

February 4, 2021

Mr. Monty Wilkinson, Acting Attorney General
Mr. John Carlin, Acting Deputy Attorney General
United States Department of Justice
950 Pennsylvania Avenue Northwest
Washington, DC 20530

[Re: January 13, 2021 DOJ Statement on PCAST report]

Dear Mr. Wilkinson and Mr. Carlin,

On January 13, 2021, the Department of Justice (“DOJ”) released a position statement¹ (hereinafter, the “Statement”) rejecting the scientific principles and values expressed in the 2016 President’s Council of Advisors on Science and Technology (“PCAST”) report, *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods*.² It is our belief that the Statement misstates basic scientific principles, misrepresents accepted approaches to methods testing, and undermines the advancement of forensic science.

Largely based on a previous publication by Senior Advisor on Forensic Science Ted Hunt,³ the DOJ’s belated statement comes over four years after the PCAST report’s publication, in the last days of the previous Administration. Over the past four years, numerous courts across the country have carefully considered the PCAST report—and the science on which it is based—and concluded that it required limiting firearms and toolmarks comparison testimony. The DOJ’s new position, as set forth in the Statement, is that these courts are mistaken, purportedly because its analysis undermines PCAST’s conclusions. However, this is inaccurate.

We accordingly respectfully request that the DOJ acknowledge the inaccuracies of the Statement and retract it immediately. We address the inaccuracies of the Statement’s arguments below.

The Statement argues that forensic pattern examination methods do not and cannot belong to the discipline of metrology because they do not rely on objective measurements misses PCAST’s

¹ U.S. Department of Justice, *United States Department of Justice Statement on the PCAST Report: Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* (2021), <https://www.justice.gov/olp/page/file/1352496/download>.

² PRESIDENT’S COUNCIL OF ADVISORS ON SCIENCE AND TECHNOLOGY, *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* (2016), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf (last visited Aug 16, 2018).

³ Ted Robert Hunt, *Scientific Validity and Error Rates: A Short Response to the PCAST Report*, 86 FORDHAM LAW REV. ONLINE (2017), <https://ir.lawnet.fordham.edu/flro/vol86/iss1/14/>.

conclusion entirely. The point is that pattern examinations can and should be based on metrology.

The Statement uses a series of definitions to argue rhetorically that traditional forensic pattern examination methods do not belong to the field of metrology because their properties cannot be measured quantitatively and can only be expressed qualitatively. But the practice area's need for objective measurements is precisely the criticism leveled—appropriately—by PCAST.

The Statement sets up a false choice between comparison and measurement. *All* comparisons are based on measurements, even if they are subjective ones. To the extent that forensic pattern disciplines do not currently use objective measurements, that is an opportunity for scientific development, rather than a reason to reject PCAST's conclusions. The PCAST report's affirmation of metrology as the foundation for forensic pattern examination methods provides a framework for strengthening these methods and ensuring they can be optimally used in the criminal legal process. Well-designed scientific procedures require that observations and comparisons be based on objective, reproducible measurements that have been shown to accurately describe the entity in question. In fact, research based on this premise has been funded by the federal government, including the DOJ.⁴

The Statement also uses its Uniform Language for Testimony and Reports (ULTRs)⁵ as purported proof of the non-metrological condition of traditional forensic comparison disciplines. But these testimonial guidelines were developed by the DOJ itself, without peer review, and have been described by the DOJ's representatives as internal documents akin to standard operating procedures. The Statement misapplies standards for reporting on and testifying to forensic testing outcomes to make arguments about its methodology or the type of information that the DOJ chooses to report. Even if relying on these internal documents to prove the DOJ's points were not entirely circular, the very same categorical conclusions from a ULTR for a forensic pattern comparison discipline have also been criticized as lacking empirical basis by former AAAS Chief Executive Officer Dr. Rush Holt, among others.⁶ The Innocence Project has likewise submitted comments to the DOJ objecting to the use of language in many ULTRs that surpasses what can be supported by rigorous scientific

⁴ For example, forensic scientists at the U.S. Army Criminal Investigation Laboratory joined with researchers and statisticians from the University of Lausanne in Switzerland, the University of South Dakota, and the University of Virginia to develop “a method for the statistical interpretation of friction ridge skin impression evidence” (See H.J. Swofford et al., *A method for the statistical interpretation of friction ridge skin impression evidence: Method development and validation*, 287 FORENSIC SCI. INT. 113–126 (2018)). The National Institute of Justice recently funded a study to quantitatively evaluate footwear impression evidence (See National Institute of Standards and Technology, *Quantitative Evaluation of Footwear Evidence: Advancing the Footwear Impression Comparison System (FICS) towards Casework Application*, NATIONAL INSTITUTE OF JUSTICE (2020), <https://nij.ojp.gov/funding/awards/djo-nij-20-ro-0010> (last visited Jan 20, 2021)). Researchers at the National Institute of Standards and Technology and Michigan State University established a scoring algorithm that assessed the quality of a latent print more accurately than human examiner participants (See Tarang Chugh et al., *Latent Fingerprint Value Prediction: Crowd-Based Learning*, 13 IEEE TRANS. INF. FORENSICS SECUR. 20–34 (2018)).

⁵ U.S. Department of Justice, *Uniform Language for Testimony and Reports (ULTRs)* (2020), <https://www.justice.gov/olp/uniform-language-testimony-and-reports> (last visited Jan 15, 2021).

⁶ Rush Holt, *Letter from AAAS CEO Rush Holt to Deputy Attorney General Rod Rosenstein* (2018), <https://www.aaas.org/sites/default/files/s3fs-public/Letter%20from%20AAAS%20CEO%20Rush%20Holt%20to%20Deputy%20Attorney%20General%20Rod%20Rosenstein.pdf> (last visited Jul 11, 2020).

studies.⁷ The ULTRs do not qualify as scientific literature in support of forensic science and practice as defined by the National Commission on Forensic Science.⁸

Despite its criticism of PCAST’s six key criteria for validation studies to establish foundational validity, the Statement does not dispute they are consistent with widely accepted scientific practice, even as it erroneously argues they are not necessary to validation study design.

The PCAST report summarizes “key criteria for validation studies to establish foundational validity” as (1) having sufficient sample size, (2) not providing participants with the answer, (3) using an *a priori* study design and analysis established in advance, (4) overseen by entities without a conflict of interest, (5) having a commitment to share data, software, and results so other scientists can review the results, and (6) demonstrated to be replicable.⁹ The Statement attempts to discredit PCAST’s statement of these basic principles of experimental design without identifying a single legitimate basis for doing so.

First, the Statement inexplicably refers to a “non-severable set of nine experimental design criteria,”¹⁰ when PCAST clearly enunciates only six, and offers no description of what the other three might be. Second, the Statement asserts without basis that these “nine” criteria are in fact inconsistent with PCAST’s own examples, internal laboratory standards, and authorities in experimental design. But this is simply untrue. These six criteria are indeed considered essential to scientific practice and their components are described in textbooks on quantitative experimental design.

Indeed, the Statement does not refute that these are non-controversial criteria, but instead discounts them based on the citations PCAST used. First, it faults PCAST for not providing citations for the six experimental design criteria, despite the fact that they are widely accepted and commonly understood in scientific research. Second, where PCAST did provide citations, the Statement tried to undermine the necessity of PCAST’s key criteria by pointing to language in those citations permitting scientists to select research approaches that were fit for their research purposes. Such a recommendation in no ways suggests that research study design can be unrestrained. The Statement then cites ISO 17025:2017 requirements for validation of laboratory methods, a laboratory accreditation standard, and uses the broad requirements for validation of laboratory development methods (e.g. “The validation shall be as extensive as is necessary to meet the needs of the given application or field of application”¹¹ and “The performance characteristics of validated methods, as assessed for the intended use, shall be relevant to the customers’ needs and consistent with specified

⁷ Public comments from Innocence Project and the Innocence Network on July 8, 2016 and August 26, 2016; Innocence Project letters to Ted Hunt and Kira Antell on November 15, 2017, March 22, 2018, September 25, 2018, May 31, 2019, and April 13, 2020.

⁸ See NATIONAL COMMISSION ON FORENSIC SCIENCE, *Views of the Commission Scientific Literature in Support of Forensic Science and Practice* 4 (2015), <https://www.justice.gov/archives/ncfs/file/786591/download> (last visited Nov 14, 2018).

⁹ PRESIDENT’S COUNCIL OF ADVISORS ON SCIENCE AND TECHNOLOGY, *supra* note 2, at 52-53.

¹⁰ U.S. Department of Justice, *Justice Department Publishes Statement on 2016 President’s Council of Advisors on Science and Technology Report [Press Release]* (2021), <https://www.justice.gov/opa/pr/justice-department-publishes-statement-2016-presidents-council-advisors-science-and> (last visited Jan 15, 2021).

¹¹ ISO/IEC, *ISO/IEC 17025:2017 General requirements for the competence of testing and calibration laboratories* (2017), at 7.2.2.1, <https://www.iso.org/standard/66912.html>.

requirements”¹²) as proof that PCAST’s criteria are incomplete. This again misleads. The Statement misrepresents the relevance of the standard: the fact that ISO 17025 relies on the scientist to determine the best study design is independent of the fact that the PCAST criteria would be among the criteria by which the general community of researchers will judge the rigor and fitness of the study design.

As further “proof” that the PCAST criteria were too narrow, the Statement cites an analytical clinical testing textbook,¹³ a popular social science research design textbook,¹⁴ and a metrology textbook¹⁵ to make the same point. However, these texts also advise that the research design of any study must be relevant for the application. Again, these principles are consistent and compatible with the PCAST report, which outlined basic criteria for foundational validation studies of forensic comparison disciplines. The Statement inappropriately interprets these texts to infer that quality research study designs are open for unfettered interpretation. While there are many different methodological frameworks, fitness for purpose *always* guides the selection of study design and, consequently, the criteria by which it will be evaluated.

The Statement attacks PCAST’s scientifically accepted recommendation to use black box studies to ascertain error rates—by criticizing the utility of error rates.

The Statement did not meaningfully explain its criticism of PCAST’s recommendation that black box studies be used to establish error rate. Instead, the DOJ report expounds on *which* error rates should be used. The response would be apparent to any researcher—the error rate produced by a study is generalizable to the sample population. A study with a large sample size, diversity of jurisdictions, and diversity of case-level difficulty will be more generalizable to the practice of a particular discipline across those variables. A study focusing on a single individual’s work will only be generalizable to that individual across the type of casework evaluated.

The primary message of the PCAST report is that the criminal legal system needs to have a sense of how accurate a forensic test is if it is to be relied upon when life and liberty are at stake. Clearly, if an evaluation is to be made of the performance of a forensic science discipline, a more generalizable error rate study would provide more utility for making that assessment. If black box studies were not the appropriate way to establish error rates, it would be difficult to understand why the National Institute of Justice (“NIJ”) with its limited forensic science research funds, chose to fund black box studies in tire evidence,¹⁶ bloodstain pattern analysis,¹⁷ and palmar friction ridge¹⁸ comparisons after the publication of the PCAST report. In the description of a forensic science

¹² *Id.* at 7.2.2.3.

¹³ WESTGARD, BASIC METHOD VALIDATION 198 (Westgard QC Inc., 3rd ed. 2008).

¹⁴ JOHN W. CRESSWELL & J. DAVID CRESSWELL, RESEARCH DESIGN: QUALITATIVE, QUANTITATIVE AND MIXED METHODS APPROACHES (2004).

¹⁵ CZICHOS ET AL., SPRINGER HANDBOOK OF METROLOGY AND TESTING 86 (Springer 2011).

¹⁶ National Institute of Justice, *A Black Box Study of the Accuracy and Reproducibility of Tire Evidence Examiners’ Conclusions*, NATIONAL INSTITUTE OF JUSTICE (2020), <https://nij.ojp.gov/funding/awards/2020-dq-bx-0026> (last visited Jan 20, 2021).

¹⁷ National Institute of Justice, *Black Box Evaluation of Bloodstain Pattern Analysis Conclusions*, NATIONAL INSTITUTE OF JUSTICE (2018), <https://nij.ojp.gov/funding/awards/2018-du-bx-0214> (last visited Sep 27, 2020).

¹⁸ National Institute of Justice, *Testing the Accuracy and Reliability of Palmar Friction Ridge Comparisons: A Black Box Study*, NATIONAL INSTITUTE OF JUSTICE (2017), <https://nij.ojp.gov/funding/awards/2017-dn-bx-0170> (last visited Jan 20, 2021).

research award to the FBI to study black box and white box evaluations, the FBI states: “The primary purpose of black box studies is to demonstrate the accuracy and reliability of a given process... Black box evaluations provide a means of quantifying forensic examinations for which quantitative models do not (yet) exist and, therefore, provide both an interim solution while such models are under development, as well as a means of validating such models.”¹⁹

The DOJ’s Statement is replete with additional mischaracterizations which warrant retraction.

In addition to the above-detailed fundamental flaws, the Statement also grossly misunderstands the intersection of forensic science and wrongful conviction. For example, the Statement argues that PCAST’s critique of closed set studies was misplaced, because “closed-set design . . . simulates real casework.” But closed set studies only simulate casework *if* the actual source of the impressions is present, a presumption that is completely unwarranted in real life casework. Knowing how examiners perform when the source is not present—that is, when they are faced with innocent persons—is a question of vital importance to our justice system. The Statement ignores that, at all our peril.

Likewise, the Statement misstates what is needed to help prevent the conviction of the innocent when it argues, without citation, that the National Research Council (“NRC”) found that “a wrongfully accused person’s best insurance against false incrimination is the opportunity to have the evidence retested.” It is true that the 1996 NRC report²⁰ concluded that retesting “provides an opportunity to identify and correct errors that might have been made during the course of analysis,” but the Statement fails to acknowledge that this recommendation was made about retesting with a *validated technique*. While retesting might root out human error in a validated technique, such as a sample mix-up in a DNA case as the NRC suggested, repeating unreliable science will never lead to reliable results.

Nor can it prevent against wrongful conviction, as the Innocence Project’s own experience with the unvalidated use of bite mark analysis shows. Steven Chaney, for example, was wrongfully convicted on the basis of bite mark evidence, despite the fact that he had his own expert who conducted an independent examination.²¹ Another exoneree, Keith Harward, was unable to find a bite mark expert to contradict the state’s false bite mark evidence in his case.²² To say that innocent people are protected from faulty forensic analyses by the availability of more faulty forensic analyses denies not only scientific realities but legal ones—and it defies common sense.

¹⁹ National Institute of Justice, *Black Box and White Box Forensic Examiner Evaluations - Understanding the Details*, NATIONAL INSTITUTE OF JUSTICE (2019), <https://nij.ojp.gov/funding/awards/djo-nij-19-ro-0010> (last visited Jan 20, 2021).

²⁰ NATIONAL RESEARCH COUNCIL, *THE EVALUATION OF FORENSIC DNA EVIDENCE* (1996), <https://doi.org/10.17226/5141> (last visited Jul 11, 2019).

²¹ *Ex parte Chaney*, 563 S.W.3d 239, 252 (Tex.Crim.App., 2018).

²² *See*, Keith Allen Harward, <https://innocenceproject.org/cases/keith-allen-harward/>

Conclusion

When Attorney General Jeff Sessions refused to renew the National Commission on Forensic Science, the academic research members of the National Commission on Forensic Science published an article raising concerns of prosecutor-led, rather than an independent scientist-led, forensic science evaluation. “Putting a prosecutor in charge of forensic science perpetuates an irreconcilable conflict-of-interest and reinforces the dominance of the prosecutorial perspective.”²³ The Statement—which presents its critique of the PCAST report as a response to recent court orders limiting firearms/toolmarks evidence in criminal cases²⁴—is emblematic of this very concern. It uses misleading rhetorical pretzels rather than scientific evidence to refute a scientific report; and all for the purpose of preserving specious litigation strategies.

As the federal government seeks to rebuild our nation’s scientific infrastructure, we urge the DOJ to recognize that forensic science is strengthened by the PCAST report and its recommendations, not threatened. Justice benefits when research scientists work in partnership with forensic scientists to develop a science-based foundation and application of methods used to analyze forensic evidence.

For these reasons, the Statement is fundamentally flawed, does not reflect basic principles of science, and attempts to politicize forensic science. It should not represent the position of the Department at large and should be retracted.

Respectfully submitted,



Christina Swarns
Executive Director
Innocence Project

²³ Suzanne Bell et al., *A call for more science in forensic science*, 115 PROC. NATL. ACAD. SCI. 4541–4544 (2018).

²⁴ U.S. Department of Justice, *Justice Department Publishes Statement on 2016 President’s Council of Advisors on Science and Technology Report [Press Release]* (2021), <https://www.justice.gov/opa/pr/justice-department-publishes-statement-2016-presidents-council-advisors-science-and> (last visited Jan 15, 2021).