

10-2-2017

It is Now Up to the Courts: "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods"

Jennifer Friedman

Jessica Brand

Follow this and additional works at: <http://digitalcommons.law.scu.edu/lawreview>



Part of the [Law Commons](#)

Recommended Citation

Jennifer Friedman and Jessica Brand, *It is Now Up to the Courts: "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods"*, 57 SANTA CLARA L. REV. 367 (2017).

Available at: <http://digitalcommons.law.scu.edu/lawreview/vol57/iss2/2>

This Article is brought to you for free and open access by the Journals at Santa Clara Law Digital Commons. It has been accepted for inclusion in Santa Clara Law Review by an authorized editor of Santa Clara Law Digital Commons. For more information, please contact sculawlibrarian@gmail.com, pamjadi@scu.edu.

**IT IS NOW UP TO THE COURTS:
“FORENSIC SCIENCE IN CRIMINAL COURTS: ENSURING
SCIENTIFIC VALIDITY OF FEATURE-COMPARISON
METHODS”**

Jennifer Friedman* and Jessica Brand**

Introduction.....	368
I. The National Academy of Sciences Report.....	370
II. The Intervening Years.....	371
III. The PCAST Report.....	374
A. Foundational Validity	375
B. Validity as Applied.....	377
C. Specific PCAST Recommendations.....	377
D. PCAST’s Addendum.....	381
IV. Specific Steps for Courts to Follow.....	382

* Jennifer Friedman is a Los Angeles County Public Defender and the Forensic Science Coordinator for the office. The views expressed in this article are her own. Jessica Brand is the Legal Director at the Fair Punishment Project. The views expressed in this article are her own.

** The authors wish to thank Gabriel Fuentes for reviewing the article and offering some very helpful suggestions.

INTRODUCTION

In this article we explore the history of what we characterize as failed attempts to reform forensic science. We describe in detail the newly issued report by the President's Advisors on Science and Technology and its attempt to evaluate the scientific validity of some of the most commonly used "feature-comparison" disciplines.¹ In that report, the committee addresses the intersection of legal admissibility and scientific validity, and it concludes that many forensic sciences do not meet the criteria for either. We then argue that forensic reform will not occur until the courts truly become gatekeepers against the admission of junk science, as the law requires. We provide a roadmap for courts to follow to properly review the admissibility of forensic science.

In its 2009 report "Strengthening Forensic Science in the United States: A Path Forward," the National Academy of Sciences (NAS) issued a scathing critique of forensic science research. "With the exception of nuclear DNA analysis," the committee of esteemed scientists wrote, "no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source."² Throughout the 350-page report, the committee reiterated that traditional forensic sciences lacked empirical data supporting the claims of individualization regularly made in the courtroom.³ The committee urged the forensic community to "develop rigorous protocols to guide these subjective interpretations and pursue equally rigorous research and evaluation programs."⁴

Yet in the last seven years, the forensic science community has made little progress validating many types of forensic analysis, and it has not scaled back the forceful conclusions of individualization—that a known sample is the source of an evidentiary sample recovered from a crime scene—regularly made by analysts in feature-comparison fields. While the federal government has started conducting scientific research into some types of forensic analysis including fingerprint comparison, that research is extremely limited.⁵ Importantly, forensic examiners

1. In the feature comparison disciplines, an examiner evaluates features or characteristics of an evidence sample, compares those features to a known and then makes a judgment about whether the evidence sample matches the known or does not match the known.

2. COMM. ON IDENTIFYING THE NEEDS OF THE FORENSIC SCIENCES CMTY., NAT'L RESEARCH COUNCIL, 228091, STRENGTHENING FORENSIC SCIENCE IN THE UNITED STATES: A PATH FORWARD 1, 7 (2009) [hereinafter NATIONAL ACADEMY OF SCIENCES REPORT].

3. *Id.* at 7.

4. *Id.* at 8.

5. See National Institute of Justice Research and Development Projects <https://www.nij.gov/topics/forensics/Pages/research-development-projects.aspx>

seem unwilling to limit the scope of their testimony. They have continued as before—and innocent people have gone to jail as a result.⁶

The recent report by the President’s Council of Advisors on Science and Technology (PCAST) recognized that forensic scientists have not heeded the warnings in the NAS report.⁷ This watershed report highlighted the lack of meaningful research establishing the scientific validity of feature-comparison forensics.⁸ For the second time in a decade, the report concluded that with few exceptions, feature-comparison scientists have not performed research establishing that examiners can do what they say they can do: reliably identify a known sample as the source of recovered trace evidence.⁹ The authors of PCAST detailed steps for forensic examiners to take to establish scientific validity, but noted that “PCAST expects that some forensic feature-comparison methods may be rejected by courts as inadmissible because they lack adequate evidence of scientific validity.”¹⁰

Given forensic scientists’ reticence to establish the accuracy and reliability of their comparison methods, courts must do just this—reject certain feature-comparison evidence—and serve as a barrier to the admission of evidence lacking an empirical foundation. Judges must understand the prerequisites for a validated scientific method, the relationship between established legal principles and scientific validity, and how to apply those principles in criminal cases. A court’s failure to understand the role of validity and reliability when evaluating the admissibility of feature-comparison evidence, and relatedly, its refusal to exclude feature-comparison evidence where the proponent does not establish its reliability or validity, calls into question the fairness and integrity of the criminal proceedings¹¹

6. According to the Innocence Project, faulty forensic science accounts for 46% of cases exonerated through DNA evidence. *Misapplication of Forensic Science*, THE INNOCENCE PROJECT, <https://www.innocenceproject.org/causes/misapplication-forensic-science/>.

7. See President’s Council of Advisors on Sci. & Tech., Exec. Office of the President, REPORT TO THE PRESIDENT, Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods 1, 1-2 (Sept. 2016) [hereinafter PCAST Report].

8. See *id.* at 4.

9. *Id.*

10. *Id.* at 122.

11. *Womack v. United States*, 350 A.2d 381, 383 (D.C.1976) (“A criminal trial is not a game, but a quest for truth.”); *State v. Behn*, 375 N.J.Super. 409, 434 (App. Div. 2005) (“the integrity of the criminal justice system is ill-served by allowing a conviction based on evidence of this quality, whether described as false, unproven or unreliable, to stand”).

I. THE NATIONAL ACADEMY OF SCIENCES REPORT

In February of 2009, the National Academy of Sciences,¹² at the direction of Congress, issued a report entitled “Strengthening Forensic Science in the United States: A Path Forward.” The NAS issued the report after Congress, in 2005, ordered it to “asses the present and future resource needs of the forensic science community,” recognizing that “there exists little to no analysis of the remaining needs of the community outside of the area of DNA.”¹³ Congress mandated that the NAS chart an agenda for the forensic science community to “ensure the reliability of the disciplines, establish enforceable standards, and promote best practices and their consistent application.”¹⁴

The members of the NAS committee included research scientists, academics, forensic scientists, pathologist, judges, a defense attorney and a former prosecutor.¹⁵ This committee heard testimony from members of the forensic science community and reviewed and evaluated numerous studies and articles submitted by forensic science stakeholders.¹⁶

After over two years of research, the NAS issued a scathing report demonstrating serious deficiencies in forensic science and in the manner in which prosecutors utilize forensic evidence in the criminal justice system. The committee found that the forensic disciplines largely lacked standardization, certification, and accreditation.¹⁷ Perhaps most importantly, the committee reached the following conclusion: “Among existing forensic methods, only nuclear DNA analysis has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between an evidentiary sample and a specific individual source.”¹⁸ Forensic scientists had conducted virtually no research establishing the validity and reliability

12. The National Academy of Sciences, an arm of the National Research Council, is a private, nonprofit, self-perpetuating society of distinguished scholars engaged in scientific and engineering research, dedicated to the furtherance of science and technology and to their use for the general welfare.

13. NATIONAL ACADEMY OF SCIENCES REPORT, *supra* note 2, at 1.

14. *Id.* at xix.

15. *Id.* at v. “The Committee was composed of a diverse and accomplished group of professionals. Seven of the 17 Committee members are prominent professionals in the forensic science community, with extensive experience in forensic analysis and practice; 11 members of the Committee are trained scientists (with expertise in physics, chemistry, biology, engineering, biostatistics, statistics, and medicine); 10 members of the Committee have Ph.Ds, 2 have MDs, 5 have JDs, and one has an M.S. in chemistry.” Harry T. Edwards, *The National Academy of Sciences Report in Forensic Sciences: What it Means for the Bench and Bar*, 51 JURIMETRICS J. 1 (2010).

16. NATIONAL ACADEMY OF SCIENCES REPORT, *supra* note 2, at xix–xx.

17. *Id.* at 6.

18. *Id.* at 100.

of most forensic science disciplines, including toolmarks, handwriting, fingerprint, hair, bitemark, and footprint analysis.¹⁹ Forensic examiners commonly reported and testified to individualization statements without empirical data supporting such statements.²⁰

The authors emphasized that for the feature-comparison fields to be generally accepted and considered valid and reliable, forensic scientists needed to conduct significant research evaluating the limitations of each discipline.²¹ It recommended that forensic scientists carefully measure the examiners' actual performance.²² The committee urged the analysts to carefully study the effects of cognitive bias²³ and human error.²⁴ It also made structural recommendations, including the creation of an independent federal entity, the National Institute of Forensic Sciences.²⁵

Because the authors wrote the report to provide suggestions to forensic practitioners about improving research to make forensic science reliable, it did not specifically address issues of legal admissibility. But the authors anticipated that the courts would utilize the report's findings when assessing the admissibility of that evidence.²⁶ The co-chair of the committee, Honorable Harry T. Edwards, stated in a presentation to judges the year after the report's publication: "[I]t seemed quite obvious . . . that if a particular forensic methodology or practice, once thought to be scientifically valid, has been revealed to lack validation or reliability, no prosecutor would offer evidence derived from that discipline without taking the new information into account and no judge would continue to admit such evidence without considering the new information regarding the scientific validity and reliability of its source."²⁷

II. THE INTERVENING YEARS

In the wake of NAS, courts largely ignored the report's findings and continued to allow forensic scientists, particularly in the pattern-impression disciplines,²⁸ to testify to individualization statements

19. NATIONAL ACADEMY OF SCIENCES REPORT, *supra* note 2, at 7-8.

20. *Id.* at 7. "With the exception of nuclear DNA analysis, however, no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source."

21. *Id.* at 8.

22. *Id.* at 24.

23. A cognitive bias is a mistake in reasoning, evaluating, remembering, or other cognitive process, often occurring as a result of holding onto one's preferences and beliefs regardless of contrary information. *Id.* at 122.

24. *Id.*

25. NATIONAL ACADEMY OF SCIENCES REPORT, *supra* note 2, at 19.

26. *See id.* at 5-6.

27. Edwards, *supra* note 10, at 5.

28. PCAST refers to these disciplines as "feature comparison methods." PCAST REPORT, *supra* note 5, at 1.

without a scientific basis for the statements.²⁹ In the meantime, there has been little progress by forensic examiners in developing research. The National Commission on Forensic Science (NCFS), a partnership between the Department of Justice (DOJ) and the National Institute of Standards and Technology (NIST), and the Organization of Scientific Area Committees (OSAC), an infrastructure of forensic scientists under NIST, have taken steps toward improving forensic research, but significantly more is needed. While the NCSF has promulgated a number of “Views Documents” and “Recommendations” to improve the reliability of forensic science evidence, there is currently no mechanism for requiring state and local labs or prosecuting agencies to adopt these recommendations.³⁰ And while OSAC has tried to improve forensic science by adopting consensus based documentary standards, as is its mission,³¹ the process of formulating them is understandably slow. Indeed one of the first standards adopted by the OSAC received significant criticism because it did not utilize scientifically rigorous language.³²

Yet judges have largely have continued with business as usual, admitting forensic evidence largely as they did prior to the NAS Report. There are several possible explanations for judges’ hesitation to restrict the use of feature-comparison testimony in court, notwithstanding the NAS critiques. One is criminal defense attorneys’ failure to understand and adequately raise the issues in pre-trial pleadings. As Judge Nancy Gertner stated: “[T]he NAS Report’s concerns will not be fully met until advocacy changes.”³³ “[I]n the face of precedent favoring the admission of [feature-comparison identifications], the defendant [will have] to do

29. See, e.g., *United States v. Johnson*, No. 14-cr-00412-THE, 2015 U.S. Dist. LEXIS 111921 (N.D. Cal. Aug. 24, 2015) (“(T)he only ‘revelation’ identified by Defendant is a 2009 report from the National Research Council that has been considered and rejected as a basis for excluding ballistics evidence by numerous courts.”)

30. <https://www.justice.gov/ncfs/work-products-adopted-commission>

31. “The aim of the Organization of Scientific Areas Committees for Forensic Science (OSAC) is to promote technically sound, consensus based, fit-for-purpose documentary standards that are based on sound scientific principles. <https://www.nist.gov/forensics/organization-scientific-area-committees-forensic-science> “NIST has established OSAC to support the development and promulgation of forensic science consensus documentary standards and guidelines, and to ensure that a sufficient scientific basis exists for each discipline.” Forensic Science: *Organization of Scientific Area Committees (OSAC) for Forensic Science*, NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY, <https://www.nist.gov/topics/forensic-science/organization-scientific-area-committees-osac>.

32. News: NIST Statement on ASTM Standard E2329-14, NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY (Mar. 17, 2016), <https://www.nist.gov/news-events/news/2016/03/nist-statement-astm-standard-e2329-14>.

33. Judge Nancy Gertner, *Commentary on the Need for A Research Culture in the Forensic Sciences*, 58 UCLA L. REV. 789, 790 (2011).

some work—produce some data or expert testimony, real evidence suggesting the limitations of [pattern-matching].”³⁴ This is no easy task, as defense attorneys often lack resources to help them navigate the complexities of litigating the admissibility of forensic science.³⁵ It may also be that too few independent experts understand the problems in the feature-comparison field are willing to testify in *Daubert* and *Frye* hearings on behalf of the defense.³⁶ It may be that both prosecutors and the judiciary do not understand how deeply flawed the feature-comparison field’s existing research is, or do not know what it means to adhere to the scientific method. Or it may be that prosecutors do not want to understand, because it will weaken their cases.

Whatever the cause, the reluctance to exclude evidence lacking a scientific foundation is disturbing. The significance of expert testimony at trial cannot be overstated. Scientific expert testimony carries with it the “aura of special reliability and trustworthiness,” creating a grave risk that jurors will receive it without a critical eye.³⁷ Perhaps because juries view forensic testimony with unfailing trust, the use of unreliable forensic science is one of the leading causes of wrongful convictions.³⁸

Many believe that without pressure from the courts, forensic scientists will never produce research proving that their fields are scientifically valid.³⁹ There is historical precedent for courts acting as a catalyst for scientific research. When prosecutors first introduced DNA evidence in the courts, DNA analysts had not yet validated the methods used in interpretation. In *People v. Castro*,⁴⁰ the New York Supreme Court ruled that one of the lab’s methods for interpreting the DNA results was not generally accepted as reliable by the scientific community.⁴¹ The Court’s ruling set in motion a wave of research and the forensic science community developed new reliable methods for reporting DNA results.⁴²

34. *Id.* at 791.

35. Peter J. Neufeld, *The (Near) Irrelevance of Daubert to Criminal Justice: And Some Suggestions for Reform*, 95 S1 AM. J. PUB. HEALTH S107, S110 (2005).

36. *Daubert* and *Frye* are the standards governing admissibility of scientific evidence in courts across the country. See *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579 (1993); *Frye v. United States*, 293 F. 1013 (D.C. Cir. 1923). The party seeking to admit scientific evidence has the burden of proving the scientific technique or method is reliable under the *Daubert* standard and generally accepted as reliable by the scientific community under *Frye*.

37. *United States v. Downing*, 753 F.2d 1224, 1236 (3d Cir. 1985); see also *United States v. Haines*, 803 F.3d 713, 730 (5th Cir. 2015) (recognizing the significance of expert testimony to juries); *People v. Kelly*, 17 Cal. 3d 24, 31 (1976) (“Lay jurors tend to give considerable weight to ‘scientific’ evidence when presented by ‘experts’ with impressive credentials.”).

38. See Brandon L. Garrett, *Judging Innocence*, 108 COLUM. L. REV. 55, 81–84 (2008).

39. PCAST REPORT, *supra* note 5, at 122-123.

40. *People v. Castro*, 545 N.Y.S.2d 985 (Sup. Ct. 1989).

41. *Id.* at 996-998.

42. Mnookin, J. *People v. Castro Challenging the Forensic Use of DNA Evidence*,

III. THE PCAST REPORT

The next major report to evaluate the state of forensic research recommended that courts do just that—serve as a gatekeeper against the admission of questionable forensic science. In 2015, the President of the United States requested the President’s Council of Advisors on Science and Technology (PCAST) to determine “whether there were additional steps on the scientific side, beyond those already taken by the Administration in the aftermath of the highly critical 2009 National Research Council report.”⁴³ The committee deliberately addressed their report not only to forensic scientists, but also to members of the criminal justice system.⁴⁴ The committee devoted one chapter of the report to “The Role of Scientific Validity in the Courts,” and the committee made specific recommendations to both the Attorney General and to the Judiciary.⁴⁵

The PCAST committee included renowned research scientists who reviewed and evaluated over 2000 publications submitted by members of the forensic science community.⁴⁶ In addition, the committee consulted with a panel of Senior Advisors including nine current or former federal judges, a former U.S. Solicitor General, a former state Supreme Court justice, two law-school deans, and two distinguished statisticians with expertise in forensic science.⁴⁷

The resulting report focused exclusively on “feature-comparison methods,” methods that attempt to determine “whether an evidentiary sample from a crime scene is or is not associated with a potential source sample from a suspect, based on the presence of similar patterns, impressions, or other features in the sample and the source.” The report examined DNA, latent fingerprints, firearms and toolmarks, bitemarks, hair comparison, and footwear.⁴⁸ All of these disciplines belong to the field of metrology, “the science of measurement and its application.”⁴⁹ The report gave considerable attention to latent fingerprints, toolmarks and firearms, and DNA, disciplines relied on most frequently in current criminal prosecutions.

Recognizing the courts’ gatekeeping role in prohibiting the admission of unreliable scientific evidence, the PCAST committee emphasized the intersection of scientific validity and legal

Journal of Scholarly Perspectives, 3(01)(2007).

43. PCAST REPORT, *supra* note 5 at x.

44. *Id.* at 1–2.

45. *Id.* at xii–xiii.

46. *Id.* at 2.

47. *Id.*

48. *Id.* at 1.

49. PCAST REPORT, *supra* note 5, at 23.

admissibility.⁵⁰ The report focused exclusively on the *Daubert*⁵¹ standard for admissibility, but its analysis applies equally to *Frye*⁵² jurisdictions.⁵³ The authors stated explicitly that for a discipline or method to be considered scientifically valid, the proponent of the evidence must show that it is foundationally valid; that the existing method can, in principle, be validly applied to achieve accurate results; and that it is valid “as applied”: that the specific analyst in this case accurately applied the method in practice.⁵⁴

A. Foundational Validity

To be “foundationally valid,” a field must utilize a method that has been subject to “empirical testing by multiple groups, under conditions appropriate to its intended use.”⁵⁵ Those studies must show that the method is “repeatable and reproducible.”⁵⁶ A method is “repeatable” if, with a known probability, an examiner can reach the same result while analyzing samples from the same sources. A method is “reproducible” if, with known probability, different examiners can obtain the identical outcome while evaluating the same samples.⁵⁷ A method, in other words, is foundationally valid if studies show it has a “reproducible and consistent procedure for (a) identifying features within evidence samples; (b) comparing the features in two samples; and (c) determining, based on the similarity between the features in two samples, whether the samples should be declared to be a proposed identification (“matching rule”).”⁵⁸ The studies must also provide “valid estimates of the method’s accuracy,” demonstrating how often an examiner is likely to draw the wrong conclusions.⁵⁹ As the PCAST committee noted, “foundational validity” is the scientific analogue to Federal Rule of Evidence 702’s requirement that expert testimony must be the product of “reliable principles and methods.”⁶⁰

The PCAST authors described two possible ways for examiners to establish foundational validity. For objective techniques, such as single source DNA analysis and interpretation, scientists establish foundational validity through research establishing the accuracy,

50. *Id.* at 4.

51. *Daubert v. Merrell Dow Pharms.*, 509 U.S. 579, 597 (1993).

52. *Frye v. United States*, 293 F. 1013 (D.C. Ct. App. 1923).

53. See PCAST Report, *supra* note 5, at 40-43.

54. *Id.* at 4-5.

55. *Id.* at 5.

56. *Id.*

57. *Id.* at 47.

58. *Id.* at 48.

59. PCAST REPORT, *supra* note 5, at 5.

60. *Id.* at 4-5.

reproducibility, and consistency of each of its individual steps.⁶¹ In DNA, for example, experts have developed population frequencies showing the uniqueness of particular genetic codes. But for techniques that are subjective and rely heavily on human judgment—for example, analysis of complex DNA mixtures, fingerprints and toolmarks—the simplest way to demonstrate validity is through black box error rate studies, which look at how often an examiner gets the right answer when properly applying a method or technique.⁶² Through these black box studies, examiners must show that they can do what they say they can do in circumstances and conditions replicating actual case work⁶³.

The committee voiced two major concerns about the forensic community's previous attempts to skirt proving foundational validity. First, the committee addressed examiners' frequent claim that training and experience could substitute for empirical studies. The report was firm: "[N]either experience, nor judgment, nor good professional practices (such as certification programs and accreditation programs, standardized protocols, proficiency testing, and codes of ethics) can substitute for actual evidence of foundational validity and reliability."⁶⁴ The report continued:

The frequency with which a particular pattern or set of features will be observed in different samples, which is an essential element in drawing conclusions, is not a matter of "judgment." It is an empirical matter for which only empirical evidence is relevant. Similarly, an expert's expression of *confidence* based on personal professional experience or expressions of *consensus* among practitioners about the accuracy of their field is no substitute for error rates estimated from relevant studies.⁶⁵

Second, the committee addressed examiners' tendencies to make claims unsupported by empirical studies. Studies that validate the field will generally also show that the field has limitations. Statements in reports or in testimony must accurately convey those limits and the method's error rates. "Statements claiming or implying greater certainty than demonstrated by empirical evidence are scientifically invalid."⁶⁶ The committee expressed concern that examiners regularly state that they are "100 percent certain" or have a "zero error rate," statements that are not scientifically defensible:

From the standpoint of scientific validity, experts should never be

61. *Id.* at 5.

62. *Id.* at 5–6.

63. *Id.* at 48.

64. *Id.* at 6.

65. PCAST REPORT, *supra* note 5, at 6.

66. *Id.* at 6.

permitted to state or imply in court that they can draw conclusions with certainty or near-certainty (such as “zero,” “vanishingly small,” “essentially zero,” “negligible,” “minimal,” or “microscopic” error rates; “100 percent certainty” or “to a reasonable degree of scientific certainty;” or identification “to the exclusion of all other sources.”⁶⁷

B. Validity as Applied

To establish “validity as applied,” the field must show that the examiner has reliably applied the method on case-like samples in the past, that she correctly applied the method in the particular case, and that she carefully reported the error rate established through empirical testing.⁶⁸ Critically, the proponent of the evidence must also show that the samples analyzed in the case are similar to those analyzed in validation studies.⁶⁹ If an examiner analyzes an eleven-person mixture, for example, and uses a method tested or validated on a three-person mixture, the proponent of the evidence has not shown “validity as applied.” Finally, the proponent of the evidence must disclose any information that may impact or influence the analyst’s conclusions because cognitive bias is of particular concern when a technique involves subjective judgment.⁷⁰ “Validity as applied” is the analogue to Federal Rule of Evidence 702’s requirement that “the expert has reliably applied the principles and methods to the facts of the case.”⁷¹

The committee then assessed whether the thousands of studies submitted by forensic scientists in the “feature-matching” disciplines established foundational validity and, if so, whether limits existed to the conclusions an examiner could draw about whether two samples matched.⁷² Out of the seven feature comparison disciplines examined, only three fields met the criterion for foundational validity: single source DNA, simple mixed DNA, and latent fingerprints.⁷³

C. Specific PCAST Recommendations

Echoing the 2009 NAS Report, the PCAST committee found that bitemark evidence lacked foundational validity.⁷⁴ The field had conducted few empirical studies to prove validity, and disturbingly, those studies found such a high false positive rate that the committee

67. *Id.* at 54.

68. *Id.* at 6.

69. *Id.*

70. *Id.* at 10.

71. Fed. R. Evid. 702(d).

72. PCAST REPORT, *supra* note 5, at 67.

73. *Id.* at 67-122.

74. *Id.* at 83-87.

concluded that the field should not waste resources to undertake further studies.⁷⁵ PCAST delivered a simple and unequivocal message to courts: bitemark comparison evidence is scientifically invalid.⁷⁶ It has not been shown to produce reliable results and should therefore be inadmissible in criminal prosecutions.⁷⁷

The hair comparison studies did not fare much better. Of those submitted, PCAST found serious flaws in almost all of their designs. Only one had relevance to the work forensic hair examiners perform for trial, asking how often forensic hair examiners erroneously associate hairs belonging to different people.⁷⁸ The results of that study were disturbing: the study found an 11% false identification rate.⁷⁹ Even more troubling, a study conducted by the Department of Justice, in consultation with the National Association of Criminal Defense Attorneys and the Innocence Project, found that hair comparison examiners provided scientifically invalid testimony in 95% of the cases reviewed.⁸⁰

PCAST similarly found forensic examiners failed to establish foundational validity for shoeprint comparison: “PCAST finds there are no appropriate empirical studies to support the foundational validity of footwear analysis to associate shoeprints with particular shoes based on specific identifying marks (sometimes called ‘randomly acquired characteristics’).”⁸¹ Such conclusions are unsupported by any meaningful evidence or estimates of their accuracy and thus are not scientifically valid.⁸²

The biggest bombshell in the PCAST report, and the one that produced substantial backlash amongst forensic examiners, involved toolmark comparison. Like many of the other forensic disciplines developed by law enforcement rather than scientists, PCAST found that the existing “validation” studies were not properly designed to substantiate the discipline.⁸³ Of the numerous studies submitted for review, only one—the “Ames” study—was properly designed.⁸⁴ In that study, the researchers asked examiners to perform analyses that generally mirrored actual case work.⁸⁵ The results were striking.

75. *Id.* at 87.

76. *Id.* at 83-87.

77. *Id.* at 87.

78. PCAST REPORT, *supra* note 5, at 118.

79. *Id.* at 121.

80. *Id.* at 30.

81. *Id.* at 117.

82. *Id.*

83. *Id.* at 11.

84. PCAST REPORT, *supra* note 5, at 110-11.

85. *Id.*

According to this study, the error rate for firearms comparison was between 1 out of 46 and 1 out of 66, a far cry from the “100 percent certainty” frequently testified to by firearms examiners.⁸⁶

The committee concluded:

PCAST finds that firearms analysis currently falls short of the criteria for foundational validity, because there is only a single appropriately designed study to measure validity and estimate reliability. The scientific criteria for foundational validity requires more than one such study to demonstrate reproducibility.⁸⁷

The committee acknowledged that “[w]hether firearms analysis should be deemed admissible based on current evidence is a decision that belongs to the courts.”⁸⁸ But it urged courts that did admit this evidence to use caution: “If firearms analysis is allowed in court, the scientific criteria for validity as applied should be understood to require clearly reporting the error rates seen in appropriately designed black-box studies (estimated at 1 in 66 or with a 95 percent confidence limit of 1 in 46, in the one such study to date).”⁸⁹

Unlike other feature-comparison fields, PCAST noted that fingerprint examiners responded to the 2009 NAS criticisms and developed well-designed validation studies. Two recent studies—the Tangen study and the Miami Dade study—provided empirical evidence of foundational validity.⁹⁰ But PCAST emphasized that both studies produced significant error rates, debunking analysts’ frequent claims to have zero error rates.⁹¹

PCAST made the following recommendations about what should and should not be acceptable testimony by fingerprint analysts:

Conclusions of a proposed identification may be scientifically valid, provided that they are accompanied by accurate information about limitations on the reliability of the conclusion—specifically, that (1) only two properly designed studies of the foundational validity and accuracy of latent fingerprint analysis have been conducted, (2) these studies found false positive rates that could be as high as 1 error in 306 cases in one study and 1 error in 18 cases in the other, and (3) because the examiners in the studies were aware they were being tested, the actual false positive rate in casework may be higher.⁹²

PCAST also recognized that claims of higher accuracy are currently

86. *Id.* at 111.

87. *Id.* at 111-12.

88. *Id.*

89. *Id.* at 112.

90. PCAST REPORT, *supra* note 5, at 93-95.

91. *Id.*

92. *Id.* at 101.

“not warranted or scientifically justified . . . [a]dditional black-box studies are needed to clarify the reliability of the method.”⁹³

Finally, PCAST determined that not all types of DNA analysis are scientifically sound. Since 2009, the types of DNA analysts examine for criminal trials has expanded exponentially. At present, for example, DNA analysts examine miniscule samples, along with extremely complex mixed DNA samples. PCAST evaluated the methodology used to interpret single source DNA, “simple” mixed DNA samples, and complex DNA mixtures⁹⁴. The committee also examined Probabilistic Genotyping, software that interprets low level DNA samples and complex mixtures.⁹⁵

Unlike the other disciplines reviewed, the analysis and interpretation of single source sample and simple mixtures (defined as mixtures that are easily be separated into a major and minor contributor) is objective. The field has developed population frequencies showing the rarity of a genetic profile. The committee found that numerous studies validate the methods used to analyze and interpret single source and simple mixed DNA samples.⁹⁶ For those two types of DNA analysis then, the field has established foundational validity.⁹⁷ The committee did note that analysts must protect against human error, and should disclose any issues affecting the quality or reliability of their analysis, as well as any information of which the analyst was aware that might influence his conclusion.⁹⁸

In contrast, the interpretation of low level or mixed DNA samples is subjective, much like many of the other disciplines evaluated. Like those other feature-comparison disciplines, subjective DNA analysis suffers from troubling infirmities.

DNA analysis of complex mixtures is inherently difficult. Such samples result in a DNA profile that superimposes multiple individual DNA profiles. Interpreting a mixed profile is different from and more challenging than interpreting a simple profile, for many reasons. It is often impossible to tell with certainty which genetic variants are present in the mixture or how many separate individuals contributed to the mixture, let alone accurately to infer the DNA profile of each one.⁹⁹

The statistical calculation used to convey the significance of a DNA

93. *Id.* at 101–102.

94. *Id.* at 78–79.

95. *Id.* at 82.

96. PCAST REPORT, *supra* note 5, at 71.

97. *Id.* at 75.

98. *Id.*

99. *Id.* at 8.

match with mixed DNA samples, known as the Combined Probability of Inclusion (CPI) statistic, is a subjective method that relies heavily on examiners' individual judgments about what is and is not real DNA.¹⁰⁰ The field has not yet established its foundational validity.¹⁰¹ And while researchers may eventually demonstrate the foundational validity of Probabilistic Genotyping Software (PGS), no current independent empirical studies existed establishing the range in which PGS produces reliable results: "At present published evidence supports the foundational validity of analysis, with some programs, of DNA mixtures of 3 individuals in which the minor contributor constitutes at least 20 percent of the intact DNA in the mixture and in which the DNA amount exceeds the minimum required level for the method."¹⁰² The PCAST committee also found that most of these feature-comparison fields also failed the test for validity as applied.

*D. PCAST's Addendum*¹⁰³

If there are any doubts about the validity of PCAST's conclusions, the events following the issuance of the report should put them to rest. When PCAST published its findings, prosecutors asserted that the conclusions were invalid, alleging that the committee ignored significant research.¹⁰⁴

The PCAST position is that the forensic science disciplines specializing in the examination of bitemarks, firearms/toolmarks, complex DNA mixtures, tire-treads, and shoe prints each lack scientific foundational support and should not be permitted for use in the criminal courtroom. However, the opinions expressed by PCAST in their report clearly and obviously disregard large bodies of scientific evidence to the contrary and rely, at times, on unreliable and discredited research.¹⁰⁵

In response, PCAST sent out a broad request asking stakeholders to submit any additional studies PCAST failed to consider that provided

100. *Id.* at 76.

101. *Id.* at 82.

102. PCAST REPORT, *supra* note 5, at 82.

103. On January 6, 2017 PCAST approved an addendum to its report in which it addressed issues raised by a number of commentators. *See* PRESIDENT'S COUNCIL OF ADVISORS ON SCI. & TECH., EXEC. OFFICE OF THE PRESIDENT, AN ADDENDUM TO THE PCAST REPORT IN FORENSIC SCIENCE IN CRIMINAL COURTS (Jan. 2017), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensics_addendum_finalv2.pdf [hereinafter ADDENDUM TO PCAST REPORT].

104. NAT'L DIST. ATTORNEYS ASSOC., *National District Attorneys Association Slams President's Council on Science and Technology Report*, (Sept. 2, 2016), <http://www.ndaa.org/pdf/NDAA%20Press%20Release%20on%20PCAST%20Report.pdf>.

105. *Id.*

empirical support for the scientific validity of the feature matching disciplines considered in the report. No one sent PCAST such studies. The Department of Justice affirmatively stated it had no such studies to submit.¹⁰⁶

The conclusions reached by PCAST are significant and important. PCAST represents an important voice within the relevant scientific community, and courts must take its conclusions seriously. Of the feature-matching methods evaluated, only latent fingerprint comparison, single source DNA, and simple mixed DNA analysis are foundationally valid. Those other fields, therefore, do not meet the evidentiary requirements of Federal Rule of Evidence 702.

IV. SPECIFIC STEPS FOR COURTS TO FOLLOW

To properly exercise their gatekeeping function, courts should follow PCAST's recommendations and carefully scrutinize any forensic evidence proffered in a criminal trial. There are several steps courts should take to properly do this job.

First, to ensure that the parties properly litigate the admissibility of forensic evidence, all parties must have access to experts with the background and training necessary to assess foundational validity and validity as applied. Courts also should not hesitate to consult statisticians and metrologists in evaluating the empirical foundation for the testimony and deciding whether to allow its admission.

As explained above, PCAST suggests that most feature-comparison sciences are not foundationally valid and should be excluded. Courts should follow that implied recommendation. If the government contends otherwise, courts must require analyzing laboratories to disclose all studies that purportedly show foundational validity and must then carefully assess whether those show empirically that a method is scientifically valid. Once those studies are provided, courts must ask the same questions as the PCAST members did: Do the studies mirror actual casework? Are there established error rates? What is the sample size?

So that courts may examine whether a science meets validity as "applied," they must require total transparency from laboratories and issue robust and detailed discovery orders, even if not requested by defense lawyers. Laboratories must provide all quality control documents, including logs of unexpected results, corrective action files, reports to accrediting and oversight bodies, audits, and any other information documenting errors or problems in the lab that could potentially affect the quality in the lab. Analysts must be open about any potential biasing information and examiners should report the

106. ADDENDUM TO PCAST REPORT, *supra* note 103, 2–3.

information to law enforcement, prosecutors, and others with case-specific information provided prior to the analysis that may have influenced their results. Labs must provide proficiency test results and, upon request, proficiency test files. They should disclose whether the samples in the case are similar to, or differ from, the samples used in the validation studies. Examiners should inform the parties how many times the examiner has conducted the type of analysis on the type of sample in the case at hand. In a DNA case, for example, if the sample is 250 picograms with four or more contributors, the analyst should disclose how many times she has analyzed a comparable sample.

In the end, we believe that courts should refuse to admit most feature-comparison sciences because, to date, the proponent cannot show those fields are reliable and valid. As Professor Jennifer Mnookin has stated, in many cases “outright exclusion may, in some cases, indeed be warranted, and should certainly, along with more modest measures, be part of the available judicial toolkit.”¹⁰⁷ If prosecutors continue to rely on “years of precedent” or an examiner’s “training and experience,” judges can be confident that the fields have made no progress since the PCAST report. In such a case, exclusion is the only legally acceptable option.

If forensic evidence is admitted, courts must place restrictions on the expert’s testimony to the scope of the forensic discipline’s validity and reliability, preventing experts from overstating the weight of the results or implying a higher degree of certainty and a lower error rate than what studies have established empirically. “[C]ourts should never permit scientifically indefensible claims such as: “‘zero’, ‘vanishingly small’, ‘essentially zero’, ‘negligible’, ‘minimal’, or ‘microscopic’ error rates; ‘100 percent certainty’ or proof ‘to a reasonable degree of scientific certainty’; identification ‘to the exclusion of all other sources’; or a chance of error so remote as to be a ‘practical impossibility’”¹⁰⁸ nor should courts permit experts to testify to a “reasonable degree of scientific (or other type of) certainty” a phrase which has no generally accepted meaning.¹⁰⁹ Courts must ensure that examiners clearly and accurately state their results, that they present the error rate for the results as set forth in the PCAST report, and that they disclose any additional limitations to their opinions. And examiners should disclose any potential biasing or contaminating information provided prior to their analysis of the evidence.

Finally, courts must instruct juries regarding the limitations of the

107. Jennifer L. Mnookin, *The Courts, the NAS, and the Future of Forensic Science*, 75 BROOK. L. REV. 1209, 1214 (2010).

108. PCAST REPORT, *supra* note 5, at 19, 145.

109. *Id.* at 30.

expert opinion. And they must keep close watch on the government and the defense, making sure that neither party misstates or overstates the expert's opinion in argument.

If courts fulfill their responsibility to ensure that only scientifically accepted evidence is presented to juries, they will not only improve the results of criminal trials, but they will also likely catalyze the scientific community to conduct the studies necessary to demonstrate scientific validity—if it can be established.