⁴ and be relevant to the specific “task at hand.”⁶⁵

As time passed, the need for a showing of validity as applied became even clearer.⁶⁶ A case in point is the Supreme Court’s 1997 decision in *General Electric Co. v. Joiner*.⁶⁷ There the plaintiffs endeavored to establish that their exposure to polychlorinated biphenyls (PCBs) at the defendant’s plant caused Mr. Joiner’s lung cancer.⁶⁸ The plaintiffs relied heavily on animal studies to prove causation.⁶⁹ However, there were marked differences between the parameters of the studies and the facts of the instant case.⁷⁰ Chief Justice Rehnquist wrote:

The [animal] studies involved infant mice that had developed cancer after being exposed to PCBs. The infant mice in the studies had had massive doses of PCBs injected directly into their peritoneums or stomachs. Joiner was an adult human being whose alleged exposure to PCBs was far less than the exposure in the animal studies. The PCBs were injected into the mice in a highly concentrated form. The fluid with which Joiner had come into contact generally had a much smaller PCB concentration of between 0-500 parts per million. The cancer that these mice developed was alveologenic adenomas; Joiner had developed small-cell carcinomas.⁷¹

The Court did not deny that the studies supported the hypothesis that massive direct injections of PCBs into certain organs of infant mice can cause certain types of cancer.⁷² However, the dispositive question in *Joiner* was whether those studies provided sufficient empirical support for an inference of causation in human beings under different conditions.⁷³ The Court stressed that the parameters of the studies were “so dissimilar to the facts presented in this litigation”⁷⁴ The Court then addressed the broader question of validity as applied:

Trained experts commonly extrapolate from existing data. But nothing in either *Daubert* or the Federal Rules of Evidence requires a district

63. *Id.*

64. *Id.* at 591; see *Harris v. Remington Arms Co.*, 398 F. Supp. 3d 1126, 1130 (W.D. Okla. 2019).

65. *Daubert*, 509 U.S. at 597.

66. See *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 146 (1997).

67. *Id.*

68. *Id.* at 139–40, 143.

69. *Id.* at 143.

70. See *id.* at 144.

71. *Joiner*, 522 U.S. at 144.

72. See *id.*

73. See *id.*

74. *Id.*

2020]

The Admissibility of Scientific Evidence

827

court to admit opinion evidence that is connected to existing data only by the *ipse dixit* of the expert. A court may conclude that there is simply too great an analytical gap between the data and the opinion proffered.⁷⁵

Three years later in 2000, Rule 702 was amended to explicitly impose a requirement for a showing of validity as applied.⁷⁶ As restyled in 2011, Rule 702(c) mandates a showing of foundational validity: the proponent must show that “the testimony is the product of reliable principles and methods.”⁷⁷ The very next subsection, 702(d), now provides that the proponent must also establish that “the expert has reliably applied the principles and methods to the facts of the case.”⁷⁸ The accompanying Advisory Committee Note echoes the pertinent passages in *Daubert* and *Joiner*; the Note mentions the need for “fit” and observes that the trial judge must inquire whether the expert has “unjustifiably extrapolated.”⁷⁹ The Note flatly asserts that the judge must find that the methodology has “been properly applied to the facts of the case” and that a “misappli[cation]” can render the expert’s testimony unreliable and inadmissible.⁸⁰

The jurisprudence in the lower courts is mixed.⁸¹ Some courts are content to reiterate the generalization that the proponent must establish that the expert has properly applied the methodology to the facts of the pending case.⁸² Other courts tend to focus on the quantitative aspects of

75. *Id.* at 146.

76. FED. R. EVID. 702, Adv. Comm. Note 2000 Amend.

77. *Id.* 702(c).

78. *Id.* See *United States v. Gomez-Paz et al.*, 2011 U.S. Dist. LEXIS 105442, 2011 WL 4345891 (D. Colo. Sept. 16, 2011) (“reliable application”).

79. FED. R. EVID. 702, Adv. Comm. Note 2000 Amend.

80. *Id.*

81. See *infra* notes 82–85.

82. *Johnson v. Arkema, Inc.*, 685 F.3d 452, 459 (5th Cir. 2012); *Benton v. Deli Mgmt.*, 396 F. Supp. 3d 1261, 1280 (N.D. Ga. 2019).

the validation studies.⁸³ For example, they stress the size of the study⁸⁴ or the error rate reported in the study.⁸⁵

The opinions described in the preceding paragraph make short shrift of the validity-as-applied issue. In particular, they overlook the qualitative aspect of the validity-as-applied issue: Do the facts of the pending case fall within the salient parameters of the validation studies such as the type of subjects in the study and the test conditions?⁸⁶ However, there are some exemplary opinions such as the 2019 federal district court opinion in *United States v. Gissantaner*.⁸⁷ Gissantaner was charged with being a

83. See *infra* notes 84–85.

84. E.g., *United States v. Plaza*, 188 F. Supp. 2d 549, 556 (E.D. Pa. 2002); *United States v. Starzeczyzel*, 880 F. Supp. 1027, 1036–37 (S.D.N.Y. 1995).

The courts' early treatment of the issue of the validity of Shaken Baby Syndrome (SBS) testimony also evidences the courts' stress on numbers. See generally Imwinkelried, *Shaken Baby Syndrome: A Genuine Battle of the Scientific and (Non-Scientific) Experts*, 46 *Crim. L. Bull.* 156 (2010) (describing the significance of shaken baby syndrome and outlining how annual number of prosecutions based on the syndrome has increased leading to thousands of convictions). The theory is that the violent shaking of an infant can cause the infant to sustain fatal brain injury even if, in the shaking incident, the infant's head does not strike any objects such as a wall. *Id.* at 3. Relying on studies involving primates and anthropomorphic models of infant heads and necks, today many biomechanical experts reject the theory; they claim that their studies show that shaking alone cannot generate the forces necessary to cause fatal brain injury. *Id.* at 8–9. However, the early cases were receptive to SBS testimony. *Id.* at 12. Those cases noted that researchers had documented tens of cases in which infants sustained fatal brain injuries, there was evidence of shaking, but there was no evidence of striking. *Id.*

While the early SBS cases relied on the large number of cases of fatal brain injury with shaking but without striking, many statistical testimony cases indicate the courts' reluctance to admit that type of testimony when the sample consists of a small number. *Pippin v. Burlington Res. Oil & Gas Co.*, 440 F.3d 1186, 1198 (10th Cir. 2006) (“only a very small sample size”); *Tinker v. Sears, Roebuck & Co.*, 127 F.3d 519, 524 (6th Cir. 1997) (“A sample which is too small”); *Fisher v. Vassar Coll.*, 70 F.3d 1420, 1446 (2d Cir. 1995) (“the small size” of the sample), cert. denied, 522 U.S. 1075 (1998); *United States v. Starks*, 99 F. Supp. 3d 227, 231 (D. Mass. 2014) (“small sample size”); *Holopirek v. Kennedy & Coe, LLC*, 303 F. Supp. 2d 1223, 1239 (D. Kan. 2004) (“small sample size”); *Jackson v. Univ. of New Haven*, 228 F. Supp. 2d 156, 165 (D. Conn. 2002) (“exceedingly small sample sizes”).

85. E.g., *Plaza*, 188 F. Supp. 2d at 556; *Williamson v. Reynolds*, 904 F. Supp. 1529, 1556–57 (E.D. Okla. 1995); *Starzeczyzel*, 880 F. Supp. at 1037; *Somers v. State*, 368 S.W.3d 528, 544 (Tex. Crim. App. 2012) (“[A]ccording to the various reports, studies, and testimony from the records of the cases cited, EMIT tests are over 95% accurate. When testing specifically for the presence of cocaine, one report stated that EMIT tests are 99% accurate.”).

86. See *supra* notes 82–85.

87. 417 F. Supp. 3d 857 (W.D. Mich. 2019).

felon in possession of a firearm.⁸⁸ “The case against Gissantaner rest[ed] fundamentally, if not entirely, on a small amount of ‘touch’ DNA taken from” the gun in question.⁸⁹ The rub was that there was a mixture of DNA from at least three sources on the gun.⁹⁰ The prosecution proffered testimony based on STRmix, a probabilistic genomic software program that is designed to identify the contributors to complex DNA mixtures:

The DNA analysis produced a report based on STRmix™ probabilistic genotyping software that Gissantaner was a 7% minor contributor of the DNA on the gun, and that it was at least 49 million times more likely that the DNA was that of Gissantaner and two unrelated, unknown individuals, than that the DNA was that of three unrelated, unknown contributors.⁹¹

A major hurdle for the prosecution was the limited parameters of the validation studies for probabilistic genotyping software such as STRmix.⁹² As PCAST observed in its 2016 report:

[C]urrent studies have adequately explored only a limited range of mixture types (with respect to number of contributors, ratio of minor contributors, and total amount of DNA). The two most widely used methods (STRMix and True Allele) appear to be reliable within a certain range, based on the available evidence Specifically, these methods appear to be reliable for three-person mixtures in which the minor contributor constitutes at least 20% of the intact DNA in the mixture and in which the DNA amount exceeds the minimum level required for the method.⁹³

After citing the PCAST report and carefully dissecting the studies investigating the foundational validity of probabilistic genotyping software, Judge Jane Neff concluded that the facts in the *Gissantaner* prosecution exceeded the “extent,”⁹⁴ “limits,”⁹⁵ “parameters,”⁹⁶ and “range”⁹⁷ of validation established in the studies.

—To begin with, it was debatable whether there were only three contributors to the mixture. One of Judge Neff’s court-appointed experts

88. *Id.* at 859.

89. *Id.*

90. *Id.* at 859–60.

91. *Id.*

92. *Gissantaner*, 417 F. Supp. 3d at 882.

93. PCAST, *supra* note 18, at 80.

94. *Gissantaner*, 417 F. Supp. 3d at 878–79.

95. *Id.* at 877.

96. *Id.* at 882.

97. *Id.* at 880–81, 883–84.

pointed out that one of the prosecution's experts "was open to the possibility that [the mixture] arose from four contributors."⁹⁸

—In addition, the prosecution experts had concluded that Gissantaner was "the lowest level contributor" to the mixture and his DNA represented only "approximately 7% of the total amount of DNA—corresponding to just less than 50 picograms of DNA"⁹⁹

—Finally, Judge Neff repeatedly stressed that the case involved "'touch' DNA—for example, tiny quantities of DNA left by multiple individuals on the steering wheel of a car."¹⁰⁰ "Touch" or Low-copy number (LCN) DNA analysis is more sensitive than more established DNA typing techniques such as STR.¹⁰¹ That increased sensitivity heightens the risk of both allelic drop-in (due to contamination) and allelic drop-out (due to the small amount of DNA being analyzed).¹⁰²

Judge Neff engaged in a rigorous analysis of the validity as applied issue.¹⁰³ She appreciated the connection between the studies establishing the foundational validity of probabilistic genotyping software and the validity as applied issue.¹⁰⁴ In her words, the parameters of the studies determine the "outer bounds"¹⁰⁵ or "outer limits"¹⁰⁶ of the proper application of the technique. They determine not only positively when it is proper to use the methodology but, just as importantly, negatively when an expert may not rely on the studies to justify the use of the methodology in a particular case.¹⁰⁷ She ultimately ruled that "under the circumstances presented in this case"¹⁰⁸—the possibility of four contributors, a seven percent contribution, and a miniscule LCN sample—the prosecution's testimony did not pass muster under the validity as applied requirement.¹⁰⁹ Simply stated, the facts of the *Gissantaner* case exceeded the range of validation established in the studies of foundational validity.¹¹⁰

98. *Id.* at 877.

99. *Gissantaner*, 417 F. Supp. 3d at 877.

100. *Id.* at 868–69.

101. *Id.*

102. *Id.* at 885.

103. *Id.* at 860–61, 875–76.

104. *See Gissantaner*, 417 F. Supp. 3d 857 at 885.

105. *Id.* at 877.

106. *Id.* at 882.

107. *Id.* at 885.

108. *Id.* at 875.

109. *Gissantaner*, 417 F. Supp. 3d 857 at 885.

110. *Id.* at 869–70.

2020] **The Admissibility of Scientific Evidence** 831

In effect, Judge Neff used the concept of range of validation to identify the limits of the type of extrapolation *Joiner* mentioned.¹¹¹

B. Validity as Applied in Science

At the outset, it is critical to distinguish between two senses of validity as applied in this context—a technical sense and a deeper, scientific sense.

1. The Technical Sense

In one obvious sense, validity as applied entails following the proper technique for applying the methodology.¹¹² Various organizations promulgate standards and guidelines for expert processes.¹¹³ When the process involves the use of instrumentation, one question is whether the expert complied with the operational guidance supplied by the manufacturer of the instrument.¹¹⁴ There are also private organizations that provide guidance.¹¹⁵ In the United States, one potential source is the American Standards Board (ASB) of the American Academy of Forensic Sciences,¹¹⁶ mentioned in *Gissantaner*.¹¹⁷ The American Society for Testing and Materials International (ASTM)¹¹⁸ and the International Organization for Standardization (ISO)¹¹⁹ perform a similar function on a global scale. Similarly, there are governmental organizations with responsibility for developing standards and best practices.¹²⁰ Initially, the Federal Bureau

111. See *id.* See also *United States v. Williams*, 382 F. Supp. 3d 928, 935, 937 (N. D. Cal. 2019) (the court excluded the analysis because it was a five-person mixture).

112. See Alice B. Lustre, Annotation, *Post-Daubert Standards for Admissibility of Scientific and Other Expert Evidence in State Courts*, 90 A.L.R.5th 453, at *3 (citing *State v. Alexander*, 364 P.3d 458 (Alaska Ct. App. 2015)) (discussing the central inquiry of *Daubert* standard for admissibility of scientific evidence which is whether the reasoning or methodology can properly be applied to the facts at issue).

113. See *infra* notes 116, 118–19; see also D. Hiep Truong, *Daubert and Judicial Review: How Does an Administrative Agency Distinguish Valid Science from Junk Science?*, 33 AKRON L. REV. 365, 367 (2000).

114. See *Deputy v. Lehman Bros.*, 345 F.3d 494, 501 (7th Cir. 2003) (discussing the expert's use of the technique employed to identify handwriting samples offered in evidence under *Daubert*).

115. See *infra* notes 116, 118–19.

116. ACADEMY STANDARDS BOARD, www.asbstandardsboard.org (last visited Jan. 14, 2020).

117. *Gissantaner*, 417 F. Supp. 3d 857.

118. ASTM INT'L, <https://www.astm.org> (last visited Jan. 14, 2020).

119. ISO, <https://www.iso.org> (last visited Jan. 14, 2020).

120. Ryan M. Goldstein, Note, *Improving Forensic Science Through State Oversight*, 90 TEX. L. REV. 225, 226–27 (2011).

of Investigation supported Technical Working Groups (TWGs).¹²¹ The TWGs gave way to the Scientific Working Groups (SWGs).¹²² In turn, the SWGs have been supplanted by the Scientific Area Committees and Subcommittees operating under the aegis of the National Institute of Standards and Technology.¹²³ These guidelines are often straightforward, for example, requiring that the instrument be heated to a certain temperature or that a particular quantity of solution be inserted.¹²⁴

2. *The Deeper Scientific Sense*

Although the right procedural steps in the correct sequence is a facet of validity as applied, there is much more to validity as applied than proper mechanics.¹²⁵ There is the more fundamental question of whether the proponent has the right to rely on the prior validation studies as support for the reliability of the specific test conducted in the instant case.¹²⁶ That right turns on the notion of the range of validation.¹²⁷

C. *The Recognition of the General Notion of Range of Validation*

The notion of a range of validation is well recognized in modern science.¹²⁸ The notion is perhaps most strictly applied to the calibration of measuring instruments in metrology.¹²⁹ A National Institute of Standards and Technology publication declares:

It is a generally accepted principle of reliable analysis that chemical analyzers should be calibrated over the full range of measurement and that measurement data be restricted to the range calibrated. It is not good measurement practice to report extrapolated data, i.e., outside the range

121. FED. BUREAU OF INVESTIGATION, *Scientific Working Groups*, <https://www2.fbi.gov/hq/lab/html/swg.htm> (last visited Mar. 10, 2020).

122. Simon A. Cole, *Who Will Regulate American Forensic Science?* 48 SETON HALL L. REV. 563, 576 (2018).

123. ORG. SCI. AREA COMMITTEES FOR FORENSIC SCI., *What is OSAC*, <https://www.nist.gov/topics/organization-scientific-area-committees-forensic-science> (last visited Jan. 14, 2020).

124. NSIT CTR. FOR NEUTRON RES., *Instrumentation*, <https://www.nist.gov/ncnr/neutron-instruments/general-info-and-layout> (last visited Mar. 10, 2020).

125. Goldstein, *supra* note 120, at 257 n.244 (discussing the need for a more comprehensive strategy for fostering and improving peer-reviewed scientific research including addressing issues of accuracy, reliability, and validity in forensic science disputes).

126. *Id.* at 227.

127. Vosk, *supra* note 37, at 106.

128. *Id.*

129. *Id.* at 106–07.

calibrated. The range of reliable calibration can be considered the range of reliable measurement and conversely.¹³⁰

The publication emphatically states:

Standards should never be used in an extrapolative mode. They should always bracket the measurement range. No measurement should be reported at a value lower or higher than the lowest or highest standard used to calibrate the measurement process.¹³¹

Bluntly stated, there is no empirical justification for having confidence in any measurement falling outside the range of calibration.¹³²

On occasion, the courts have recognized that the range of calibration defines the limit of reliable measurement.¹³³ In one Pennsylvania case,¹³⁴ the court applied the concept to the calibration of breath testing machines:

. . . those devices' operational calibration and consequent display of a BAC reading cannot be reliably and scientifically verified due to the limited operation field calibration range of 0.05% to 0.15%. Thus, the utilization of any instrument reading above or below that limited . . . range cannot, as a matter of science and therefore as a matter of law, satisfy the Commonwealth's burden of proof beyond a reasonable doubt on an essential element of a charged offense.¹³⁵

The metrological teaching about the range of calibration is an illustration of a broader principle.¹³⁶ As a general proposition, if the available research validates an expert methodology only within a certain range¹³⁷ or under specified conditions,¹³⁸ without more there is no empirical justification for extrapolating beyond that range.¹³⁹ It is, of course, possible that the extrapolation is correct.¹⁴⁰ However, if the litmus test is adequate empirical support, the extrapolation runs afoul of the test because there is no empirical support beyond the established range.¹⁴¹

130. John K. Taylor, U.S. Dep't of Commerce, *Standard Reference Materials: Handbook for SRM Users*, Special Pub. 260-100, 5 (Sept. 1985).

131. *Id.*; see also Int'l Org. for Standardization, *Linear Calibration Using Reference Materials*, ISO 11095 § 5.3.2 (1996).

132. Vosk, *supra* note 37, at 107.

133. Commonwealth v. Schildt, No. 2191-CR 2010, Opinion (Dauphin Co. Ct. of Common Pleas (Dec. 31, 2012)) (cited in Vosk, *supra* note 37, at 107).

134. *Id.*

135. *Id.*

136. PCAST, *supra* note 18, at 81.

137. *Id.* at 80-82.

138. *Id.* at 57.

139. *Id.* at 82.

140. *Id.* at 96 (discussing speculation beyond a certain range).

141. PCAST, *supra* note 18, at 82.

D. Specific Examples of the Notion of Range of Validation

There are numerous examples of the range of validation in scientific evidence.

—*Gissantaner* is an illustration with respect of probabilistic genotyping software.¹⁴² As previously stated, the 2016 PCAST Report canvassed the available validation studies and concluded that the empirical data in those studies validated the use of software such as STR mix only when: There are no more than three contributors, the minor contributor is responsible for at least twenty percent of the mixture, and the DNA mixture is of a certain minimum size.¹⁴³ After surveying the studies, Judge Neff agreed with PCAST's conclusion.¹⁴⁴ She consequently excluded the proffered prosecution testimony because it exceeded the range of validation in several respects: There may have been four contributors, the minor contributor represented only seven percent of the mixture, and the size of the mixture was only trace or LCN.¹⁴⁵

—The notion also comes into play with respect to drug identification techniques.¹⁴⁶ One of the most popular techniques is thin layer chromatography (TLC).¹⁴⁷ In a TLC test, the analyst extracts some of the unknown drug and spots the extract on the bottom of a glass plate coated with an adsorbent such as silica gel.¹⁴⁸ The analyst then places the plate in a tank containing a solution.¹⁴⁹ The solution creeps up the plate by capillary action in roughly the same way as a liquid moves up a blotter.¹⁵⁰ After a predetermined time period, the analyst removes the plate

142. *Gissantaner*, 417 F. Supp. 3d 857.

143. PCAST, *supra* note 18, at 80. See Bauer, Butt, Hornyak & Perlin, *Validating TrueAllele Interpretation of DNA Mixtures Containing up to Ten Unknown Contributors*, 65 J. FORENSIC SCI. 380 (2020).

144. *Gissantaner*, 417 F. Supp. 3d at 885.

145. *Id.* The American Association of Forensic Sciences Standards Board has promulgated an approved American National Standard, ANSI/ASB Standard 020 First Edition: Standard for Validation Studies of DNA Mixtures, and Development and Verification of a Laboratory's Mixture Interpretation Protocol (2018). At several points the standard emphasizes that analysts must respect the "limitations" of their methodology and not apply the methodology to a fact situation exceeding the "parameters" of the extant validation standards. See §§ 3.2, 3.5, 4.3, 4.3.1, 4.3.2, 4.33. and 4.4.1. Press, *Two New Forensic DNA Standards Added to the OSAC Registry*, AAFS News Alert for 5/15/2020 (citing Standard 020, the article states that "the new standards require that labs not interpret DNA mixtures that go beyond what they have validated and verified").

146. I P. GIANNELLI, E. IMWINKELRIED, A. ROTH & J. CAMPBELL MORIARTY, SCIENTIFIC EVIDENCE § 23.02(d) (5th ed. 2012).

147. A. MOENSSENS, B. DESPORTES & S. BENJAMIN, SCIENTIFIC EVIDENCE IN CIVIL AND CRIMINAL CASES, §15.06.2.e (7th ed. 2017).

148. Imwinkelried, *Debate*, *supra* note 23, at 30.

149. *Id.*

150. *Id.*

from the tank.¹⁵¹ The analyst next sprays the plate with a visualizing agent.¹⁵² After the spray, a streak of a particular length and color appears.¹⁵³ The distance can be stated as an Rf value—the ratio to front.¹⁵⁴ To compute the Rf value, the analyst compares the distance traveled by the known solution (the front) and the distance covered by the unknown solution.¹⁵⁵ If the unknown traveled only half as far as the solution, the Rf value would be .50.¹⁵⁶ The Rf value is some evidence of the identity of the drug.¹⁵⁷ To identify the drug, the analyst next compares the unknown's Rf value to a published database of Rf values.¹⁵⁸ At this point, the notion of range of validation comes into play:

Rf values are valid only for particular sets of conditions, namely, certain combinations of adsorbent and solvent front. Unless the analyst [reviews the validation studies to learn] the conditions under which a published Rf value was produced, he or she cannot use that value for comparison with the Rf value in [the] test run in the laboratory. If the published value was obtained with a different adsorbent or solvent, the comparison is invalid. You cannot validly compare an apples test to published data for oranges.¹⁵⁹

—Gas chromatography (GC) is another technique commonly used in forensic drug identification.¹⁶⁰ The notion of range of validation is also relevant here.¹⁶¹ GC has a superior capability to separate the components of an unknown drug compound.¹⁶² In a GC test, the analyst reduces the compound to gaseous form and sends the gas through a specially coated column.¹⁶³ The unknown is heated to a certain temperature before insertion in the column, and a carrier gas moves the unknown through the column.¹⁶⁴ Different drugs migrate through the column at varying speeds or retention times (Rt); when each component exits the

151. *Id.*

152. *Id.*

153. Imwinkelried, *Debate*, *supra* note 23, at 30.

154. *Id.*

155. *Id.* at 30–31.

156. *Id.* at 31.

157. *Id.*

158. Imwinkelried, *Debate*, *supra* note 23, at 31.

159. 2 GIANNELLI, IMWINKELRIED, ROTH & CAMPBELL MORIARTY, *supra* note 47, at § 23.02[d] 632.

160. Imwinkelried, *Debate*, *supra* note 23, at 31.

161. MOENSSENS, DESPORTES & BENJAMIN, *supra* note 147, at § 15.06.2.b.

162. Imwinkelried, *Debate*, *supra* note 23, at 31–32.

163. *Id.* at 31.

164. *Id.*

column, a detector records the Rt.¹⁶⁵ Like the Rf value in a TLC test, the Rt in a GC test is a clue to the identity of the drug.¹⁶⁶ The analyst checks the unknown's Rt against a published database of Rt values for various drugs.¹⁶⁷ However, before using the database, the analyst must know the temperature, coating, and carrier gas in the prior studies compiling the Rt values.¹⁶⁸ There is a direct parallel to TLC testing:

The analyst must know the[] column conditions before using the published Rt. Without that knowledge, the analyst could be using an apples published time to evaluate an oranges test result.¹⁶⁹

—Just as forensic chemists attempt to identify contraband substances in drug prosecutions, in arson investigations the analyst may attempt to identify an accelerant in the fire debris.¹⁷⁰ To do so, the analyst may employ a gas chromatograph/mass spectrometer (GC/MS).¹⁷¹ In this instrument, a gas chromatograph (GC) is conjoined with a mass spectrometer (MS).¹⁷² After the unknown components emerge or elute from the GC column, they enter the MS where they are fragmented by ions.¹⁷³ The instrument produces a mass spectrum displaying peaks identifying the fragments.¹⁷⁴ Arson investigators have extensive libraries of spectra for various accelerants.¹⁷⁵ For example, there is an Ignitable Liquids Reference Collection (ILRC) for GC-MS.¹⁷⁶ However, the vast majority of the spectra in the ILRC are for relatively clean, pristine accelerant samples.¹⁷⁷ The sample collected at the fire scene may not be in that condition.¹⁷⁸ In many cases, the dangerous conditions at the scene might delay the immediate collection of the debris samples; the fire may have compromised the integrity of the structure where the fire occurred, and it may be unsafe for investigators to enter the premises until the

165. Joanna Gin & Edward J. Imwinkelried, *Gas Chromatography-Mass Spectrometry (GC/MS): In Scientific Evidence, Even "Gold Standard" Techniques Require Close Scrutiny*, 56 CRIM. L. BULL. 109 (2020).

166. *Id.*

167. *Id.*

168. *See id.*

169. 2 GIANNELLI, IMWINKELRIED, ROTH & CAMPBELL MORIARTY, *supra* note 47, at § 23.02[d] 638.

170. Gin & Imwinkelried, *supra* note 165, at 18.

171. MOENSSENS, DESPORTES & BENJAMIN, *supra* note 147, at § 6.11.

172. Gin & Imwinkelried, *supra* note 165, at 7, 10.

173. 2 GIANNELLI, IMWINKELRIED, ROTH & CAMPBELL MORIARTY, *supra* note 47, at § 23.03[c].

174. Gin & Imwinkelried, *supra* note 165, at 6, 9, 14.

175. *Id.*

176. *Id.*

177. *Id.*

178. *Id.*

structure is shored up.¹⁷⁹ During the delay, the debris can be subjected to weathering and/or microbial degradation.¹⁸⁰ Weathering may cause the evaporation of the sample, and microbial activity can degrade the sample.¹⁸¹ More to the point, weathering and microbial activity can change the appearance of the spectrum produced by the MS.¹⁸² Although the ILRC contains spectra for well over 1,000 ignitable liquids, the collection includes only forty-six weathered samples and a mere twenty-eight samples that have been subjected to microbial activity.¹⁸³ When the sample has been weathered or attacked by microbial activity, it exceeds the established range of validation to consult the database of spectra of pristine samples to identify an accelerant in the debris.¹⁸⁴

—The range of validation also helps explain the evolution of the case law on the voiceprint or sound spectrography technique.¹⁸⁵ In this technique, spectrography is used to generate a visual display of a person's speech.¹⁸⁶ The spectrum indicates the speaker's frequency (vertical axis), time (horizontal axis), and amplitude (relative darkness).¹⁸⁷ Some of the early cases were receptive to sound spectrography testimony.¹⁸⁸ However, the courts gradually began to realize the limited parameters of the early validation studies.¹⁸⁹ *People v. Law* is illustrative.¹⁹⁰ In that case, the court acknowledged that there been some validation studies seemingly indicating the accuracy of the sound spectrography technique.¹⁹¹ However, the court pointed out that in those studies, the

179. Gin & Imwinkelried, *supra* note 165.

180. *Id.*

181. *Id.*

182. J. DEHAAN & D. ICOVE, *KIRK'S FIRE INVESTIGATION* 582 (7th ed. 2012); E. STUAFFER, R. NEWMAN & J. DOLAN, *FIRE DEBRIS ANALYSIS* 320 (2008); R. NEWMAN, *INTERPRETATION OF LABORATORY DATA IN FIRE INVESTIGATION* 186 (Niamh Nic Daeid ed. 2004); Turner Williams, Sigman & Goodpaster, *A Comprehensive Study of the Alteration of Ignitable Liquids by Weathering and Microbial Degradation*, 63 J. FORENSIC SCI. 58 (2018); Turner & Goodpaster, *The Effect of Microbial Degradation of Ignitable Liquids*, 394 J. ANALYTICAL & BIOANALYTICAL CHM. 363 (2009).

183. Gin & Imwinkelried, *supra* note 165.

184. *See id.* The American Society for Testing and Materials International (ASTM) has issued E1618-19 Standard Test Method for Ignitable Liquid Residues in Extracts from Fire Debris Samples by Gas Chromatography-Mass Spectrometry. Section 9.1.1 of the standard states that the chromatogram must be matched with a chromatogram "obtained under similar conditions"

185. 1 GIANNELLI, IMWINKELRIED, ROTH & CAMPBELL MORIARTY, *supra* note 47, at § 10.05.

186. Gin & Imwinkelried, *supra* note 165, at 175.

187. *Id.* at § 10.05, at 620.

188. *E.g.*, *People v. Straehle*, 279 N.Y.S.2d 115 (Sup. Ct. Westchester Cty. 1967).

189. *People v. Law*, 40 Cal. App. 3d 69, 83 (1974).

190. *Id.* at 69.

191. *Id.* at 73.

persons spoke normally with no attempt to disguise their voice.¹⁹² In *Law*, the prosecution attempted to use sound spectrography to identify a caller who had obviously attempted to disguise his voice.¹⁹³ Without expressly referring to a “range of validation,” the court correctly found that the validation studies cited by the prosecution did not justify the extrapolation of the technique to a case involving a deliberately distorted voice.¹⁹⁴

—The advent of facial recognition techniques may pose a range of validation questions strikingly similar to the issue previously raised by sound spectrography.¹⁹⁵ NIST has now tested 127 facial recognition systems from forty-four companies.¹⁹⁶ The systems use different algorithms.¹⁹⁷ Many of the systems have been trained with data sets of exclusively Caucasian male adult faces.¹⁹⁸ It is hardly surprising then that researchers have found significantly higher error rates when the person in question was: (1) a woman;¹⁹⁹ (2) an older person (who were several times more likely to be misidentified than a person under 30 years of age),²⁰⁰ or (3) a person of color.²⁰¹

If a proponent offers facial recognition testimony based on a system trained exclusively on Caucasian male adult faces, the judge should pause before admitting the testimony in a

192. *Id.*

193. *Id.* at 72.

194. *Law*, 40 Cal. App. 3d at 88–89.

195. Drew Harwell, *Oregon Became a Testing Ground for Amazon's Facial-Recognition Policing. But What if Rekognition Gets it Wrong?*, WASH. POST (Apr. 30, 2019), <https://www.washingtonpost.com/technology/2019/04/30/amazons-facial-recognition-technology-is-supercharging-local-police/>.

196. *Id.*

197. *Id.* (“The federal agency that assesses facial-recognition algorithms, the National Institute of Standards and Technology, recently said it had tested 127 systems from 44 companies on their ‘scalability to large populations’ and accuracy in identifying ‘noncooperative subjects’ photographed ‘in the wild.’”).

198. Edward C. Baig, *Should We Ban Facial Recognition? From Companies to Cities, Debate Over Privacy Rages On*, U.S.A. TODAY (May 22, 2019), <https://www.usatoday.com/story/tech/2019/05/21/facial-recognition-your-face-may-key-unlock-future/3679717002/>.

199. NIST, REPORT ON THE EVALUATION OF 2D STILL IMAGE FACE RECOGNITION ALGORITHMS 15, <https://www.nist.gov/publications/report-evaluation-2d-still-image-face-recognition-algorithms> (last visited Jan. 22, 2020).

200. *Id.* at 52.

201. Elaine McCardle, *About Face: Northeastern Law Professors Urge Confronting the Dangers of Facial Recognition Technology*, NORTHEASTERN L. 12, 15 (2019).

case involving a woman, an elderly person, or a person of color.²⁰²

E. Incorporating the Range of Validation Concept into an Approach to Determining the Validity as if the Applied of Proffered Expert Testimony

Even after finding foundational validity under Federal Rule of Evidence (c), the trial judge must determine whether the proponent has satisfied Rule 702(d) by also showing validity as applied.²⁰³ It is submitted that in order to intelligently make that determination, the judge should proceed in the following manner: (a) initially, the judge should review the validation studies to determine the range of validation; (b) the judge ought to then determine whether the facts in the instant case fall outside that range; (c) if the facts in the pending case differ from the parameters of the range of validation, the judge must determine whether that difference is salient; and (d) finally, the judge should make his or her ultimate ruling on the question of validity as applied.

—*Determining the range of validation.* The judge’s first step must be determining the range of validation established by the available empirical studies.²⁰⁴ The 2016 PCAST Report, *Daubert*, *Joiner*, and amended Rule 702(d) concur that the proponent of expert testimony must show validity as applied as well as foundational validity.²⁰⁵ The judge cannot be content or satisfied by the proponent’s proof that there have been multiple validation studies or even that those studies not only were of large size but also yielded impressive accuracy rates.²⁰⁶ That proof may be sufficient to establish foundational validity, but it misses the mark of validity as applied.²⁰⁷ As we have seen, in order to demonstrate validity as applied, the proponent must present additional testimony about such qualitative parameters of the validation studies as the composition of the

202. Patrick Grother et al., *Face Recognition Vendor Test (FRVT) Part 3: Demographic Effects*, NATL. INST. STAND. TECH. INTERAG., 1, 2 (Dec. 12, 2019) (finding a higher incidence of false positives in facial recognition tests when the subject was not white and not a man).

203. FED. R. EVID. 702.

204. See, e.g., *Gissantaner*, 417 F. Supp. 3d at 869–70 (examining the reliable range of probabilistic genotyping computer programs).

205. PCAST, *supra* note 18, at 43; *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 594–95 (1993); *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 146–47 (1997). See also *United States v. Williams*, 382 F. Supp. 3d 928, 935, 937 (N.D. Cal. 2019) (the court excluded the interpretation of a complex DNA mixture involving five contributors); *United States v. Gomez-Paz et al.*, 2011 U.S. Dist. LEXIS 105441, 2011 WL 4345891 (D. Colo. Sept. 16, 2011) (“reliable application”).

206. PCAST, *supra* note 18, at 19.

207. *Id.*

database and the conditions under which the validation study was conducted.²⁰⁸ Without the benefit of such information, the judge cannot intelligently resolve the validity as applied question.²⁰⁹ In the words of the *Joiner* Court, the judge must decide whether the expert has overreached by engaging in unsupportable “extrapolat[ion].”²¹⁰ The Advisory Committee Note accompanying the 2000 amendment to Rule 702 reiterates that the judge must inquire whether the expert has “unjustifiably extrapolated.”²¹¹ It is impossible to assess the propriety of the extrapolation unless the judge is informed of the parameters of the studies that the expert is extrapolating from.²¹²

Suppose that the proponent withholds the details about the studies’ parameters during pretrial discovery and does not disclose those details when the judge rules on the opponent’s objection to the admission of the proponent’s testimony.²¹³ At that juncture, there should be no need for the opponent to move for discovery.²¹⁴ *Joiner* and the 2000 Advisory Committee Note make it clear that evaluating the soundness of the expert’s extrapolation is an essential part of the judge’s admissibility analysis.²¹⁵ If the proponent neglects to provide the judge with the necessary details about the validation studies, the judge should rule the proponent’s testimony inadmissible—no matter how impressive the showing of foundational validity.²¹⁶ Admissibility is a both/and proposition: The proponent must demonstrate validity as applied as well as foundational validity.²¹⁷

—*Determining whether the test conducted in the instant case falls within the range of validation.* After reviewing the information about the earlier validation studies, the judge shifts his or her attention to the

208. *Id.* at 56.

209. *Id.*

210. *Joiner*, 522 U.S. at 146.

211. FED. R. EVID. 702, Adv. Comm. Note 2000 Amend.

212. *See generally id.*

213. The study in question might be unpublished. In *Daubert*, the plaintiffs relied on an unpublished reanalysis of epidemiological data. *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 584 (1993). The Court indicated that an expert may rely on unpublished studies. *Id.* at 593–94.

214. FED. R. EVID. 702, Adv. Comm. Note 2000 Amend. (“Rule 702 simply requires that: (1) the expert be qualified; (2) the testimony address a subject matter on which the factfinder can be assisted by an expert; (3) the testimony be reliable; and (4) the testimony ‘fit’ the facts of the case.”).

215. *See Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 142 (1997); *see also id.*

216. FED. R. EVID. 702, Adv. Comm. Note 2000 Amend. (“[T]he proponent has the burden of establishing that the pertinent admissibility requirements are met. . .”).

217. PCAST, *supra* note 18, at 19.

test in the instant case.²¹⁸ Suppose that the proponent establishes that in the overwhelming majority of cases, a TLC test using a certain adsorbent and a particular solvent will yield an Rf value of .63 for the drug suspected in the instant case. The proponent then presents testimony that the TLC test in the instant case produced the same Rf value. As a matter of law, that testimony is insufficient to establish validity as applied.²¹⁹ The crucial question is whether the facts in the instant case fall within the range of validation, and the judge cannot resolve that question unless the proponent presents the judge with information about the adsorbent and solvent utilized in the present TLC test.²²⁰ As previously stated, unless the proponent furnishes the judge with that information, the judge cannot tell whether he or she is being asked to use “apples” validation studies as the basis for inferring the reliability of an “oranges” test in the instant case.²²¹

As in the previous step in analysis, the proponent should bear the burden of proof on the issue.²²² *Joiner* and the 2000 Advisory Committee Note teach that the judge must inquire into the soundness of the expert’s extrapolation.²²³ The judge cannot meaningfully compare the extrapolation (the instant TLC test result) to the things being extrapolated from (the prior validation studies) unless the judge knows the parameters of

218. FED. R. EVID. 702, Adv. Comm. Note 2000 Amend. (“The amendment specifically provides that the trial court must scrutinize not only the principles and methods used by the expert, but also whether those principles and methods have been properly applied to the facts of the case.”).

219. PCAST, *supra* note 18, at 56.

220. *Id.*

221. FED. R. EVID. 702, Adv. Comm. Note 2000 Amend.

222. *Id.* (“[T]he proponent has the burden of establishing that the pertinent admissibility requirements are met. . .”).

223. *See* Gen. Elec. Co. v. *Joiner*, 522 U.S. 136, 142 (1997); *see also id.* The Advisory Committee is considering a proposed amendment to Rule 702 that, in pertinent part, would provide:

(b) Forensic Expert Witnesses

If a witness is testifying on the basis of a forensic examination [conducted to determine whether an evidentiary sample is similar or identical to a source sample] [or: “testifying to a forensic identification”], the proponent must prove the following in addition to satisfying the requirements of Rule 702(a):

(1) The witness’s method is repeatable, reproducible, and accurate—as shown by empirical studies conducted under conditions appropriate to its intended use

Memorandum from Daniel J. Capra, Reporter, Advisory Comm. on Evidence Rules, to Advisory Comm. on Evidence Rules, Symposium on Forensic Expert Testimony, *Daubert*, and Rule 702 (Oct. 1, 2017), in Advisory Comm. on Rules of Evidence 371, 380 (2017), <http://www.uscourts.gov/sites/default/files/a3.0pdf> (<https://perma.cc/VJ5T-RAG3>).

both the instant test and the prior studies.²²⁴ In order to intelligently evaluate a comparison, the judge must have the essential information about both terms of the comparison.²²⁵ The 2019 Justice in Forensic Algorithms Act, proposed by Representative Mark Takano, recognizes that insight; the Act would guarantee a defendant discovery of “documentation of [the] conditions under which software was used relative to the conditions under which the software was tested.”²²⁶ However, in principle, if the proponent denies the judge the requisite information about the test in the instant case, the opponent has no need to seek further discovery.²²⁷ In that state of the record, the opponent is entitled to have the judge rule the proponent’s evidence inadmissible.²²⁸ The proponent has failed to establish validity as applied.²²⁹

—*If the parameters of the instant test differ from those of the prior validation studies, determining whether the difference is significant enough to preclude a finding of validity as applied.* In a recent article, two respected commentators, Mr. Timothy Lau and Professor Alex Biedermann, noted an important similarity between common-law precedent reasoning and scientific analysis.²³⁰ After urging scientists to rely on a “nearest neighbors” approach to assess AI output, they write:

Lawyers and judges are well familiar with this process flow. When confronted with a legal problem, they search through the precedent for similar cases. They then compare the legal problem at hand with the precedent cases to look for similarities or differences in the underlying facts. They then use the outcomes of the precedent cases to guide them to a . . . decision.²³¹

224. See FED. R. EVID. 702, Adv. Comm. Note 2000 Amend. (“[T]he trial court must scrutinize not only the principles and methods used by the expert, but also whether those principles and methods have been properly applied to the facts of the case.”).

225. PCAST, *supra* note 18, at 56 (“Determining whether an examiner has actually reliably applied the method requires that the procedures actually used in the case, the results obtained, and the laboratory note be made available for scientific review by others.”).

226. H.R. 4368, 116th Cong. (2019). Section 2(f) of the Act would amend Federal Rule of Criminal Procedure 16(a)(i) to add (h)(vi) granting such discovery.

227. See FED. R. EVID. 702, Adv. Comm. Note 2000 Amend. Because the proponent of the evidence bears the burden, the proponent’s failure to provide the judge with the requisite information about the test at issue should result in a ruling that the test is inadmissible, and the opponent’s discovery needs as they relate to that test are moot.

228. *Id.*

229. *Id.*

230. Timothy Lau & Alex Biedermann, *Assessing AI Output in Legal Decision-Making with Nearest Neighbors*, 2020 PENN. ST. L. REV. (forthcoming 2020) (draft at 11), <https://ssrn.com/abstract=3459870>.

231. *Id.*

In the common-law system, courts are to promote the public interest by formulating rules that serve legitimate social policies.²³² For example, a common-law court might develop a certain definition of an “offer” for contract disputes because the definition effectuates a certain complex of policies including as the protection of reasonable expectations. In other words, there is a policy rationale for the development of the definition (the rule). Assume that in the seminal case establishing “the precedent,” the court announces that definition for a fact situation involving a written proposal with certain language. Weeks, months, or years later the court is presented with a second contract dispute. The question presented is whether the court should extend the precedent from the seminal case to the second fact situation.²³³ It is sometimes said that the judge must decide whether the second case is “on point”²³⁴ or “on all fours” with the first case.

The point, though, is that the second case is never “on all fours” in the absolute sense that the facts are identical. To some extent there will always be differences between the historical circumstances of the two cases. The decisive question is this: Are the factual differences merely superficial, or do they effect a policy distinction? Despite the factual differences, is the judge confident that it would serve the same complex of social policies to invoke the precedent in the same case? In the common-law system, the judge focuses on the question of the policy significance of the factual differences. If the differences are “frivolous”²³⁵ or insignificant, under the doctrine of *stare decisis*²³⁶ the judge will ordinarily extend the precedent to the new case. In contrast, when the differences are significant in terms of policy, the judge “distinguishes” the prior precedent and refuses to apply the precedent in the new case.²³⁷

There is an important parallel in the scientific analysis of the validity as applied question. Just as there will always be some factual differences between two legal cases, there will inevitably be some differences between the conditions under which the validation studies were conducted and the conditions surrounding the test in the instant case.²³⁸ Hence, the decisive question is: Are the factual differences merely superficial, or do

232. E. BODENHEIMER, J. OAKLEY & J. LOVE, AN INTRODUCTION TO THE ANGLO-AMERICAN LEGAL SYSTEM: READINGS AND CASES 79-81 (4th ed. 2004).

233. *Id.* at 81-82.

234. *Id.* at 80.

235. *Id.*

236. *Id.* at 81-82.

237. BODENHEIMER, OAKLEY & LOVE, *supra* note 232, at 80-81.

238. Dale A. Nance, *Reliability and the Admissibility of Experts*, 34 SETON HALL L. REV. 191, 210 (1993) (“[T]he particular task at hand in a lawsuit is never replicated in research.”).

they effect a reliability distinction? Despite the factual differences, can the judge be reasonably confident that like the tests in the validation studies, the instant test will yield a reliable outcome? Now the center of attention is the reliability significance of the factual differences. If the differences are insignificant, the judge ought to permit the expert to extrapolate to the test in the pending case.

In theory, it should be easier to decide the propriety of a scientific expert's extrapolation than the soundness of the extension of a legal precedent. The former decision can certainly be made in a more objective fashion. On the one hand, there is obviously a substantial measure of subjectivity in the judge's assessment of the policy impact of the factual differences between two cases. Policies are not only intangible but also largely incommensurable. On the other hand, it is at least theoretically possible for an expert to empirically determine whether changing a parameter of prior validation studies can render a test outcome unreliable. Given enough time and resources, the expert can construct an experiment in which ground truth is known to learn whether the modification of that parameter yields an erroneous test outcome.

However, in practice the litigants may be strapped for time and resources. As the *Daubert* Court remarked, the law "must resolve disputes finally and quickly."²³⁹ In litigation, a party may not have the luxury of conducting a time-consuming experiment.²⁴⁰ Furthermore, an indigent defendant may have special difficulty obtaining the financial resources needed to hire an expert to conduct the sort of experiment described in the preceding paragraph.²⁴¹ In this light, how feasible will it be for the judge to determine whether the factual differences between the parameters of the validation studies and those of the instant test are salient enough to warrant excluding the proponent's testimony?

In some cases, the validation studies themselves will answer that question; the studies may contain enough information to enable the judge to make the determination. In addition to specifying the range in which the methodology yielded accurate results, the study might indicate a cutoff below which or a ceiling above which the methodology's results were unreliable. The investigators might have encountered and reported unreliable outcomes below or above the cutoff. The study's contents provide credible evidence that the difference is salient and that exceeding the parameter warrants the exclusion of the proponent's testimony.

239. *Daubert v. Merrell Dow Pharm., Inc.*, 509 U.S. 579, 597 (1993).

240. *Id.*

241. 1 GIANNELLI, IMWINKELRIED, ROTH & CAMPBELL MORIARTY, *supra* note 47, at Ch.4.

However, in other cases, the proponent might be able to make a plausible showing that the difference is inconsequential. By way of example, assume that in the pending case, the expert utilized a TLC test to identify an unknown drug and that the color of the streak and the Rf value indicated that the drug was cocaine. The problem is that while the expert used one solvent in the instant TLC test, the expert compared the Rf value in the test to a database based on TLC tests conducted with a different solvent.²⁴² The opponent makes the familiar argument that the expert is comparing an “apples” test result to an “oranges” database.²⁴³ At first blush, the judge might find that argument attractive and be inclined to bar the testimony about the TLC test in the pending case. However, before the judge excludes the testimony, the proponent comes forward with testimony that in TLC tests for heroin, the two solvents consistently produced identical Rf values. Here the additional testimony makes it reasonable for the judge to conclude that the factual difference is immaterial. Based on that conclusion, the judge could admit the proponent’s testimony.

In still other cases, if the prior validation studies do not contain the necessary information and the proponent does not provide the judge with enough information to make the determination, the judge has another option: Under Federal Rule of Evidence 706, the judge could appoint an expert for the court to assist the judge in making the determination.²⁴⁴ In *Gissantaner*, Judge Neff did precisely that.²⁴⁵ In order to help her decide whether the differences between the circumstances of the instant probabilistic genotyping test and the conditions in the validation studies were material, the judge appointed two experts, Dr. Michael Coble and Professor Dan Krane.²⁴⁶ Judge Neff relied heavily on their reports and testimony in ultimately concluding that the differences—the possibility of a four-person mixture, a minor contributor representing only seven percent of the mixture, and a “touch” or LCN DNA quantity—dictated a finding that validity as applied was lacking.²⁴⁷ However, Rule 706 appointments are relatively rare.²⁴⁸

By process of elimination, we come to the fact situation in which: The prior validation studies do not indicate whether a change in the parameter in question renders the test unreliable, the proponent has not

242. *Shatkin v. McDonnell Douglas Corp.*, 727 F.2d 202, 208 (2d Cir. 1984).

243. *Id.*

244. FED. R. EVID. 706.

245. *Gissantaner*, 417 F. Supp. 3d at 874.

246. *Id.*

247. *Id.* at 874–75.

248. Andrew Jurs, *Expert Evidence*, 143–44 (2019) (In one survey, only twenty-two percent of the responding judges indicated that they had ever appointed an independent expert).

provided credible testimony that the change is immaterial, and the judge has decided against appointing an expert under Rule 706 to assist the judge in the determination.²⁴⁹ What should be the outcome in this situation? Here the judge should arguably exclude the evidence. In general, *Daubert* announces that the proponent has the burden of proving that his or her foundational testimony satisfies Rule 702.²⁵⁰ *Joiner*'s discussion of extrapolation implied and the 2000 amendment to Rule 702 made it an express mandate that under Rule 702, the proponent has the burden of proving validity as applied as well as foundational validity.²⁵¹

CONCLUSION

It has been relatively clear since the Supreme Court's 1997 *Joiner* decision that the proponent of expert testimony must make a showing of validity as applied as well as foundational validity.²⁵² Unfortunately, in the past, many courts have glossed over that distinction.²⁵³ In their review of the proponent's validation studies, they have tended to focus on such considerations as the size of the study and the overall accuracy rate reported by the researchers.²⁵⁴ Those quantitative factors are highly relevant to a judge's decision on the issue of foundational validity, but they shed little light on the question of validity as applied.²⁵⁵

When the question is validity as applied, the courts must scrutinize both the validity studies and the test in the instant case more closely.²⁵⁶ As Part II explained, following Judge Neff's example in *Gissantaner*, the trial judge must review the earlier studies to determine the range of validation empirically established in the studies.²⁵⁷ Again, a hypothesis about the validity of a methodology is a conditional proposition: when certain factors or conditions are specified, what is the likely outcome of the use of the methodology? The judge must identify those factors and then inquire whether the same conditions obtained in the test are also found in

249. Gen. Elec. Co. v. Joiner, 522 U.S. 136, 144–45 (1997).

250. Daubert v. Merrell Dow Pharm., Inc., 509 U.S. 579, 592–93 (1993).

251. Gen. Elec. Co. v. Joiner, 522 U.S. 136, 146 (1997).

252. *Id.*

253. See Black, Ayala & Saffran-Brinks, *supra* note 45, at 722 (“The analysis used by pre-*Daubert* courts that applied the Rules in lieu of Frye typically involves balancing various enumerated factors, albeit without any guidance on how the factors relate to each other or how they fit into a coherent picture of the way science actually works.”).

254. *Id.*

255. *Id.*

256. *Id.*

257. See *supra* notes 103–11 and accompanying text.

the pending case.²⁵⁸ A validation study supports an inference of reliability, satisfying Rule 702 only under the conditions of the study. By happenstance, an extrapolation beyond the conditions of the study may indeed be correct, but without more, the study furnishes no empirical support or justification for the extrapolation.²⁵⁹ The expert is making an epistemological claim of a scientific nature, and there must be an empirical warrant for the claim.

In the post *Daubert* era, the jurisprudence has evolved significantly.²⁶⁰ In the period immediately after *Daubert*, some proponents believed that it was sufficient to assert that validation studies had established the reliability of the expert's methodology.²⁶¹ However, in *Joiner*²⁶² and *Kumho*,²⁶³ the Court was insistent that a trial judge may not accept an expert's conclusory, *ipse dixit* assertions of reliability. Those opinions pressured proponents to go into more detail about the contents of validation studies in their foundational testimony.²⁶⁴ As we have seen, at this juncture, proponents began providing judges with additional information about such details as the size of such studies and overall accuracy rates.²⁶⁵ That quantitative information helped judges make more informed decisions about foundational validity, but at this point in the evolution of the jurisprudence, proponents did not fully appreciate the significance of the requirement for proof of validity as applied, recognized in *Joiner*, and expressly prescribed by the 2000 amendment to Rule 702.²⁶⁶ That requirement necessitates another step in the evolution of the jurisprudence, that is, a new stage of much more in depth analyses of the qualitative facets of the validation studies.

Hopefully, this article will help refine litigants' and judges' understanding of the demands of the requirement for proof of validity as applied. There are already some exemplary decisions such as Judge Neff's opinion in *Gissantaner*. In that case, Judge Neff took the validity as applied requirement very seriously and conducted a model, critical analysis.²⁶⁷ This article has attempted to expand on the analysis in *Gissantaner*

258. See *supra* notes 54–55. See also *United States v Williams*, 382 F. Supp. 3d 928, 935, 937 (N.D. Cal. 2019) (the court excluded the expert's testimony when the expert attempted to use GlobalFiler to analyze a sample with five contributors).

259. See Black, Ayala & Saffran-Brinks, *supra* note 45, at 761.

260. See *supra* note 81 and accompanying text.

261. See *supra* note 82 and accompanying text.

262. *Gen. Elec. Co. v. Joiner*, 522 U.S. 136, 146 (1997).

263. *Kumho Tire Co., Ltd. v. Carmichael*, 526 U.S. 137, 157 (1999).

264. See *id.* at 156; see also *Joiner*, 522 U.S. at 146.

265. See *supra* notes 84–85 and accompanying text.

266. *Id.*

267. See *supra* notes 257–59 and accompanying text.

and offer trial judges a systematic approach to evaluating validity as applied: first reviewing the contents of the validation studies to determine the range of validation, then identifying the respects in which the conditions of the test in the instant case exceeded that range, next deciding whether the differences between the conditions in the validation studies and the circumstances of the test in the pending case were salient, and finally making the ultimate admissibility ruling—bearing in mind that the burden of making a distinct showing of validity as applied is squarely on the shoulders of the proponent.

As Chief Justice Rehnquist correctly observed in *Joiner*, experts “commonly” resort to extrapolation.²⁶⁸ They do so in order to establish the connection between their technique or theory and the specific facts of the case, which is the underlying premise of their opinion.²⁶⁹ Unfortunately, in many of the post-*Joiner* cases, in which experts have done so, the courts have conducted only a superficial analysis of the validity as applied to the issue, notably the propriety of the extrapolation.²⁷⁰ Few courts have taken *Joiner* and Rule 702(d) as seriously as *Gissantaner*.²⁷¹ In *Gissantaner*, Judge Neff took to heart the guidance in the Advisory Committee Note to the 2000 amendment to Rule 702 that imposed the validity as applied requirement:

The amendment specifically provides that the trial court must scrutinize not only the principles and methods used by the expert, but also whether those principles and methods have been properly applied to the facts of the case. As the court noted in *In re Paoli R.R. Yard PCB Litig.*, 35 F.3d 717, 745 (3d Cir. 1994), “any step that render the analysis unreliable . . . renders the expert’s testimony inadmissible. This is true whether the step completely changes a reliable methodology or merely misapplied that methodology.”²⁷²

As a matter of logic, proof of validity as applied is every bit as essential as proof of foundational validity. It is imperative that discerning judicial evaluations of the validity as applied issue such as the analysis in

268. *Joiner*, 522 U.S. at 146.

269. Faigman, Monahan & Slogobin, *Group to Individual (G2i) Inference in Scientific Expert Testimony*, 81 U. CHICAGO L.REV. 417 (2014).

270. See *supra* notes 82–85.

271. See *Gissantaner*, 417 F. Supp. 3d 857. But see *United States v. Williams*, 382 F. Supp. 3d 928, 935, 937 (N.D. Cal. 2019) (although Judge Orrick’s analysis is not as extensive as Judge Neff’s, in essence Judge Orrick applied the concept of validity as applied).

272. FED. R. EVID. 702, Adv. Comm. Note 2000 Amend.

2020] **The Admissibility of Scientific Evidence** 849

Gissantaner become as commonplace as experts' attempts to engage in extrapolation.²⁷³

273. Again, although the validity as applied analysis is not as in depth as the analysis in *Gissantaner*, *United States v. Williams*, 382 F. Supp. 3d 928 (N.D. Cal. 2019) is another step in the right direction. Like *Gissantaner*, *Williams* deals with probabilistic genotyping software.