REPORT TO THE PRESIDENT

Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods

Executive Office of the President
President's Council of Advisors on
Science and Technology

September 2016

881

Organization of Scientific Area Committees (OSAC)
Firearms and Toolmarks Subcommittee

Response to the President's Council of Advisors on Science and Technology (PCAST) Call for Additional References Regarding its Report "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods"

14 December 2016

The Organization of Scientific Area Committees (OSAC)[1] Firearms and Toolmarks Subcommittee is composed of sixteen forensic practitioners with a combined 307 years of forensic science experience. The practitioners are drawn from federal, state, county, local and private laboratories from across the country. Additionally, the subcommittee includes four non-practitioners with backgrounds in metrology, statistics, and computer science. The subcommittee's composition meets OSAC's goals of diversity of both forensic practitioners and non-practitioners. Given the responsibility of the subcommittee for informing the process of developing standards and guidelines for the forensic discipline of firearm and toolmark identification, we feel it necessary to respond to the report published by the President's Council of Advisors on Science and Technology (PCAST) and the subsequent Request for Information (RFI) distributed by PCAST co-chair Dr. Eric Lander on December 2, 2016.

The PCAST report addresses numerous subjects and seven disciplines of forensic science. We will limit our response to those portions addressing firearm and toolmark identification. We disagree with PCAST's conclusion 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." This response will outline why we find PCAST's analysis to be inaccurate.

---

[1] The purpose of the Organization of Scientific Area Committees (OSAC) is "...to strengthen the nation's use of forensic science by providing technical leadership necessary to facilitate the development and promulgation of consensus-based documentary standards and guidelines for forensic science, promoting standards and guidelines that are fit-for-purpose and based on sound scientific principles, promoting the use of OSAC standards and guidelines by accreditation and certification bodies, and establishing and maintaining working relationships with other similar organizations." https://www.nist.gov/topics/forensic-science/about-osac

1 Black-Box (Validation) Study Analysis

PCAST analyzed nine firearm black-box studies and concluded that firearms identification "falls short of the criteria for foundational validity."[2] We disagree with their position because it ignores critical details within each study and their review falls short in understanding the research value these studies provide when considered in totality. Additionally, other validation studies have been performed that were not addressed by PCAST.[3,4,5,6,7,8]

1.1 Introduction

Black-box studies (a common type of validation study) use ground truth to evaluate the soundness and accuracy of examinations.   PCAST required that a validation study be of "black-box" design and that samples be examined completely independently of each other. PCAST set the following criteria for determining if a forensic science discipline is scientifically valid: 1) at least two black-box studies that allow for the calculation of a False Positive Error Rate (FPR) and 2) an error rate less than 5%[9]. There is no reference or justification to support that this is a generally-accepted standard.

The studies examined by PCAST were categorized into four different types: "within-set," "set-to-set," "partly open set," and "independent/open."   Within these categories, PCAST examined nine validation studies and discounted the data from eight due to test design. PCAST also made errors when summarizing these studies. They did not accurately count the number of responses, or left data out, from four of the nine validation studies used for their analysis. A summary of the errors can be found in Appendix A.

---

[2] PCAST Report "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods"," (September, 2016) Finding 6, pg 112.

[3] Lyons, D. J. "The Identification of Consecutively Manufactured Extractors." *AFTE Journal*, Vol. 41, No. 3 (2009): 246-256.

[4] Bunch, S. G., and D. Murphy. "A Comprehensive Validity Study for the Forensic Examination of Cartridge Cases." *AFTE Journal*, Vol. 35, No. 2 (2003): 201-203.

[5] Mayland, B. and C. Tucker. "Validation of Obturation Marks in Consecutively Reamed Chambers." *AFTE Journal*, Vol. 44, No. 2 (2012): 167-169.

[6] Fadul, T. G. "An Empirical Study to Evaluate the Repeatability and Uniqueness of Striations/Impressions Imparted on Consecutively Manufactured Glock EBIS Gun Barrels." *AFTE Journal*, Vol. 43, No 1 (2011): 37-44.

[7] Cazes, M. and J. Goudeau. "Validation Study Results from Hi-Point Consecutively Manufactured Slides." *AFTE Journal*, Vol. 45, No. 2 (2013): 175-177.

[8] A listing and summary of additional supportive research, and validation studies pertaining to non-firearm toolmarks, can be found in the SWGGUN Admissibility Resource Kit (ARK). https://afte.org/resources/swggun-ark/testability-of-the-scientific-principle

[9] PCAST Report "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods"," (September, 2016) "Finding 6", pp 112, Appendix A, pg 152.

Below we summarize PCAST's analysis and why we disagree with their findings.

1.2 Within-set Studies

PCAST summarized two "within-set" validation studies.[10,11] The PCAST committee could not calculate a False Positive Error Rate (FPR) using these studies, so they did not use them to measure the validity of firearm and toolmark identification.

The dismissal of these studies does not accurately reflect the scientific value of the research. A total of 1037 different-source comparisons were performed. No false identifications or false eliminations were reported by any of the participants. Therefore, these two studies provide empirical and independent support that the overall error rate for firearm and toolmark identification is low, despite the inability to calculate a false positive error rate.

1.3 Set-to-Set Comparison/Closed Set Studies

PCAST summarized four "closed set" studies.[12,13,14,15] PCAST is critical of these test designs because each comparison is not independent of the others. The assumption is that examiners may be able to deconstruct the test design, and PCAST likens this to the same logic as solving a "Sudoku" puzzle.[16] The analogy used by PCAST misrepresents the challenge posed by these tests. First, three of the studies (Brundage et al., Hamby et al., Fadul et al.) used consecutively manufactured firearms. Consecutively manufactured firearms have been shown to have the potential for subclass characteristics, which are toolmarks that sometimes carry over, with very

---

[10] Smith, E. "Cartridge case and bullet comparison validation study with firearms submitted in casework." *AFTE Journal*, Vol. 37, No. 2 (2005): 130-5. There were a total of 16 same-source comparisons and 704 different-source comparisons in this study. 13 of the 16 same-source comparisons were correctly identified and 3 were inconclusive. There were no false identifications or false eliminations reported.

[11] DeFrance, C.S., and M.D. Van Arsdale. "Validation study of electrochemical rifling." *AFTE Journal*, Vol. 35, No. 1 (2003): 35-7. There were a total of 45 same-source comparisons and 333 different-source comparisons. 42 of the 45 same-source comparisons were correctly identified and 3 were inconclusive. There were no false identifications or false eliminations.

[12] Stroman, A. "Empirically determined frequency of error in cartridge case examinations using a declared double-blind format." *AFTE Journal,* Vol. 46, No. 2 (2014):157-175.

[13] Brundage, D.J. "The identification of consecutively rifled gun barrels." *AFTE Journal*, Vol. 30, No. 3 (1998): 438-44.

[14] Fadul, T.G., Hernandez, G.A., Stoiloff, S., and S. Gulati. "An empirical study to improve the scientific foundation of forensic firearm and tool mark identification utilizing 10 consecutively manufactured slides." *AFTE Journal.* Vol. 45, No. 4 (2013): 376-93.

[15] Hamby, J.E., Brundage, D.J., and J.W. Thorpe. "The identification of bullets fired from 10 consecutively rifled 9mm Ruger pistol barrels: a research project involving 507 participants from 20 countries." *AFTE Journal,* Vol. 41, No. 2 (2009): 99-110.

[16] PCAST Report "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods" (September, 2016) Section 5.5, pp 106. PCAST was quoting Jeff Salyards, Director of the Defense Forensic Science Center.

little change or variation, from one machined part to the next on the same production line.[17,18,19] Qualified examiners are able to recognize these marks so as not to use them for conclusions of identification. Though consecutively manufactured firearms are not likely to be encountered in actual casework, the authors used them in an attempt to create a worst-case scenario (i.e. potential best known non-matches). Additionally, each test used *more* questioned samples than knowns (15 questioned samples from 10 consecutively manufactured firearms). Therefore, taking these tests was not as simple as figuring out a few of the correct answers and then deducing the rest. Since these tests used consecutively manufactured samples, it was just as important to know if examiners could correctly identify samples as it was to know if samples were falsely identified. This is the reason at least one true match was provided with each questioned cartridge case.

Another study discounted by PCAST was conducted by Stroman et al. This validation study used cartridge cases that had been fired in Smith & Wesson pistols. While this study did not use consecutively-manufactured samples, the firearms were the same make and model and had documented subclass characteristics on the firearms' ejectors. Again, these are potentially difficult samples and provide the opportunity for false positive errors, yet none were observed.

In each of these four studies, the authors attempted to create tests with potentially challenging samples. Each of these studies provide insight into the overall error rate (see Appendix A for more details about each study). The fact that few false positive errors occur is strong evidence in support of the discipline of firearm and toolmark identification. These studies present evidence that firearm and toolmark examiners can reliably and accurately associate questioned toolmarks to the correct source tool. Though the test design does not fit the model proposed by PCAST, these studies present valuable performance estimates and should not be disregarded. When viewed collectively, these studies are independent of each other and show a low overall error rate among the tested examiners. This provides strong support for the overall validity of firearm and toolmark identification.

---

[17] Weller, T.J., Zheng, X.A., Thompson, R.M., and F. Tulleners. "Confocal microscopy analysis of breech face marks on fired cartridge cases from 10 consecutively manufactured pistol slides." *Journal of Forensic Sciences,* Vol. 57, No. 4 (2012): 912-17. This study has documented subclass characteristics among the 10 consecutively manufactured pistol slides. An eleventh pistol slide, that was not part of the consecutive batch, no longer has the same subclass toolmarks.

[18] Miller J., Beach G. "Toolmarks: Examining The Possibility of Subclass Characteristics" *AFTE Journal*, Vol 32, No 4: 296-345.

[19] Subclass characteristics are features that may be produced during manufacture that are consistent among items fabricated by the same tool in the same approximate state of wear. These features are not determined prior to manufacture and are more restrictive than class characteristics. AFTE Glossary, 6th Edition.

1.4 Partly Open Set

PCAST summarized another validation study and categorized it as "partly open."[20] We would like to highlight the fact that this study also uses consecutively manufactured samples and, as described above, provides examiners with test samples which are most likely to have similar toolmarks since the firearms used to create them were sequentially manufactured with the same tools.

PCAST's statistical analysis of this report focused solely on two unknowns that had no matching known. This analysis is incomplete, and differs from the analysis used by PCAST in the "set-to-set/closed set" and "open set" studies where all "conclusive" responses were used to calculate the False Positive Error Rate.

The authors' reported error rate (0.7%) was low and this study provides an additional independent study establishing that firearm and toolmark examiners can accurately associate questioned toolmarks to the correct source tool.

1.5 Open Set

PCAST summarized another validation study and categorized it as "open."[21]

Each test taker in this study was instructed to work independently and not collaborate with other test takers. These instructions negate an important quality assurance step used in most accredited forensic laboratories: the peer review process known as verification[22]. Verification is a reevaluation of a comparison by another qualified examiner to ensure there is sufficient data to support the conclusion. Many laboratories accomplish this by direct reexamination of the evidence, while others use representative photographs of sufficient quality for the verification step. The errors reported in this paper may have been caught if verification were allowed. This suggests the true false positive error rate may be lower than calculated in this study. We would like to highlight that Baldwin et al. discusses this point in their study (emphasis added):

> "This finding does not mean that 1% of the time each examiner will make a false-positive error. Nor does it mean that 1% of the time laboratories or agencies would report false positives, **since**

---

[20] Fadul, T.G., Hernandez, G.A., Stoiloff, S., and S. Gulati. "An empirical study to improve the scientific foundation of forensic firearm and tool mark identification utilizing consecutively manufactured Glock EBIS barrels with the same EBIS pattern." National Institute of Justice Grant #2010-DN-BX-K269, December 2013.

[21] Baldwin, D.P., Bajic, S.J., Morris, M., and D. Zamzow. "A study of false-positive and false-negative error rates in cartridge case comparisons." Ames Laboratory, USDOE, Technical Report #IS-5207 (2014) afte.org/uploads/documents/swggun-false- postive-false-negative-usdoe.pdf.

[22] In the other validation studies discussed above, verification was also unlikely because test takers were not to collaborate with other test takers.

1059

> **this study did not include standard or existing quality assurance procedures, such as peer review or blind reanalysis**. What this result does suggest is that quality assurance is extremely important in firearms analysis and that an effective QA system must include the means to identify and correct issues with sufficient monitoring, proficiency testing, and checking in order to find false-positive errors that may be occurring at or below the rates observed in this study. "[23]

It should be noted that PCAST used the data from the study to recalculate a false positive error rate by using only exclusion conclusions and omitting the inconclusive results. This resulted in a rise in the calculated error rate from 1.01% to 1.5%. The different error rates provide different answers for different questions: The Baldwin et al. error rate estimates how often non-matching cartridge cases are falsely identified, while PCAST's error rate estimates the proportion of definitive (i.e. not inconclusive) results that are incorrect when non-matching cartridge cases are examined.

Baldwin et al. provide a discussion about inconclusive results (emphasis added):[24]

> "If the examiner does not find sufficient matching detail to uniquely identify a common source for the known and questioned samples, and there are no class characteristics such as caliber that would preclude the cases as having been fired from the same-source firearm, **a finding of inconclusive is an appropriate answer (and not counted as an error or as a non-answer in this study).** The underlying rationale for this finding of inconclusive is that the examiner is unable to locate sufficient corresponding individual characteristics to either include or exclude an exhibit as having been fired in a particular firearm and the possible reasons are numerous as to why insufficient marks exist. As is determined in this study, there are also a significant number of times that the firearm fails to make clear and reproducible marks (which very well might have happened for a questioned case)."

Baldwin et al. found the rate of poor quality mark production to be 2.3% (+/- 1.4%). This rate is double the calculated false positive error rate. This provides support for the use of inconclusive results in the calculation of error rates.

We would like to highlight the fact that the Baldwin study found "all but two of the 22 false identification calls were made by five of 218 examiners."[25] This indicates when errors do occur, they may be committed by the same few examiners. This supports the need for rigorous

---

[23] Baldwin et al. Pg 18.
[24] Baldwin et al. Pg 6
[25] Baldwin et al. Pg 16.

training, periodic proficiency testing, continuing education and thorough laboratory quality control measures.

1.6 Smith et al. Study

The final validation study examined by PCAST was the Smith et al. study, in which the authors created a test that mimics casework[26]. PCAST concluded this study was insufficient to test the validity of firearm identification:

> "While interesting, the paper clearly is not a black-box study to assess the reliability of firearms analysis to associate ammunition with a particular gun, and its results cannot be compared to previous studies."[27]

PCAST recognizes the study as being new and novel. We disagree with their observation that since the study is not a "black-box" design then the study does not provide support for the validity of firearm identification. In the test design that PCAST requires, test takers examine only one questioned sample at a time, independent of other questioned samples. While we understand this test design allows for easier statistical analysis, one to one comparisons are not an accurate representation of actual casework. A typical examination for a firearm examiner entails opening a package of evidence with dozens of items and attempting to associate or disassociate the items. This study tested that process by forcing examiners to make all of the typical decisions they would make in casework, rather than conducting a series of examinations on isolated pairs of specimens. The test takers were presented with bullets and cartridge cases of various ammunition types, and asked to perform both class and individual characteristic evaluations. They were not given any information about the source of any of the items.

Test takers were faced with a real-world scenario and performed very well. Although not stated in the PCAST footnote referencing this article, the overall error rate for this study was 0.303%.

1.7 Conclusions

PCAST reviewed nine validation studies and through their criteria, elected to discount eight of those studies. Two of those disregarded studies (the "within-set" design) had no false positive results. Five of the disregarded studies had very few false positives (see Appendix A) and the last study (which attempted to replicate casework) found a low overall error rate (0.303%).

---

[26] Smith, T., Smith, G.A., Snipes, J.B. "A Validation Study of The Bullet and Cartridge Case Comparisons Using Samples Representative of Actual Casework." Journal of Forensic Sciences, Vol. 61, No. 4: 939-946

[27] PCAST Report "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods" (September, 2016), footnote #335.

When PCAST set criteria for the validity of a forensic science discipline, they chose an arbitrary threshold of having at least two black box studies. The black-box test design favored by PCAST requires that each questioned sample be examined independently from each other. Examiners are not faced with completely independent examinations when they analyze evidence in a case. It is not realistic, if trying to replicate casework, to have fifteen or twenty individual sets of comparisons, each of which is made independent of each other. The PCAST-proposed design may make sense from a purely statistical standpoint, but does not simulate the practical task of an examiner performing casework. The OSAC subcommittee believes that various types of tests are valuable and can provide meaningful information regarding the potential error rates[28].

2.0 Subjective and Objective Methods

PCAST defines objective feature comparison methods as "methods consisting of procedures that are each defined with enough standardized and quantifiable detail that they can be performed by either an automated system or human examiners exercising *little or no judgment*" (emphasis added). PCAST defines subjective methods as "methods including key procedures that involve *significant human judgment*[29]" (emphasis added).

In fact, all disciplines, including firearm and toolmark identification, require some human judgment or interpretation of results. Implementation of more objective techniques may make those interpretations easier, but judgment will still be required.

We agree, however, with the goal of continuing to research and implement more objective analytical methods. One of our subcommittee's task groups is writing standards that will assist industry and crime laboratories with the validation and implementation of new technology. Additionally, there is a growing body of research using three-dimensional instrumentation and advanced machine-learning algorithms to compare toolmarks. The research fails to disprove the foundational premise of firearm and toolmark identification: that fired ammunition components can be associated to (or eliminated from) the originating firearm through the comparison of microscopic toolmarks. In fact, the recent research provides strong objective

---

[28] Different test designs estimate different error rates. For example: when examining evidence from an officer involved shooting where each officer admits to firing their firearm: error rates based on data from "set to set/closed-set" studies may be more appropriate while the Smith et. al. study may provide a better estimate for an examination of numerous items with no questioned firearm. All of these studies have the potential to provide a relevant error rate estimates and the "true" error rate may not be the same for each situation.

[29] PCAST Report "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods" (September 2016) Section 4.1, pp. 46-47.

support for this premise. The PCAST committee was provided with 25 citations by our subcommittee documenting this work; however, their report only cites two studies.

3.0 AFTE Theory of Identification is Circular Logic

PCAST states that the AFTE Theory of Identification is circular logic. PCAST's summary of the theory makes it sound circular:

> "It declares that an examiner may state that two toolmarks have a "common origin" when their features are in "sufficient agreement." It then defines "sufficient agreement" as occurring when the examiner considers it a "practical impossibility" that the toolmarks have different origins."[30]

The PCAST Report makes the AFTE Theory sound circular by ignoring the basis for "sufficient agreement." This is based on a misunderstanding of what constitutes "sufficient agreement." They claim it is an arbitrary point at which the examiner considers it a "practical impossibility." PCAST seems to believe that this "practical impossibility" is arbitrarily decided by the examiner, thus making the theory sound circular. This is incorrect. The sufficient agreement threshold is exhibited when the amount of agreement is greater than best known non-matches established by the community and conveyed to each examiner through a lengthy and extensive training program. That is, it is not an arbitrary point. In fact, by definition, no non-matches can ever have more similarity than the sufficient agreement point. When the basis for the ground truth is included, the AFTE Theory is not circular.

4.0 Focus on Training and Experience Rather Than Empirical Demonstration of Accuracy

PCAST quote:

> "Many practitioners hold an honest belief that they are able to make accurate judgments about identification based on their training and experience.[31]"

In all professions, proper training and experience is critical. Firearm and toolmark identification is like other applied sciences (e.g. medicine, engineering) that require training to become proficient and experience to further refine and maintain that proficiency. There is only so much that textbooks can teach, and structured training (like residency for physicians) is a critical aspect of developing proficiency. It is through rigorous training that examiners develop their criteria for what constitutes an elimination, an identification, or an inconclusive result. They

---

[30] PCAST Report "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods" (September 2016) Section 4.7, pp. 60.

[31] PCAST Report "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods" (September 2016) Section 4.7, pp 60-61.

learn and understand the differences in microscopic agreement between toolmarks created by the same source (a known match) and toolmarks created by different sources (a known non-match) and how that understanding factors into any conclusion of elimination, inconclusive, or identification. Examiners do not memorize all patterns that have been observed, as suggested in the PCAST report.

5.0 Conclusion

The Firearms and Toolmarks Subcommittee of OSAC fundamentally disagrees with the conclusions regarding the firearm and toolmark identification discipline presented in the PCAST report. Four major points have been put forth in this response. First, we disagree with the premise that a structured black-box study is the only useful way to gain insight into both the foundations of firearm and toolmark identification and examiner error rates. Taken collectively, the published studies support the underlying principles of firearm and toolmark examination and the fact that examiner error rates are quite low. PCAST's critique of these studies included several misunderstandings. Second, PCAST's dismissal of methods employing a subjective component discounts the core scientific methods that have been used for hundreds of years. Third, PCAST misunderstands and misquotes the AFTE Theory of Identification. PCAST's summary of the AFTE Theory of Identification leaves out important provisions. Fourth, PCAST minimizes the value of training and experience. The training received by firearm examiners includes both subjective and objective components and is comparable to the domain-specific rigor of other applied scientific fields.

We do not agree that firearm identification "...falls short of the criteria for foundational validity." However, we do agree that a hallmark of any scientific endeavor is ongoing research and technology development. Indeed, our subcommittee, which is tasked with writing standards and providing guidance to the profession, would not exist if it was believed that the field of firearm identification is flawless and requires no improvement. As such, we are hopeful that the path forward from the PCAST report is a renewed commitment to research in the forensic sciences, continued testing of foundational principles, and a more robust collaboration between the academic and forensic practitioner communities.

Appendix A
Errors and Omissions in PCAST Summaries of Firearms and Toolmarks Validation Studies

PCAST incorrectly summarized four of the nine validation studies used in their analysis of firearm and toolmark identification. For clarity, we first repeat some of the terms used by PCAST to illustrate how they (and we) calculated these error rates.

> "The results of a given empirical study can be summarized by four values: the number of occurrences in the study of true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN)"[32]

PCAST used the following formula to calculate the "maximum likelihood estimate of FPR": FP/(FP+TN).[33] For those unfamiliar with statistics, we recalculate the FPR for the Baldwin et al. study. There were a total of 2178 different-source comparisons performed: 1421 were declared elimination, 735 were reported as inconclusive, and there were 22 false positives reported. PCAST did not use inconclusive results in their statistical treatment (as we discussed in Section 1.5). Therefore, PCAST's FPR calculation for the Baldwin et al. study is: FPR= 22/(1421+22). This equals 0.015, or 1.5%. Conversely, recognizing that inconclusive results are appropriate[34], Baldwin, et al. included inconclusive results in their calculations, as follows: FPR= 22/(1421+735+22). This equals 0.010, or 1.0%.[35]

For the "set-to-set/closed" studies, PCAST used correct identifications in lieu of using true negatives[36]. PCAST does not explain or justify why they did this. The error rates reported by PCAST for the "set-to-set/closed" studies found in Table 2 on page 111 of the PCAST report are not false positive error rates and should not be reported as such.

Below we summarize the errors made by PCAST in their assessment of four of the nine studies.

---

[32] PCAST Report "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods", Appendix A, pg 152.

[33] PCAST Report "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods", Appendix A, pg 152.

[34] Baldwin et al., pg 6.

[35] Baldwin et al., pg 16.

[36] See footnote 327 of PCAST report: "Of the 10,230 answers returned across the three studies, there were there were <u>10,205 correct assignments</u>, 23 inconclusive examinations and 2 false positives."

Brundage Study

The PCAST summary of the Brundage Study is (emphasis added):

> In this study, bullets were fired from 10 consecutively manufactured 9 millimeter Ruger P-85 semi-automatic pistol barrels. Each of 30 examiners received a test set containing **20 questioned** bullets to compare to a set of **15 standards**, containing at least one bullet fired from each of the 10 guns. Of the **300 answers returned**, there were no incorrect assignments and one inconclusive examination.

This is not correct. The Brundage study consisted of 15 questioned bullets compared to a set of 10 standards (two test fired bullets from each standard set). This test was sent to 30 examiners and 450 answers returned (30 examiners x 15 questioned bullets) with no false positives and one inconclusive conclusion.

Hamby Study

The Hamby Study was a continuation of the Brundage study. Hamby et al. used the same firearm and ten consecutively manufactured barrels to produce an additional 240 test sets. The PCAST summary of this study states (emphasis added):

> In this study, bullets were fired from 10 consecutively rifled Ruger P-85 barrels. Each of **440 examiners** received a test set consisting of 15 questioned bullets and two known standards from each of the 10 guns. Of the **6600 answers returned, there were 6593** correct assignments, seven inconclusive examinations and no false positives.

This study combined the conclusions from the Brundage study, and additional results collected with both the original Brundage test sets and the 240 new test sets. If we subtract the original 30 responses from the Brundage study, the Hamby et al. article reports an additional 477 examiners having completed the test, for a total of 7155 answers with 7148 correct assignments and 7 inconclusive conclusions.

Fadul Pistol Slides Study

The PCAST summary of the Fadul Pistol Slides Study:

> In this study, bullets were fired from 10 consecutively manufactured semi- automatic 9mm Ruger pistol slides. Each of 217 examiners received a test set consisting of 15 questioned cartridge cases and two known cartridge cases from each of the 10 guns. Of the 3255 answers returned, there were 3239 correct assignments, 14 inconclusive examinations and two false positives.

This summary is correct; however, it is incomplete because it only includes Phase 1 of the study. It does not include the second phase of the study, the durability study. Results for Phase 1 and 2 are included in the same report. In Phase 2, an additional 114 examiners participated. The examiners received 5 more questioned cartridge cases (after the firearm had been fired 1000 times) and were asked to compare these cartridge cases to the 10 cartridge cases from the knowns that were previously received. A total of 570 answers were returned with 564 correct assignments, 5 inconclusive and one false positive.

Fadul EBIS Barrels Study
The PCAST summary of this study states (emphasis added):

> The 165 examiners in the study were asked to assign a collection of **15 questioned** samples, fired from **10 pistols**, to a collection of known standards; two of the 15 questioned samples came from a gun for which known standards were not provided.

This is not correct. Each test consisted of two known standards from each of the 8 pistols and 10 questioned samples. One of the known pistols had no matching questioned samples. Additionally, two of the unknowns had no matching known pistol.

Fadul et al. reported an <u>overall</u> error rate of 0.7% (95% lower bound 0.2%, 95% upper bound 1.2%).

13

AN ADDENDUM TO THE PCAST REPORT ON FORENSIC SCIENCE IN CRIMINAL COURTS

On September 20, 2016, PCAST released its unanimous report to the President entitled "*Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods*." This new document, approved by PCAST on January 6, 2017, is an addendum to the earlier report developed to address input received from stakeholders in the intervening period.

Background

PCAST's 2016 report addressed the question of when expert testimony based on a forensic feature-comparison method should be deemed admissible in criminal courts.[1] We briefly summarize key aspects of the previous report.

Forensic feature-comparison methods

PCAST chose to focus solely on forensic feature-comparison methods. These methods seek to determine whether a questioned sample is likely to have come from a known source based on shared features in certain types of evidence. Specific methods are defined by such elements as:

(i)   the type of evidence examined (e.g., DNA, fingerprints, striations on bullets, bitemarks, footwear impressions, head-hair);

(ii) the complexity of the sample examined (e.g., a DNA sample from a single person vs. a three-person mixture in which a person of interest may have contributed only 1%); and

(iii) whether the conclusion concerns only "class characteristics" or "individual characteristics" (e.g., whether a shoeprint was made by a pair of size 12 Adidas Supernova Classic running shoes vs. whether it was made by a *specific* pair of such running shoes).

The U.S. legal system recognizes that scientific methods can assist the quest for justice, by revealing information and allowing inferences that lie beyond the experience of ordinary observers. But, precisely because the conclusions are potentially so powerful and persuasive, the law requires scientific testimony be based on methods that are scientifically valid and reliable.[2]

Requirement for empirical testing of subjective methods

In its report, PCAST noted that the *only* way to establish the scientific validity and degree of reliability of a *subjective* forensic feature-comparison method—that is, one involving significant human judgment—is to test it *empirically* by seeing how often examiners actually get the right answer. Such an empirical test of a subjective forensic feature-comparison method is referred to as a "black-box test." The point reflects a central tenet underlying all science: *an empirical claim cannot be considered scientifically valid until it has been empirically tested*.

If practitioners of a subjective forensic feature-comparison method claim that, through a procedure involving substantial human judgment, they can determine with reasonable accuracy whether a particular type of evidence came from a particular source (e.g., a specific type of pistol or a specific pistol), the claim cannot be considered scientifically valid and reliable until one has tested it by (i) providing an adequate number of examiners with an adequate number of test problems that resemble those found in forensic practice and (ii) determining whether they get the right answer with acceptable

---

[1] As noted in the report, PCAST did not address the use of forensic methods in criminal *investigations*, as opposed to in criminal prosecution in courts.
[2] See discussion of the Federal Rules of Evidence in Chapter 3 of PCAST's report.

frequency for the intended application.[3] While scientists may debate the precise design of a study, there is no room for debate about the absolute requirement for empirical testing.

Importantly, the test problems used in the empirical study define the specific bounds within which the validity and reliability of the method has been established (e.g., is a DNA analysis method reliable for identifying a sample that comprises only 1% of a complex mixture?).

<u>Evaluation of empirical testing for various methods</u>

To evaluate the empirical evidence supporting various feature-comparison methods, PCAST invited broad input from the forensic community and conducted its own extensive review. Based on this review, PCAST evaluated seven forensic feature-comparison methods to determine whether there was appropriate empirical evidence that the method met the threshold requirements of "scientific validity" and "reliability" under the Federal Rules of Evidence.

- In two cases (DNA analysis of single-source samples and simple mixtures; latent fingerprint analysis), PCAST found that there was clear empirical evidence.
- In three cases (bitemark analysis; footwear analysis; and microscopic hair comparison), PCAST found *no empirical studies whatsoever* that supported the scientific validity and reliability of the methods.
- In one case (firearms analysis), PCAST found only one empirical study that had been appropriately designed to evaluate the validity and estimate the reliability of the ability of firearms analysts to associate a piece of ammunition with a specific gun. Because scientific conclusions should be shown to be reproducible, we judged that firearms analysis currently falls short of the scientific criteria for scientific validity.
- In the remaining case (DNA analysis of complex mixtures), PCAST found that empirical studies had evaluated validity within a limited range of sample types.

Responses to the PCAST Report

Following the report's release, PCAST received input from stakeholders, expressing a wide range of opinions. Some of the commentators raised the question of whether empirical evidence is truly needed to establish the validity and degree of reliability of a forensic feature-comparison method.

The Federal Bureau of Investigation (FBI), which clearly recognizes the need for empirical evidence and has been a leader in performing empirical studies in latent-print examination, raised a different issue. Specifically, although PCAST had received detailed input on forensic methods from forensic scientists at the FBI Laboratory, the agency suggested that PCAST may have failed to take account of some relevant empirical studies. A statement issued by the Department of Justice (DOJ) on September 20, 2016 (the same day as the report's release) opined that:

> The report does not mention numerous published research studies which seem to meet PCAST's criteria for appropriately designed studies providing support for foundational validity. That omission discredits the PCAST report as a thorough evaluation of scientific validity.

Given its respect for the FBI, PCAST undertook a further review of the scientific literature and invited a variety of stakeholders—including the DOJ—to identify any "published . . . appropriately designed

---

[3] The size of the study (e.g., number of examiners and problems) affects the strength of conclusions that can be drawn (e.g., the upper bound on the error rate). The acceptable level of error rate depends on context.

studies" that had not been considered by PCAST and that established the validity and reliability of any of the forensic feature-comparison methods that the PCAST report found to lack such support. As noted below, DOJ ultimately concluded that it had no additional studies for PCAST to consider.

PCAST received written responses from 26 parties, including from Federal agencies, forensic-science and law-enforcement organizations, individual forensic-science practitioners, a testing service provider, and others in the US and abroad.[4] Many of the responses are extensive, detailed and thoughtful, and they cover a wide range of topics; they provide valuable contributions for advancing the field. PCAST also held several in-person and telephonic meetings with individuals involved in forensic science and law enforcement. In addition, PCAST reviewed published statements from more than a dozen forensic-science, law-enforcement and other entities.[5] PCAST is deeply grateful to all who took the time and effort to opine on this important topic.

In what follows, we focus on three key issues raised.

Issue: Are empirical studies truly necessary?

While forensic-science organizations agreed with the value of empirical tests of subjective forensic feature-comparison methods (that is, black-box tests), many suggested that the validity and reliability of such a method could be established *without* actually empirically testing the method in an appropriate setting. Notably, however, none of these respondents identified any *alternative* approach that could establish the validity and reliability of a subjective forensic feature-comparison method.

PCAST is grateful to these organizations because their thoughtful replies highlight the fundamental issue facing the forensic sciences: *the role of empirical evidence*. As noted in PCAST's report, forensic scientists rightly point to several elements that provide critical foundations for their disciplines. However, there remains confusion as to whether these elements can suffice to establish the validity and degree of reliability of particular methods.

(i) The forensic-science literature contains many papers describing variation among features. In some cases, the papers argue that patterns are "unique" (e.g., that no two fingerprints, shoes or DNA patterns are identical if one looks carefully enough). Such studies can provide a valuable *starting point* for a discipline, because they suggest that it may be worthwhile to attempt to develop reliable methods to identify the source of a sample based on feature comparison. However, such studies—no matter how extensive—can *never* establish the validity or degree of reliability of any particular method. Only empirical testing can do so.

(ii) Forensic scientists rightly cite examiners' experience and judgment as important elements in their disciplines. PCAST has great respect for the value of examiners' experience and judgment: they are critical factors in ensuring that a scientifically valid and reliable method is practiced correctly. However, experience and judgment alone—no matter how great—can *never* establish the validity or degree of reliability of any particular method. Only empirical testing of the method can do so.[6]

---

[4] www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensics_2016_additional_responses.pdf.

[5] www.whitehouse.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensics_2016_public_comments.pdf.

[6] Some respondents, such as the Organization of Scientific Area Committees' Friction Ridge Subcommittee, suggested that forensic science should be considered as analogous to medicine, in which physicians often treat patients on the basis of experience and judgment even in the absence of established empirical evidence. However, the analogy is inapt. Physicians act with a patient's consent for the patient's benefit. There is no legal requirement, analogous to the requirement imposed upon expert testimony in court by the Federal Rules of Evidence, that physician's actions be based on "reliable principles and methods." Physicians may rely on hunches; experts testifying in court about forensic feature-comparison methods may not.

1070

7340d2d7-67ae-4b31-9c9d-0419dd510c7a   20220314-17090

(iii) Forensic scientists cite the role of professional organizations, certification, accreditation, best-practices manuals, and training within their disciplines. PCAST recognizes that such practices play a critical role in any professional discipline. However, the existence of good professional practices alone—no matter how well crafted—can *never* establish the validity or degree of reliability of any particular method. Only empirical testing of the method can do so.

PCAST does not diminish in any way the important roles of prior research and other types of activities within forensic science and practice. Moreover, PCAST expresses great respect for the efforts of forensic practitioners, most of whom are devoted public servants. It is important to emphasize, however, contrary to views expressed by some respondents, that there is no "hierarchy" in which empirical evidence is simply the <u>best</u> way to establish validity and degree of reliability of a subjective feature-comparison method. In science, empirical testing is the <u>only</u> way to establish the validity and degree of reliability of such an empirical method.

Fortunately, empirical testing of empirical methods is feasible. There is no justification for accepting that a method is valid and reliable in the absence of appropriate empirical evidence.

Issue: Importance of other kinds of studies

In its response to PCAST's call for further input, the Organization of Scientific Area Committees' Friction Ridge Subcommittee (OSAC FRS), whose purview includes latent-print analysis, raised a very important issue:

> While the OSAC FRS agrees with the need for black box studies to evaluate the overall validity of a particular method, the OSAC FRS is concerned this view could unintentionally stifle future research agendas aimed at dissecting the components of the black box in order to transition it from a subjective method to an objective method. If the PCAST maintains such an emphasis on black box studies as the *only* means of establishing validity, the forensic science community could be inundated with predominantly black box testing and potentially detract from progress in refining other foundational aspects of the method, such as those previously outlined by the OSAC FRS, in an effort to identify ways to emphasize objective methods over subjective methods (see www.nist.gov/topics/forensic-science/osac-research-development-needs.) Given the existing funding limitations, this will be especially problematic and the OSAC RFS is concerned other foundational research will thus be left incomplete.

PCAST applauds the work of the friction-ridge discipline, which has set an excellent example by undertaking both (i) path-breaking black-box studies to establish the validity and degree of reliability of latent-fingerprint analysis, and (ii) insightful "white-box" studies that shed light on how latent-print analysts carry out their examinations, including forthrightly identifying problems and needs for improvement. PCAST also applauds ongoing efforts to transform latent-print analysis from a subjective method to a fully objective method. In the long run, the development of objective methods is likely to increase the power, efficiency and accuracy of methods—and thus better serve the public.

In the case of subjective methods whose validity and degree of reliability have <u>already been established</u> by appropriate empirical studies (such as latent-print analysis), PCAST agrees that continued investment in black-box studies is likely to be less valuable than investments to develop fully objective methods. Indeed, PCAST's report calls for substantial investment in such efforts.

The situation is different, however, for subjective methods whose validity and degree of reliability has not been established by appropriate empirical studies. If a discipline wishes to offer testimony based on a subjective method, it must first establish the method's validity and degree of reliability—which can only be done through empirical studies. However, as the OSAC FRS rightly notes, a discipline could follow an alternative path by abandoning testimony based on the subjective method and instead developing an objective method. Establishing the validity and degree of reliability of an objective method is often more straightforward. PCAST agrees that, in many cases, the latter path will make more sense.

Issue: Completeness of PCAST's evaluation

Finally, we considered the important question, raised by the DOJ in September, of whether PCAST had failed to consider "numerous published research studies which seem to meet PCAST's criteria for appropriately designed studies providing support for foundational validity."

PCAST re-examined the five methods evaluated in its report for which the validity and degree of reliability had not been fully established. We considered the more than 400 papers cited by the 26 respondents; the vast majority had already been reviewed by PCAST in the course of the previous study. At the suggestion of John Butler of the National Institute of Standards and Technology (NIST), we also consulted INTERPOL's extensive summary of the forensic literature to identify additional potentially relevant papers.[7] Although our inquiry was undertaken in response to the DOJ's concern, DOJ informed PCAST in late December that it had no additional studies for PCAST to consider.

Bitemark analysis

In its report, PCAST stated that it found no empirical studies whatsoever that establish the scientific validity or degree of reliability of bitemark analysis as currently practiced. To the contrary, it found considerable literature pointing to the unreliability of the method. None of the respondents identified any empirical studies that establish the validity or reliability of bitemark analysis. (One respondent noted a paper, which had already been reviewed by PCAST, that studied whether examiners agree when measuring features in dental casts but did not study bitemarks.) One respondent shared a recent paper by a distinguished group of biomedical scientists, forensic scientists, statisticians, pathologists, medical examiners, lawyers, and others, published in November 2016, that is highly critical of bitemark analysis and is consistent with PCAST's analysis.

Footwear analysis

In its report, PCAST considered feature-comparison methods for associating a shoeprint with a specific shoe based on randomly acquired characteristics (as opposed to with a class of shoes based on class characteristics). PCAST found no empirical studies whatsoever that establish the scientific validity or reliability of the method.

The President of the International Association for Identification (IAI), Harold Ruslander, responded to PCAST's request for further input. He kindly organized a very helpful telephonic meeting with IAI member Lesley Hammer. (Hammer has conducted some of the leading research in the field—including a 2013 paper, cited by PCAST, that studied whether footwear examiners reach similar conclusions when they are presented with evidence in which the identifying features have already been identified.)

---

[7] The INTERPOL summaries list 4232 papers from 2010-2013 and 4891 papers from 2013-2016, sorted by discipline, see www.interpol.int/INTERPOL-expertise/Forensics/Forensic-Symposium.

Hammer confirmed that no empirical studies have been published to date that test the ability of examiners to reach correct conclusions about the source of shoeprints based on randomly acquired characteristics. Encouragingly, however, she noted that the first such empirical study is currently being undertaken at the West Virginia University. When completed and published, this study should provide the first actual empirical evidence concerning the validity of footwear examination. The types of samples and comparisons used in the study will define the bounds within which the method can be considered reliable.

Microscopic hair comparison

In its report, PCAST considered only those studies on microscopic hair comparison cited in a recent DOJ document as establishing the scientific validity and reliability of the method. PCAST found that none of these studies provided any meaningful evidence to establish the validity and degree of reliability of hair comparison as a forensic feature-comparison method. Moreover, a 2002 FBI study, by Houck and Budowle, showed that hair analysis had a stunningly high error rate in practice: Of hair samples that FBI examiners had found in the course of actual casework to be microscopically indistinguishable, 11% were found by subsequent DNA analysis to have come from different individuals.

PCAST received detailed responses from the Organization of Scientific Area Committees' Materials Subcommittee (OSAC MS) and from Sandra Koch, Fellow of the American Board of Criminalistics (Hairs and Fibers). These respondents urged PCAST not to underestimate the rich tradition of microscopic hair analysis. They emphasized that anthropologists have published many papers over the past century noting differences in average characteristics of hair among different ancestry groups, as well as variation among individuals. The studies also note intra-individual differences among hair from different sites on the head and across age.

While PCAST agrees that these empirical studies describing hair differences provide an encouraging starting point, we note that the studies do not address the validity and degree of reliability of hair comparison as a forensic feature-comparison method. What is needed are empirical studies to assess how often examiners incorrectly associate similar but distinct-source hairs (i.e., false-positive rate). Relevant to this issue, OSAC MS states: "Although we readily acknowledge that an error rate for microscopic hair comparison is not currently known, this should not be interpreted to suggest that the discipline is any less scientific." In fact, this is the central issue: the acknowledged lack of any empirical evidence about false-positive rates indeed means that, as a *forensic feature-comparison method*, hair comparison lacks a scientific foundation.

Based on these responses and its own further review of the literature beyond the studies mentioned in the DOJ document, PCAST concludes that there are no empirical studies that establish the scientific validity and estimate the reliability of hair comparison as a forensic feature-comparison method.

Firearms analysis

In its report, PCAST reviewed a substantial set of empirical studies that have been published over the past 15 years and discussed a representative subset in detail. We focused on the ability to associate ammunition not with a class of guns, but with a specific gun within the class.

The firearms discipline clearly recognizes the importance of empirical studies. However, most of these studies used flawed designs. As described in the PCAST report, "set-based" approaches can inflate examiners' performance by allowing them to take advantage of internal dependencies in the data. The

most extreme example is the "closed-set design", in which the correct source of each questioned sample is always present; studies using the closed-set design have underestimated the false-positive and inconclusive rates by more than 100-fold. This striking discrepancy seriously undermines the validity of the results and underscores the need to test methods under appropriate conditions. Other set-based designs also involve internal dependencies that provide hints to examiners, although not to the same extent as closed-set designs.

To date, there has been only one appropriately designed black-box study: a 2014 study commissioned by the Defense Forensic Science Center (DFSC) and conducted by the Ames Laboratory, which reported an upper 95% confidence bound on the false-positive rate of 2.2%.[8]

Several respondents wrote to PCAST concerning firearms analysis. None cited additional appropriately designed black-box studies similar to the recent Ames Laboratory study. Stephen Bunch, a pioneer in empirical studies of firearms analysis, provided a thoughtful and detailed response. He agreed that set-based designs are problematic due to internal dependencies, yet suggested that certain set-based studies could still shed light on the method if properly analyzed. He focused on a 2003 study that he had co-authored, which used a set-based design and tested a small number of examiners (n=8) from the FBI Laboratory's Firearms and Toolmarks Unit.[9] Although the underlying data are not readily available, Bunch offered an estimate of the number of truly independent comparisons in the study and concluded that the 95% upper confidence bound on the false-positive rate in his study was 4.3% (vs. 2.2% for the Ames Laboratory black-box study).

The Organization of Scientific Area Committee's Firearms and Toolmarks Subcommittee (OSAC FTS) took the more extreme position that all set-based designs are appropriate and that they reflect actual casework, because examiners often start their examinations by sorting sets of ammunition from a crime-scene. OSAC FTS's argument is unconvincing because (i) it fails to recognize that the results from certain set-based designs are wildly inconsistent with those from appropriately designed black-box studies, and (ii) the key conclusions presented in court do not concern the ability to sort collections of ammunition (as tested by set-based designs) but rather the ability to accurately associate ammunition with a specific gun (as tested by appropriately designed black-box studies).

Courts deciding on the admissibility of firearms analysis should consider the following scientific issues:
(i)   There is only a single appropriate black-box study, employing a design that cannot provide hints to examiners. The upper confidence bound on the false-positive rate is equivalent to an error rate of 1 in 46.
(ii) A number of older studies involve the seriously flawed closed-set design, which has dramatically underestimated the error rates. These studies do not provide useful information about the actual reliability of firearms analysis.
(iii) There are several studies involving other kinds of set-based designs. These designs also involve internal dependencies that can provide hints to examiners, although not to the same extent that closed-set designs do. The large Miami-Dade study cited in the PCAST report and the small studies cited by Bunch fall into this category; these two studies have upper confidence bounds corresponding to error rates in the range of 1 in 20.

From a scientific standpoint, scientific validity should require at least two properly designed studies to ensure reproducibility. The issue for judges is whether one properly designed study, together with

---

[8] PCAST also noted that some studies combine tests of both class characteristics and individual characteristics, but fail to distinguish between the results for these two very different questions.
[9] PCAST did not select the paper for discussion in the report owing to its small size and set-based design, although it lists it.

ancillary evidence from imperfect studies, adequately satisfies the legal criteria for scientific validity. Whatever courts decide, it is essential that information about error rates is properly reported.

DNA analysis of complex mixtures

In its report, PCAST reviewed recent efforts to extend DNA analysis to samples containing complex mixtures. The challenge is that the DNA profiles resulting from such samples contain many alleles (depending on the number of contributors) that vary in height (depending on the ratios of the contributions), often overlap fully or partially (due to their "stutter patterns"), and may sometimes be missing (due to PCR dropout). Early efforts to interpret these profiles involved purely subjective and poorly defined methods, which were not subjected to empirical validation. Efforts then shifted to a quantitative method called combined probability of inclusion (CPI); however, this approach also proved seriously problematic.[10]

Recently, efforts have focused on an approach called probabilistic genotyping (PG), which uses mathematical models (involving a likelihood-ratio approach) and simulations to attempt to infer the likelihood that a given individual's DNA is present in the sample. PCAST found that empirical testing of PG had largely been limited to a narrow range of parameters (number and ratios of contributors). We judged that the available literature supported the validity and reliability of PG for samples with three contributors where the person of interest comprises at least 20% of the sample. Beyond this approximate range (i.e. with a larger number of contributors or where the person if interest makes a lower than 20% contribution to the sample), however, there has been little empirical validation.[11]

A recent controversy has highlighted issues with PG. In a prominent murder case in upstate New York, a judge ruled in late August (a few days before the approval of PCAST's report) that testimony based on PG was inadmissible owing to insufficient validity testing.[12] Two PG software packages (STRMix and TrueAllele), from two competing firms, reached differing[13] conclusions about whether a DNA sample in the case contained a tiny contribution (~1%) from the defendant. Disagreements between the firms have grown following the conclusion of the case.

PCAST convened a meeting with the developers of the two programs (John Buckleton and Mark Perlin), as well as John Butler from NIST, to discuss how best to establish the range in which a PG software program can be considered to be valid and reliable. Buckleton agreed that empirical testing of PG software with different kinds of mixtures was necessary and appropriate, whereas Perlin contended that empirical testing was unnecessary because it was mathematically impossible for the likelihood-ratio approach in his software to incorrectly implicate an individual. PCAST was unpersuaded by the latter argument. While likelihood ratios are a mathematically sound concept, their application requires

---

[10] Just as the PCAST report was completed, a paper was published that proposed various rules for the use of CPI. See Bieber, F.R., Buckleton, J.S., Budowle, B., Butler, J.M., and M.D. Coble. "Evaluation of forensic DNA mixture evidence: protocol for evaluation, interpretation, and statistical calculations using the combined probability of inclusion." *BMC Genetics*. bmcgenet.biomedcentral.com/articles/10.1186/s12863-016-0429-7. While PCAST agreed that these rules are *necessary*, PCAST did not review whether these rules were sufficient to ensure reliability and took no position on this question.

[11] The few studies that have explored 4- or 5-person mixtures often involve mixtures that are derived from only a few sets of people (in some cases, only one). Because the nature of overlap among alleles is a key issue, it is critical to examine mixtures from various different sets of people. In addition, the studies involve few mixtures in which a sample is present at an extremely low ratio. By expanding these empirical studies, it should be possible to test validity and reliability across a broader range.

[12] See McKinley, J. "Judge Rejects DNA Test in Trial Over Garrett Phillips's Murder." New York Times, August 26, 2016, www.nytimes.com/2016/08/27/nyregion/judge-rejects-dna-test-in-trial-over-garrett-phillipss-murder.html. The defendant was subsequently acquitted.

[13] Document updated on January 17, 2017.

making a set of assumptions about DNA profiles that require empirical testing.[14] Errors in the assumptions can lead to errors in the results. To establish validity with a range of parameters, it is thus important to undertake empirical testing with a *variety* of samples in the relevant range.[15]

PCAST received thoughtful input from several respondents. Notably, one response[16] suggested that the relevant category for consideration should be expanded from "complex mixtures" (defined based on the number of contributors) to "complex samples" (defined to include also samples with low amounts of template, substantial degradation, or significant PCR inhibition, all of which will also complicate interpretation). We agree that this expansion could be useful.

The path forward is straightforward. The validity of specific PG software should be validated by testing a diverse collection of samples within well-defined ranges. The DNA analysis field contains excellent scientists who are capable of defining, executing, and analyzing such empirical studies.

When considering the admissibility of testimony about complex mixtures (or complex samples), judges should ascertain whether the published validation studies adequately address the nature of the sample being analyzed (e.g., DNA quantity and quality, number of contributors, and mixture proportion for the person of interest).

Conclusion

Forensic science is at a crossroads. There is growing recognition that the law requires that a forensic feature-comparison method be established as scientifically valid and reliable before it may be used in court and that this requirement can only be satisfied by actual empirical testing. Several forensic disciplines, such as latent-print analysis, have clearly demonstrated that actual empirical testing is feasible and can help drive improvement. A generation of forensic scientists appears ready and eager to embrace a new, empirical approach—including black-box studies, white-box studies, and technology development efforts to transform subjective methods into objective methods.

PCAST urges the forensic science community to build on its current forward momentum. PCAST is encouraged that NIST has already developed an approach, subject to availability of budget, for carrying out the functions proposed for that agency in our September report.

In addition, progress would be advanced by the creation of a cross-cutting Forensic Science Study Group—involving leading forensic and non-forensic scientists in equal measure and spanning a range of feature-comparison disciplines—to serve as a scientific forum to discuss, formulate and invite broad input on (i) empirical studies of validity and reliability and (ii) approaches for new technology development, including transforming subjective methods into objective methods. Such a forum would complement existing efforts focused on developing best practices and informing standards and might strengthen connections between forensic disciplines and other areas of science and technology. It might be organized by scientists in cooperation with one or more forensic and non-forensic science organizations—such as DFSC, NIST, IAI, and the American Association for the Advancement of Science.

---

[14] Butler noted that one must make assumptions, for each locus, about the precise nature of reverse and forward stutter and about the probability of allelic dropout.

[15] Butler noted that it is important to consider samples with different extents of allelic overlap among the contributors.

[16] This response was provided by Keith Inman, Norah Rudin and Kirk Lohmueller.

★ ★ ★

GOVERNMENT OF THE DISTRICT OF COLUMBIA
METROPOLITAN POLICE DEPARTMENT

March 26, 2007

TO: Commander   FES #91-01407/TS
     Office of the Superintendent of Detectives   FES #96-01556/TS
                                                     CSIB #91-05961
                                                     CCN #202-465

ATTENTION: Fifth District
            Det. Eddie Voysest
            Violent Crimes Branch

RE: K.E. Simmons
     C.K. Pixley
     Samantha Gillard
     Decedents – Homicide

TYPE OF EXAM: Firearms

The following items of evidence were personally delivered on March 16, 2007 by Technician Alfred P. Holmes.

Item	#5	one (1) bullet
Items	#6-#9	four (4) cartridge cases
Item	#11	one (1) lead fragment and one (1) jacket fragment
Item	#14	one (1) bullet
Items	#15-#23	nine (9) cartridge cases
Item	#48	one (1) bullet
Item	#49	one (1) bullet fragment
Items	#50-#51	two (2) bullets
Items	#58-#61	four (4) bullets
Item	#62	one (1) bullet fragment
Items	#69A-#69B	three (3) bullet jackets
Item	#69D	two (2) lead bullet cores, nine (9) lead fragments and six (6) jacket fragments

Note: Items #69A- #69D were designated by the writer.

The following items of evidence were personally delivered on March 20, 2007 by Detective Eddie Voysest and processed under FES# 96-01556/TS.

Item #1 one (1) pistol with magazine

RESULTS OF EXAMINATION:

Item #1 is a Glock, caliber 9mm Luger semiautomatic pistol, model 17, serial number MP2341DC. This firearm is the property of the Metropolitan Police Department and is issued to Officer Duane Fowler. Item #1 was found to be in normal operating condition. All safety design mechanisms are intact and functional.

Items #5, #14, #51, #58, #60, #61, #69A, #69B and #69C are three (3) caliber 9mm Luger copper jacketed bullets and six (6) nickel jacketed bullets which were identified as having been fired from the same barrel rifled with six (6) grooves, right twist. Due to differences in rifling characteristics, Items #5, #14, #51, #58, #60, #61, #69A, #69B and #69C could not have been fired from Item #1 pistol. Among the firearms, which may produce similar rifling impressions are caliber 9mm Luger pistols marketed by Astra, Beretta, Hi-Point, Intratec, Jennings and Taurus.

Items #6 through #9 and #15 through #23 are thirteen (13) caliber 9mm Luger cartridge cases, Winchester and Remington brands, which were identified as having been fired in the same firearm. Due to differences in rifling characteristics, Items #6 through #9 and #15 through #23 could not have been fired from Item #1 pistol.

Items #48, #50, #59 and #70 are four (4) caliber .22 lead bullets which were identified as having been fired from the same barrel rifled with eight (8) grooves, right twist. Among the firearms, which may produce similar rifling impressions are caliber .22 revolvers and pistols marketed by Arminius, Cooey Arms Lorcin, Mossberg, RG, Rohm, Tanfoglio and Walther.

Items #11, #49, #62 and #69D consist of ten (10) lead fragments, seven (7) nickel jacket fragments and two (2) copper bullet fragments which have no identifiable characteristics.

Item #69D consists of two (2) lead bullet cores, which is consistent with having separated from its jacket.

Please arrange to have a member of your unit pick up the evidence which is being held in the Firearms Examination Section. In the event a suspect firearm is recovered, please re-submit evidence.

(b)(6) per EOUSA

Travis Spinder
Firearms and Toolmark Examiner

CURRICULUM VITAE

TRAVIS Y. SPINDER
Forensic Science Supervisor-Firearm & Toolmark Section

Business Address: (b)(6) per EOUSA
Business Phone: (b)(6) per EOUSA
Desk: (b)(6) per EOUSA
Fax: (406) 549-1067
Born: (b)(6) per EOUSA

EDUCATION

B.A. Sociology/Criminology – University of Montana, Missoula, MT 1997

CERTIFICATIONS

Association of Firearm and Tool Mark Examiners (AFTE) Certifications
     -Firearm Evidence Examination and Identification (December 19, 2012)

CURRENT FIELD OF ACTIVITY

Forensic Science Supervisor-Firearm & Toolmark Section, Montana Department of Justice,
Division of Forensic Science, Missoula, MT – September 2007 to Present

PREVIOUS EXPERIENCE

Forensic Firearm and Toolmark Examiner, Montana Department of Justice, Division of Forensic
Science, Missoula, MT – May 2002 to September 2007

Forensic Firearm and Toolmark Contractor, Southwestern Institute of Forensic Sciences,
Criminal Investigation Laboratory, Dallas, TX – May 2005 to September 2009

Forensic Firearm and Toolmark Contractor, Metropolitan Police Department, Firearm
Examination Section, Washington D.C. – June 2, 2003 to August 26, 2003 & March 26, 2007 to
September 24, 2007

Forensic Firearm and Toolmark Examiner, Southwestern Institute of Forensic Sciences, Criminal
Investigation Laboratory, Dallas, TX – August 1998 to May 2002

Updated – June 6, 2016
1079

Curriculum Vitae   Travis Y. Spinder

PROFESSIONAL ORGANIZATION/AWARDS

Association of Firearm and Tool Mark Examiners (Distinguished Member 2002)

Scientific Working Group for Firearms and Toolmarks - SWGGUN (November 2006 – November 2012)

National Shooting Sports Foundation - Shot Show - Safety Advisor (February 2008 – Present)

American Society of Crime Laboratory Directors/Laboratory Accreditation Board - Firearm/Toolmarks Proficiency Review Committee (June 2008 – September 2013)

Association of Firearm and Tool Mark Examiners - Bylaws Committee (June 2008 – Present)

Association of Firearm and Tool Mark Examiners - Board of Admissions Committee (September 2009 – June 2011)

Association of Firearm and Tool Mark Examiners - Board of Directors (June 2012 – Present)

Virginia Department of Forensic Science - Scientific Advisory Committee (October 2013 – Present)

Montana Department of Justice Forensic Science Division – Employee of the Year 2015

Association of Firearm and Tool Mark Examiners – President (June 2016 – Present)

TECHNICAL/SPECIALIZED TRAINING

Resident training course in the field of Firearm and Toolmark Examination, Montana Department of Justice, Division of Forensic Science, Missoula, MT – May 1997- July 1998

Passed Competency Testing at the Southwestern Institute of Forensic Sciences in Firearm and Toolmark Examination, Dallas, TX – August 1998

Beretta Armorers School offered by Beretta, Tampa, FL – July 1998

NIBIN/Drugfire Training Course, Rosslyn, VA – March 1999

Heckler & Koch Armorers School offered by Heckler & Koch, St. Louis, MO – June 2000

NIBIN/IBIS Training Course, Largo, FL – June 2001

Bureau of Alcohol, Tobacco, Firearms and Explosives (ATF) Serial Number Restoration Course,

2

Curriculum Vitae   Travis Y. Spinder
Dallas, TX – August 2001

Ruger Armorers School offered by Ruger, Denton, TX – November 2001

Smith & Wesson "SW99" Armorers School offered by Smith & Wesson, Coeur d'Alene, ID – October 2002

Colt "Rifle, Carbine & SMG" Armorers School offered by Colt, Missoula, MT – March 2004

"Trends in Ammunition" Northwest Association of Forensic Scientists, Missoula, MT – April 2004

"Shooting Scene Reconstruction" Northwest Association of Forensic Scientists, Missoula, MT – April 2004

"ISO Standards and Firearm and Toolmarks" offered by ASCLD/LAB at AFTE 2007, San Francisco, CA – May 2007

Federal Bureau of Investigation (FBI) Gunpowder and Gunshot Residue School, Spokane, WA – August 2008

"Trajectory Measurement/Documentation" offered by Michael Haag at AFTE 2010, Henderson, NV – May 2010

PROFESSIONAL TRAINING CONFERENCES

29th Annual Training Conference, Association of Firearm and Tool Mark Examiners, Tampa, FL – July 1998

31st Annual Training Conference, Association of Firearm and Tool Mark Examiners, St. Louis, MO – June 2000

Northwest Association of Forensic Scientists, Missoula, MT – April, 2004

35th Annual Training Conference, Association of Firearm and Tool Mark Examiners, Vancouver, BC, Canada – May 2004

37th Annual Training Conference, Association of Firearm and Tool Mark Examiners, Springfield, MA – June 2006

38th Annual Training Conference, Association of Firearm and Tool Mark Examiners, San Francisco, CA – May 2007

Updated – June 6, 2016
1081

7340d2d7-67ae-4b31-9c9d-0419dd510c7a   20220314-17101

Curriculum Vitae   Travis Y. Spinder

39th Annual Training Conference, Association of Firearm and Tool Mark Examiners, Honolulu, HI – May 2008

40th Annual Training Conference, Association of Firearm and Tool Mark Examiners, Miami, FL – June 2009

41st Annual Training Conference, Association of Firearm and Tool Mark Examiners, Henderson, NV – May 2010

42nd Annual Training Conference, Association of Firearm and Tool Mark Examiners, Chicago, IL – June 2011

43nd Annual Training Conference, Association of Firearm and Tool Mark Examiners, Buffalo, NY – June 2012

44th Annual Training Conference, Association of Firearm and Tool Mark Examiners, Albuquerque, NM – June 2013

45th Annual Training Conference, Association of Firearm and Tool Mark Examiners, Seattle, WA – May 2014

46th Annual Training Conference, Association of Firearm and Tool Mark Examiners, Dallas, TX – May 2015

FIREARM/AMMUNITION FACTORY TOURS

The Hunting Shack (Ammunition), Stevensville, MT – April 1998

Cooper Firearms, Stevensville, TX – April 1998

Blount Inc. (CCI & Speer Ammunition), Lewiston, ID – May 1998

Hi-Point Firearms, Mansfield, OH – December 2000

Shilen Barrel, Ennis, TX – March 2001

Outback (Outback Shooting Range - Custom Ammunition), Cumby, TX – March 2001

Smith & Wesson Firearms, Springfield, MA – June 2006

Savage Arms, Springfield, MA – June 2006

MasterPiece Arms, Carrollton, GA – November 2007

Curriculum Vitae   Travis Y. Spinder

Advanced Armament (Silencers), Norcross, GA – November 2007

Glock Firearms, Smyrna, GA – November 2007

Olympic Arms, Olympia, WA – April 2009

Rainier Ballistics, Tacoma, WA – April 2009

DCA Inc., Barrington, IL – November 2009

Klein Tools, Lincolnshire, IL & Skokie, IL – November 2009

Red Jacket Firearms, Baton Rouge, LA – April 2010

Ithaca Gun Company, Upper Sandusky, OH – November 2010

Hi-Point Firearms, Mansfield, OH – November 2010

Bitterroot Valley Ammunition and Components, Stevensville, MT – December 2010

Kel-Tec CNC Industries, Inc (Firearms) , Cocoa, FL – April 2012

Diamondback Firearms, Cocoa, FL – April 2012

FORENSIC LABORATORY TOURS

Idaho State Police Forensic Services Laboratory – Coeur d'Alene, ID

Washington State Patrol Forensic Laboratory Services – Spokane, WA

Florida Department of Law Enforcement Crime Lab – Tampa, FL

U.S. Army Criminal Investigation Laboratory – Forest Park, GA

Ohio Bureau of Criminal Identification and Investigation Crime Lab – London, OH

Ohio Bureau of Criminal Identification and Investigation Crime Lab – Richfield, OH

Columbus Police Department Crime Laboratory – Columbus, OH

Indianapolis - Marion County Forensic Services Agency – Indianapolis, IN

Updated – June 6, 2016
1083

Curriculum Vitae   Travis Y. Spinder

Federal Bureau of Investigation Laboratory Services – Quantico, VA

Georgia Bureau of Investigation – Decatur, GA

U.S. Army Criminal Investigation Laboratory (new facility) – Forest Park, GA

Washington State Patrol Forensic Laboratory Services (new facility) – Cheney, WA

Washington State Patrol Forensic Laboratory Services – Seattle, WA

Oregon State Police Forensic Services Division – Clackamas, OR

Southwestern Institute of Forensic Sciences (new facility) – Dallas, TX

Idaho State Police Forensic Services Laboratory (new facility) – Coeur d'Alene, ID

Louisiana State Police Crime Laboratory – Baton Rouge, LA

Miami Valley Regional Crime Laboratory – Dayton, OH

Virginia Department of Forensic Science Central Laboratory – Richmond, VA

FORENSIC SCIENCE PUBLICATIONS

T.Y. Spinder, "Suppressed Ruger 10/22" AFTE Journal, Volume 33, Number 4, pp. 332.

T.Y. Spinder, S.B. Allen and D.S. Engel, "Comet Tailing" AFTE Journal, Volume 33, Number 4, pp. 336-337.

T.Y. Spinder and S.B. Allen, "Full-Auto Intratec or Not" AFTE Journal, Volume 34, Number 1, pp. 49.

T.Y. Spinder and S.B. Allen, "Specialty Shotgun Ammunition from All Purpose Ammunition" AFTE Journal, Volume 34, Number 1, pp. 53.

FORENSIC SCIENCE PRESENTATIONS

T.Y. Spinder, "Effects of 5,000 Ejector-to-Breechface Strike of a Single Shot Shotgun" presented at the 29th Annual Meeting of the Association of Firearm and Toolmark Examiners, Tampa, FL – July 1998

Curriculum Vitae   Travis Y. Spinder

T.Y. Spinder, "1999 Firearm Proficiency Test Overview" presented at the 31st Annual Meeting of the Association of Firearm and Toolmark Examiners, St. Louis, MO – June 2000

T.Y. Spinder, "1999 Toolmark Proficiency Test Overview" presented at the 31st Annual Meeting of the Association of Firearm and Toolmark Examiners, St. Louis, MO – June 2000

T.Y. Spinder, "1999 Firearm Research Test Overview" presented at the 31st Annual Meeting of the Association of Firearm and Toolmark Examiners, St. Louis, MO – June 2000

TESTIMONY

Called by both Prosecution and Defense to provide Expert Testimony pertaining to Firearm and Toolmark Examinations 167 times in the following courts:

> Criminal District Court (Texas, Wyoming & Montana)
> Juvenile Court (Texas)
> United States Federal District Court (Montana, Texas, Washington, D.C.)

7

STATEMENT OF STEPHEN G. BUNCH

I, Stephen G. Bunch, , state the following to be true to the best of my knowledge:

1. Since 2002, I have served as a Supervisory Physical Scientist (Unit Chief) at the Federal Bureau of Investigation (FBI) in Quantico, Virginia. My principal duties as Unit Chief involve managing the Firearms and Toolmarks Unit.

2. I began my employment with the FBI in 1996. From 1996 to 1999, I assisted Firearms and Toolmark examiners as a Physical Science Technician. In 1999, I was formally qualified by the FBI Laboratory as a Physical Scientist. My principal duties involved examining firearms and toolmarks related evidence, reporting results to contributing agencies, and sometimes testifying to findings in court. I served in that capacity until becoming Unit Chief in 2002.

3. I earned a B.S. in mechanical engineering in 1978 and a M.A. in History in 1988 from University of Missouri. I also earned a Ph.D in history from University of Illinois in 1995.

4. Dating back to 1996, I have received a great deal of specialized training in the area of firearms and toolmark identification. A listing of forensic education, workshops, in-service training, and any other sort of specialized training is set forth in my attached resume (Statement of Qualification).

5. From 2001 to date, I have been a regular member of the Association of Firearm and Tool Mark Examiners (AFTE). I also serve as a member of the Scientific Working Group for Firearms-Toolmarks identification (SWGGUN).

6. A listing of my peer reviewed publications are listed on my attached resume.

7. Over the course of my career, including my training, I have conducted between 600 and 800 comparison examinations of firearms evidence. I have conducted approximately an additional 300 confirmatory firearms examination comparisons.

8. I have been qualified as an expert witness in the area of Firearms and Toolmark Identification on several occasions in federal and local courts in various jurisdictions, including Easton, Pennsylvania; Pittsburgh, Pennsylvania; Hammond, Indiana; White Plains, New York; Topeka, Kansas; Pierre, South Dakota; Grand Forks, North Dakota; Cheyenne, Wyoming; Rapid City, South Dakota; and Williamsport, Pennsylvania.

9. Firearm identification has been a forensic discipline since the 1930s. Firearms identification is a subset of the broader forensic discipline known as tookmark identification. Toolmark examiners are trained to examine the marks left by tools on any variety of surfaces in an attempt to "match" a toolmark to a particular tool that made the mark. Firearms are simply a subset of tools that impart marks on bullets and cartridge cases. Firearm and toolmark identification is based upon two propositions:

1

Proposition #1:   Toolmarks imparted to objects by different tools will rarely if ever display agreement sufficient to lead a qualified examiner to conclude the objects were marked by the same tool. That is, a qualified examiner will rarely if ever commit a false positive error (misidentification).

Proposition #2:   Most manufacturing processes involve the transfer of rapidly changing or random marks onto work pieces such as barrel bores, breech faces, firing pins, screwdriver blades, and the working surfaces of other common tools. This is caused principally by the phenomena of tool wear and chip formation or by electrical/chemical erosion. Microscopic marks on tools may then continue to change from further wear, corrosion, or abuse.

10. In the field of toolmark identification, toolmarks imparted onto bullets and cartridge cases generally are the easiest to identify to a particular tool, i.e., to a particular firearm, because ammunition is cycled through a firearm in a predictable manner. Other tools commonly analyzed in connection with criminal investigations, such as knives, are less easily analyzed because they can impart toolmarks at different angles and with varying degrees of force .

11. A cartridge is made up of four main parts: the bullet, the case, the propellant, and the primer. The case is the covering that holds all of the cartridge components together. The bullet itself is the projectile propelled from the weapon. The propellant rests behind the bullet and very rapidly burns upon ignition. The primer is the component at the rear of the case that starts the reaction when the cartridge is fired.

12. When a gun is fired, the interior of the barrel of the gun imparts "rifling" impressions onto the bullet. The barrel of a gun is manufactured to impart a twist on a bullet as it travels, to ensure firing accuracy. The inside of a gun barrel is imprinted with cuts running the length of the barrel. The cuts within the barrel are called "grooves" and the raised surfaces are called "lands." Those rifling characteristics create marks on the bullet as it travels down the barrel. The raised lands cut into the surface of the bullet. Likewise, the bullet also fills the recessed grooves. The corresponding impressions left on the bullet as it travels through the barrel are depressed "lands impressions" and raised "groove impressions." The twist imparted on a bullet can be either left or right, depending on the direction of the lands and grooves.

13. Before a gun is fired, the base of the cartridge abuts the breech of the gun as the cartridge rests in the chamber. When the gun is fired, the cartridge slams into the breech, thereby leaving "breech face marks." An instant before this, the firing pin strikes the primer at the base of the cartridge, initiating the reaction that causes the bullet to fire. The firing pin contact creates a "firing pin impression" on the primer itself.

14. Examiners are trained to observe 3 types of markings, known also as "characteristics," which are imparted onto bullets and cartridge cases:

7340d2d7-67ae-4b31-9c9d-0419dd510c7a   20220314-17107

1.   Class characteristics;

2.   Subclass characteristics; and

3.   Individual characteristics.

15. Firearm class characteristics imparted to a fired bullet or cartridge case allow an examiner to narrow the class of firearm possibilities to certain types of guns made by certain manufacturers. For a fired bullet, the class characteristics consist of the caliber (diameter) of the bullet, the number of land and groove impressions, the direction of twist of the land and groove impressions, and the width of the land and groove impressions. In the case of a spent cartridge case, the examiner looks primarily for the class characteristic displayed by the firing pin impression on the primer. There are several types of firing pin impressions, including, among others, circular, rectangular, hemispherical, and elliptical.

16. On the other end of the spectrum from class characteristics are individual characteristics. Individual characteristics consist of microscopic, random imperfections in the barrel or firing mechanism created by the manufacturing process, wear, corrosion, or abuse. These unintended characteristics are initially caused by changes in the tool as it makes each barrel on the production line. Individual characteristics typically fall into two categories: (1) striated marks made by movement of the bullet within a gun's barrel (typically appearing as scratches), and (2) impressed marks that are pressed into a surface. A spent bullet usually has striated marks, created as it moves through the barrel of the gun. A spent cartridge case, on the other hand, can have both impressed and striated marks. Prior to firing, the process of feeding the cartridge into the chamber can create striated marks. Once the firearm is fired, impressed marks are created on the cartridge case by the guns's firing pin and breech. With semi-automatic weapons, additional marks can be made as the case is expelled from the gun. A spent cartridge is pulled backwards by the "extractor," which can leave striated marks on the case. Subsequently, the "ejector" kicks the case out of the gun, often leaving an impressed mark.

17. A third type of characteristic straddles the line between class and individual characteristics. These are subclass characteristics. These characteristics can exist within a particular production run in the manufacturing process of a certain brand of firearm. Subclass characteristics can occasionally arise from imperfections in a machine tool that persist during the production of multiple firearm components; from extreme hardness differences between the machine tool and the workpieces; or occasionally from particular manufacturing processes such as casting or molding. They cannot be considered class characteristics because they are not common to all units of a particular make and model of firearm. Nor are they individual characteristics because they persist throughout a period of manufacturing.

18. Qualified examiners are trained to distinguish subclass characteristics from individual characteristics, because a true identification may not be made from subclass characteristics. As I discuss later in this affidavit, because potential issues of subclass characteristics are limited to firearms manufactured in the same part of the manufacturing process, researchers have undertaken validity studies specifically designed to test whether firearms examiners could distinguish spent bullets and spent cartridge casings from consecutively manufactured firearms. In each case, examiners were able to match the bullets

3

and cartridge cases to specific firearms, with either no reported errors, i.e., no instances of "false positives," or with error rates under 1%.

19. Since the inception of firearms and toolmark identification as a forensic discipline, firearms examiners have been using a method known as "pattern matching."

20. According to the theory of firearms identification, a qualified examiner often can determine whether two bullets or two cartridge cases came from the same firearm (inconclusive results are fairly common, however). This can be achieved based on an examiner's training and expertise. A conclusion that two cartridge components have a "common origin" can be reached when the examiner concludes that sufficient similarity exists between the patterns on the components.

21. This theory of firearms identification has been utilized throughout the field of firearms and toolmark identification for decades. In 1992, the Association of Firearms and Toomark Examiners (AFTE) memorialized the theory of identification in an attempt to explain the basis of opinions of common origin in toolmark comparisons. The AFTE theory of Identification states:

1. The theory of identification as it pertains to the comparison of toolmarks enables opinions of common origin to be made when the unique surface contours of two toolmarks are in "sufficient agreement."

2. This "sufficient agreement" is related to the significant duplication of random toolmarks as evidenced by a pattern or combination of patterns of surface contours. Significance is determined by the comparative examination of two or more sets of surface contour patterns comprised of individual peaks, ridges and furrows. Specifically, the relative height or depth, width, curvature and spatial relationship of the individual peaks, ridges and furrows within one set of surface contours are defined and compared to the corresponding features in the second set of surface contours. Agreement is significant when it exceeds the best agreement demonstrated between toolmarks known to have been produced by different tools and is consistent with agreement demonstrated by toolmarks known to have been produced by the same tool. The statement that "sufficient agreement" exists between two toolmarks means that the agreement is of a quantity and quality that the likelihood another tool could have made the mark is so remote as to be considered a practical impossibility.

3. Currently the interpretation of individualized/identification is subjective in nature, founded on scientific principles and based on the examiners training and experience.

22. The concept of "uniqueness" can be misleading. No two sets of toolmarks are identical. Said differently, all toolmarks are different on some level. The marks on the items, however, need not be identical for an examiner to declare a match. There need only be "sufficient agreement" between the marks based on the examiner's training and experience.

23. Pattern matching is done by inspecting bullets or cartridge casings under a split-screen

4

comparison microscope, with typical magnifications of 10X-50X. This instrument has been used in this field of forensic science since the 1930s.

24. There are generally four conclusions that examiners reach when conducting an examination:

1. IDENTIFICATION – meaning the toolmarks have been produced by the same tool;

2. INCONCLUSIVE – meaning the toolmarks may or may not have been produced by the same tool;

3. ELIMINATION – meaning the toolmarks were not produced by the same tool;

4. UNSUITABLE – meaning the evidence is unsuitable for examination.

25. In making an identification, an examiner utilizes sound examination methods by employing the precepts of empirical research or study in the comparison of two toolmarks. Each examiner undergoes standardized technical training that develops cognitive skills to recognize, differentiate, and understand the patterns of marks and their meaning. The method of pattern matching makes it possible for an examiner to make an individual association or identification conclusion.

26. Validation studies have repeatedly demonstrated that consecutively manufactured firearms produce individual toolmarks that can be distinguished from one another and can be matched to a single firearm, to a high degree of reliability. However, there is no way to be absolutely (100%) certain of any identification without comparing a particular set of marks to marks created by every firearm produced since the invention of the modern day firearm. Such an endeavor is impossible. Because an examiner cannot rule out with absolute certainty the highly unlikely event that two different firearms produce indistinguishable individual characteristics, an examiner, if asked, must properly qualify an identification. One way an examiner can qualify his or her identification is to conclude that the match is one of "practical certainty," rather than one of "absolute certainty." Practical certainty means that the determination of identity correlates to features whose frequency (or likelihood) of reoccurrence by another tool is so remote that it can be considered practically impossible. Another way to properly qualify an identification is to state that the examiner has matched a toomark to a particular firearm "to a reasonable degree of scientific certainty." Either qualification communicates the examiner's high level of certainty without overstating the significance of the match.

27. Presumptive validity checks involve examiners who investigate a new manufacturing technique to check for indications of "subclass" marking. That is, to see if a tool imparts marks on objects that persist in highly similar form, and that could possibly result in examiners committing false-positive errors for the reason of this similarity. Most of the time the answer is no. On the infrequent occasions when the answer is yes, the results are published or publicized and examiners are thereby informed to be careful about these circumstances.

28. However, the "gold standard" for testing the scientific validity of examiner claims is by means

5

of comprehensive, "black-box" validity tests involving control examiners as participants. In these tests it is known with absolute certainty where each of the test components came from. Of particular interest to researchers is the rate at which an examination results in a "false positive," meaning a false identification (or false match). Over the past decade, firearm examiners, using the same methods and identification criteria as those in actual casework, have consistently reached correct conclusions based upon the samples before them. Usually the error rate was zero. The only published tests that contained a mis-identification error(s) involved marks produced by tools other than firearms. The Scientific Working Group for Firearms and Toolmarks (SWGGUN) has tracked the most recent studies, which can be summarized as follows:

STUDY	ERROR RATE
Brundage (1998)	0%
Bunch & Murphy (2003)	0%
De France (2003)	0%
Thompson & Wyant (2003)	0.78%
Smith (2005)	0%
Orench (2005)	0%

Copies of the above-referenced studies are attached hereto.

29. Consecutively manufactured firearms are the most likely to produce similar microscopic marks on bullets or cartridge cases (subclass marks), for the reason that machine tool wear is at a minimum in moving from one workpiece to the next. Thus, the possibility of a false-positive conclusion that two bullets came from the same firearm is highest with bullets that were fired from two different but consecutively manufactured firearms. Validity tests using consecutively manufactured specimens, however, have not undermined the basic underpinnings of firearm and toolmark identification. For example, research has revealed that the fine, microscopic marks on bullets from consecutively manufactured barrels are readily distinguishable.

30. Another type of black-box test is a proficiency test. These are quality assurance devices designed to test an examiner's competence, or the competence of a laboratory system, not test directly the validity of a theory or technique. There are many drawbacks to these tests when used for validity and error rate purposes. Some of these are the following: anyone who pays the fee may participate in these tests, including attorneys and examiner-trainees; they are not as blind as gold-standard validity tests; participants' responses are linked to him or her and thus are not anonymous; and returns are not mandatory. Firearms proficiency tests, unsurprisingly, show higher error rates than validity tests, with an overall average in the range of 1% - 3%.

31. Thus, contrary to critics' assertions, subclass marks in practice are by no means a serious problem for firearms and toolmarks examiners. This is partly for the reasons given above; namely, (1) that examiners are always alert to new manufacturing techniques that could possibly produce subclass marks, and publish any positive findings to the community at large in order that practicing examiners can take

6

special care in cautionary situations; (2) examiners are trained to remain alert to potential subclass issues, even when research may be silent on particular circumstances; and (3) by all accounts, subclass marks appear to be rare in actual casework, as they are in validity and proficiency tests '

And this leads to additional relevant observations. Indeed, if subclass marks were a significant problem, then doubtless such problems would materialize in black-box testing, especially for those tests involving consecutively-manufactured items. In the big picture, all types of errors are captured by black-box tests, whether they be comprehensive validity tests or proficiency tests, and whether the errors be from theoretical weaknesses, subclass marks, or from human errors stemming from incompetency, lack of training, or quality assurance mishaps (transposing control and evidence samples, for example). But the record so far is that error rates are not high, and in the best designed tests are very low. Were subclass marks a significant problem, error rates would doubtless be well into the double digits or at least consistently in the high single digits. But they are not.

32. It should be noted that the vast majority of forensic laboratories in the United States and abroad have standard operating procedures (technical protocols) that set forth in detail the proper examination procedures, and that these procedures are highly similar across laboratories. Throughout forensic firearms laboratories, once an identification is made, the industry "best practices" provides for a firearms examiner to document and explain the identification through either a photograph or narrative text, describing the primary areas on which the identification was based. Best practices also provides for identifications, or representative identifications, to be confirmed by at least one other examiner. Proper technical and administrative review further insures that the results of the technique are reliable. It should be noted that the practice of confirming identifications suggests that the error rates for validation and proficiency tests may be higher than for actual casework.

33. A small percentage of the community of forensic firearms and toolmark examiners uses a method involving the observation of "consecutive matching striae" (CMS). In principle, CMS can add some quantification to an examination to support an examiner's conclusion of identification. Under CMS, an examiner looks at the number of consecutive striae that match between the bullets being compared. A "run" of striae is essentially a cluster of matching striae that are adjacent to one another. According to the principles of CMS, correspondence between one six-line run of striae or two three-ling runs are enough to make an identification. CMS applies only to striated marks, not to breechface or other impressed marks. For this reason, CMS is only used on fired bullets and not on cartridge cases.

34. Thus, CMS and pattern matching are not mutually exclusive. In practice rather, CMS is merely an extension of pattern matching. CMS is still a method in development and does not undermine the validity or acceptance of traditional pattern matching. In fact, as demonstrated by the attached SWGGUN Survey Summary, the majority of firearms examiners continue to utilize pattern matching and not CMS methods. Also, the vast majority of firearms examiners who use CMS do so in conjunction with, or in addition to, traditional pattern matching.

35. Unlike the small ridges on fingers, a tool will change over time from wear and thus leave different marks on, for example, bullets. In bullets fired through a barrel in sequential fashion, bullet #1

7

may or may not display significant microscopic correspondence to bullet #2000. But this in no sense diminishes the reliability of examiner conclusions or the validity of the examination technique. As microscopic similarities/correspondence diminish in the firing sequence, an inconclusive result becomes increasingly likely. However, this changing-tool phenomenon does not increase the likelihood of false positive errors.

36. The late researcher Alfred Biasotti recently has been selectively quoted (reference attached) in such a way that suggests reliable bullet comparisons were problematic. The passage in Biasotti's 1959 article reads more fully as follows:

> "Two basic types of data were recorded: (1) The total line count and total matching lines per land or groove mark from which the percent matching lines were derived (Tables 3 and 4); and (2) the frequency of occurrence of each serious of consecutive matching lines for which probability estimates were calculated (Figures 4 to 8).
>
> "Dealing first with the data for percent matching lines given by Tables 3 and 4, it will be seen that the average percent match for bullets fired from the same gun ranged from 36 to 38% for lead bullets and from 21 to 24% for metal-cased bullets. For bullets fired from different guns (not tabulated) 15 to 20% matching lines per land or groove mark was frequently found. Relatively speaking, this data indicates that even under such ideal conditions the average percent match for bullets from the same gun is low and the percent match for bullets from different guns is high, which should illustrate the limited value of percent matching lines without regard to consecutiveness.... [Alfred A. Biasotti, A Statistical Study of the Individual Characteristics of Fired Bullets. *Journal of Forensic Sciences*; volume 4, number 1, 1959]

The point Biasotti was making was that there is no value in counting the percentage of matching lines (straie) in a bullet comparison, which is a fact understood by firearms examiners for about as long as firearms identification has been practiced. Biasotti went on to re-affirm instead that it is consecutiveness that matters:

> "Since no two objects are ever absolutely identical, a realizable or practical identity must be based on the occurrence of a sufficiently high number of corresponding individual characteristics having a very low probability of having occurred as a result of chance, and therefore must be the result of a common cause. It should be obvious that consecutiveness; viz., the compounding of a number of individual characteristics, is the very basis of all identities.

And it is the consecutiveness of matching straie that counts for much when examiners use the traditional pattern-matching methods in their examinations.

37. Firearms and toolmark identification involves some degree of subjectivity when an examiner looks for a high degree of correspondence in patterns. Doubtless the methodology is similar to matching dental records to a particular person. It is also analogous to the manner in which we recognize people in

8

everyday life. When we see a friend or relative in public we are able to make an identification based upon patterns of features that match our memories. Familiarity with a particular subject is what enables us to make an identification of a face with a high level of confidence. This explains why parents of identical twins can typically distinguish between their children with practical certainty. Similarly, a medical researcher may know each mouse by name. The practiced eye of the firearms examiner is trained to recognize corresponding marks on bullets and cartridge casings. It should be noted that all sciences involve some elements of subjectivity, whether it's taking readings from an analog instrument; or interpreting epidemiological data, for example; or interpreting the meaning of a fossil or bone; or a physician diagnosing a fever. Subjectivity is not tantamount to unscientific, nor to unreliability. Each theory or technique, whether more or less subjective/objective, must be empirically tested on its own terms to determine its level of validity and reliability.

38. The AFTE theory of firearms identification merely adopted and articulated traditional principles of pattern matching that have enjoyed broad acceptance within the forensic firearms community for decades. Traditional pattern matching is practiced by firearms and toolmark examiners in forensic laboratories throughout the world. According to the SWGGUN Survey Summary of laboratories in the United Sates, 98% of laboratories that responded to the study utilize traditional pattern matching.

39. The defense points out that "a number of known none-matched test fires from different firearms" have been observed to appear "near the top of the [same gun] candidate list" in large image databases. This is completely predictable in any large database of ballistic images. But there is no evidence that an increase in similar-images in databases has lead to misidentifications under a comparison microscope. No identifications are made based upon matches in a database alone. Any positive identifications relating to individual characteristics are made under a comparison microscope. Matches in the National Integrated Ballistic Information Network (NIBIN) are merely a starting point for further examination.

40. Defense has cited an affidavit by William Tobin, a former FBI metallurgist. Although metallurgy and materials science may provide a general understanding of the manufacturing processes for products such as firearms and common tools, they do not provide a detailed knowledge of the firearm and toolmark identification process and the conclusions that can be drawn from examinations under a comparison microscope. There is no indication that Mr. Tobin has performed any firearms identification casework or undergone any formal training in the field. From his metallurgy background, Mr. Tobin makes broad assertions about what conclusions *cannot* be drawn in the field of firearms and toolmark identification, but provides no specific evidence or research studies to support his assumptions.

Based upon Mr. Tobin's affidavit, the defense appears to be asserting that in order to conduct a reliable examination, firearms examiners must watch the production of every single item or have a detailed knowledge of every manufacturing facility and its processes. However, as noted above, firearms examiners are well aware of the issue of subclass markings and are continually investigating new and novel manufacturing processes to insure that such marks are not produced, or if they are, that the examiner community is alerted to them via publication or other means. Further, if in casework an examiner examines items for which he knows subclass marks could potentially have occurred, best practices dictates that he

9

account for this and insure that the strength of his conclusions correspond to the strength or weaknesses of the underlying evidence examined; or, alternatively, before stronger identification conclusions justifiably could be drawn, that he conduct additional and detailed research in his particular case to insure no subclass marks were produced. (It should be noted, however, that if subclass marks are suspected, it is highly likely they are present on only one surface area of a specimen. For example, if breechface marks in a particular instance are known to be problematic, then the examiner would not conclude identify unless there were sufficient microscopic correspondence in non-subclass firing pin impressions, chamber marks, etc.)

(b)(6) per EOUSA

Stephen G. Bunch, Ph.D

State of Virginia

SS:

County of Stafford

Before me the undersigned, a Notary Public for __Stafford__ County, State of Virginia, personally appeared STEPHEN G. BUNCH, and he being first duly sworn by me upon his oath, says that the statements contained in the above affidavit are true.

Signed and sealed this 29th day of May 2008

(b)(6) per EOUSA
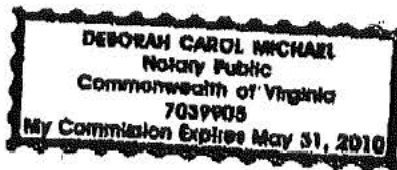
My commission expires May 31, 2010

DEBORAH CAROL MICHAEL
Notary Public
Commonwealth of Virginia
7039905
My Commission Expires May 31, 2010

10

1095

7340d2d7-67ae-4b31-9c9d-0419dd510c7a   20220314-17115

SUPPLEMENTAL STATEMENT OF STEPHEN G. BUNCH

I, Stephen G. Bunch, , state the following to be true to the best of my knowledge:

1. As noted in my prior sworn statement in this matter, I am currently the Supervisory Physical Scientist (Unit Chief) at the Federal Bureau of Investigation (FBI) in Quantico, Virginia. The purpose of this supplemental statement is to respond to certain issues raised in the defense's most recent pleading in U.S. v. Worsley, 2003 FEL 6856.

2. The defense asserts that firearms examiners have taken fundamental assumptions of uniqueness and reproducibility for granted (Defendant's Reply at 6). However, as indicated in my earlier statement, these principles have been demonstrated, and continue to be demonstrated, by researchers in the form of validation studies and proficiency tests. If both uniqueness and reproducibility did not exist, validity and proficiency tests would result in either inconclusive results or in false positives/negatives. Instead, firearm examiners have repeatedly demonstrated an ability to properly identify markings to specific firearms, often under extreme (worst case scenario) conditions, with no error rate at all or with very low error rates.

3. The defense cites to my discussion regarding the word "uniqueness" in an effort to create a controversy within the scientific community (Defendant's Reply at 8). There is no controversy whatsoever. In all forensic disciplines objects will display uniqueness at some microscopic level. This fact is easily illustrated. For example, if one were to look at two "identical" bar codes affixed to the same product type, those bar codes would be read as "identical" by the bar code scanner. However, microscopic examination of the two bar codes would certainly reveal countless differences. This is analogous to the similarities of "individual characteristics" on bullets and cartridge cases which provide sufficient similarities to conclude identity, but, at some level, are clearly unique. The word "uniqueness" is simply used in the firearms and toolmark identification discipline to signify a level of similarity that allows for identifications.

4. The defense intimates that I have acknowledged that the firearms and toolmark identification community is divided over the use of traditional pattern matching versus Consecutive Matching Striation Criteria (CMS) (Defendant's Reply at 11). However, the mere fact that a minority of firearms examiners choose to quantify their identifications using CMS, does not in any way invalidate the traditional use of pattern matching. Based upon my participation in AFTE and SWGGUN, it is my strong belief that CMS practitioners, such as Ron Nichols or Bruce Moran, would wholeheartedly concur that CMS in no way invalidates tradition pattern matching.

5. The defense overstates the difficulties presented with the issue of subclass marks. In fact, validation studies involving Ruger pistols, like the weapon at issue in this case, have not revealed any issue of subclass markings and firearms examiners have demonstrated an ability under rigorous testing conditions to identify cartridge cases and bullets to Ruger pistols. See Erich D. Smith, Cartridge Case and Bullet Comparison Validation Study with Firearms Submitted in Casework, AFTE Journal, Vol. 36, No. 4, p. 130 (Fall 2004); David J. Brundage, The Identification of Consecutively Rifled Gun Barrels, AFTE Journal, Vol. 30, No. 3, p.438 (Summer 1989). The articles cited by the defense merely reference two instances of subclass markings that have been widely known to practitioners in the firearms community

1

for quite some time. See Matty W. And Johson T., A comparison of manufacturing marks on Smith & Wesson firing pins, AFTE Journal 1984, 16(3), 51-56; Thompson E. Phoenix, Arms (Raven) breech face toolmarks, AFTE Journal 1994, 26(2), 134-134. In fact, I first learned about the studies cited by the defense during my examiner training period. The article by M.S. Bonfonti & J. De Kinder merely supports my earlier contention that firearms examiners are readily alerted, via publication or otherwise, to issues relating to subclass characteristics.

7. The defense cites to an article by Bruce Moran in support of its contention that "for decades ... firearms examiners failed to recognize the existence of sub-class characteristics" and that "it was the discovery of errors that, in the 1980s and 1990s, led the AFTE to adopt its classification system." (Defendant's Reply at 13). This assertion implies that errors in past decades resulted from the existence of unrecognized subclass characteristics. However, the Moran article does not link past errors to a failure to recognize subclass characteristics. In my experience, subclass marks have not materialized as a problem in casework or validity studies. Indeed, if the existence of subclass markings was an important and pervasive problem, then error rates in validity studies involving consecutively manufactured firearms would have been far higher, or at least present. The same can be said for the error rates in proficiency tests.

(b)(6) per EOUSA

Stephen G. Bunch, Ph.D

State of Virginia

SS:

County of Stafford

Before me the undersigned, a Notary Public for ___Stafford___ County, State of Virginia, personally appeared STEPHEN G. BUNCH, and he being first duly sworn by me upon his oath, says that the statements contained in the above affidavit are true.

Signed and sealed this 8th day of July 2008

(b)(6) per EOUSA

My commission expires _____

NANCY E. WRIGHT
Notary Public
Commonwealth of Virginia
343702
My Commission Expires Apr 30, 2012

2
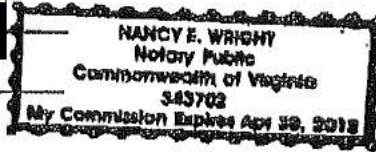
SECOND SUPPLEMENTAL STATEMENT OF STEPHEN G. BUNCH

I, Stephen G. Bunch, , state the following to be true to the best of my knowledge:

1. As noted in my prior sworn statement in this matter, I am currently a Supervisory Physical Scientist (Unit Chief) at the Federal Bureau of Investigation (FBI) Laboratory in Quantico, Virginia. The purpose of this supplemental statement is to respond to certain issues raised in the defense's most recent pleading in United States v. Willie Gayden, 2006 CFI 27899.

2. Throughout my 12.5-year tenure in the FBI Firearms and Toolmarks Unit, I am not aware of a single instance in which an independent examination by a defense expert has found what was later confirmed to be a misidentification by one of the firearms examiners at the FBI Laboratory.

3. In the recent report issued by the National Academy of Sciences ("NAS") entitled, *Strengthening Forensic Science in the United States: A Path Forward*, the section devoted to toolmarks references the following five citations to articles that appear in scientific journals: A.A. Biasotti, *A statistical study of the individual characteristics of fired bullets*, Journal of Forensic Sciences 4:34 (1959); A.A. Biasotti and J. Murdock, *"Criteria for Identification" or "state of the art" of firearms and tool marks identification*, Journal of the Association of Firearms and Tool Mark Examiners 16(4):16 (1984); J. Miller and M.M. McLean, *Criteria for identification of tool marks*, Journal of the Association of Firearms and Tool mark Examiners, 30(1):15 (1998); J.J. Masson, *Confidence level variations in firearms identification through computerized technology*, Journal of the Association of Firearms and Tool Mark Examiners 29(1):42 (1997); R.G. Nichols, *Defending the scientific foundations of the firearms and tool mark identification discipline: Responding to recent challenges*, Journal of Forensic Sciences, 52(3);586-594 (2007). None of these articles questions the ability of a trained firearms examiner to match toolmarks to a particular firearm. Three of the articles explore the development of an objective-criteria decision-making scheme for effecting bullet identifications commonly referred to as the CMS (consecutive matching striations) method. As noted in my previous affidavit, advocates of CMS do not question the scientific validity of traditional pattern matching. To the contrary, CMS is used by its practitioners in conjunction with pattern matching when effecting bullet identifications. When effecting identifications by observing impressed marks, examiners who practice the CMS method continue to use traditional pattern matching but without the CMS decision criteria.

These articles also explore the application of statistics to firearms identifications. Once again, nothing said in these articles questions the ability of a trained firearms examiner to match toolmarks to a particular firearm. Over the past few decades, there have been attempts by some researchers and firearms examiners, including myself, to explore the application of statistical methods to firearms and toolmark identification. These efforts in no way invalidate the application of the traditional pattern matching methodology. I'm unaware of any authors of peer-reviewed research in this area who have asserted that the lack of a statistical probability attached to an identification conclusion for a toolmark to a firearm somehow invalidates the identification conclusion.

The article by Masson merely points out that as a database of images of bullets and/or cartridge cases becomes larger and larger, there will be increasing numbers of non-matches that appear more and

1

more similar in their microscopic correspondence. This is just common sense. The author argues that this reality could be used as a valuable training tool (i.e., for trainees to observe known-non-matches that are very similar, so that they then would not identify on actual comparisons displaying lesser degrees of striation correspondence—thus resulting in a lower risk of misidentification).

The article by Nichols simply addresses criticism by non-firearms examiners.

Most articles in peer-reviewed scientific journals about the field of firearms and toolmark identification are written by individuals who are trained firearms examiners. The same is true for the totality of validation studies in the field. For these reasons, most researchers who have explored ways to improve this area of forensic science or who have sought to validate the field have themselves practiced the traditional pattern matching methodology.

4. The NAS report also cites to: J.E. Hamby, D.J. Brundage, and J.W. Thorpe, *The identification of bullets fired from 10 consecutively rifled 9mm Ruger pistol barrels – A research project involving 46 participants from 19 countries*. This study is a continuation of a prior validation study (referenced in my original affidavit) and demonstrates that the pattern matching methodology also is practiced by qualified firearm examiners from around the world and with equally reliable findings (zero false identifications).

5. The defense has made reference to comments made by Dr. Chris Hassell, Assistant Director of the FBI Laboratory, during a panel discussion at an April 2009 Arizona conference on forensic evidence. In particular, the defense states that Dr. Hassell remarked that, in light of the NAS forensic report, the FBI Laboratory's firearms/toolmarks examiners no longer reach conclusions of individualization. I was not present at the conference. However, the FBI Laboratory's firearms/toolmarks examiners have and will continue to reach conclusions of identity (or individualization) for those comparisons in which they observe sufficient microscopic agreement to justify this conclusion.

An examiner cannot rule out with absolute certainty the highly unlikely event that two different firearms produce indistinguishable individual characteristics; therefore, when asked, an examiner must properly explain the meaning of an identification conclusion. One way to do this is to explain that the match is one of "practical certainty," rather than one of "absolute certainty." Practical certainty means that the determination of identity correlates to features whose likelihood of reproduction by another tool is so low that such a possibility can be practically ignored. Another way to properly qualify an identification is to state that the examiner has matched a toolmark to a particular firearm "to a reasonable degree of scientific certainty." Either qualification communicates the examiner's high level of assurance without overstating the significance of the match.

Stephen G. Bunch, Ph.D

2

1099

7340d2d7-67ae-4b31-9c9d-0419dd510c7a   20220314-17119
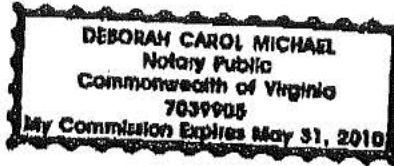
State of Virginia

SS:

County of Stafford

Before me the undersigned, a Notary Public for _Stafford_ County, State of Virginia, personally appeared STEPHEN G. BUNCH, and he being first duly sworn by me upon his oath, says that the statements contained in the above affidavit are true.

Signed and sealed this _5th_ day of June 2009

(b)(6) per EOUSA

My commission expires _May 31, 2010_

DEBORAH CAROL MICHAEL
Notary Public
Commonwealth of Virginia
7039905
My Commission Expires May 31, 2010

3

1100

7340d2d7-67ae-4b31-9c9d-0419dd510c7a   20220314-17120

FYI - background document on NIST Scientific Foundation Reviews

From:	"Cavanagh, Richard R. Dr. (Fed)" <(b) (6) >
To:	"Antell, Kira M. (OLP)" <(b) (6) "Hunt, Ted (ODAG)" <(b) (6)
Cc:	"Butler, John M. (Fed)" <(b) (6) >
Date	Thu, 06 Sep 2018 15 46 47 0400
Attachments:	NIST Scientific Foundation Reviews Background - DRAFT Sept 6.docx (88.77 kB)

Dear Ted and Kira,

As the Department of Justice is a valued partner with our forensic science activities at NIST, we are providing this document to you for your information. It is intended to provide NIST's perspective and expectations of the Scientific Foundations Reviews that we have undertaken This is just an FYI and we are not requesting formal DOJ feedback on it We plan to share this document publicly (potentially with further revisions) in draft form later this month along with a press release asking for feedback from interested parties

Thanks,
Rich

RE: Proposed Presentations for FRE Conference

From: "Hunt, Ted (ODAG)" <(b) (6)

To: "Antell, Kira M. (OLP)" <(b) (6)   "Young, Cynthia (USAMA)" <(b) (6)   >, "Goldsmith, Andrew (ODAG)" <(b) (6)   "Hafer, Zachary (USAMA)"
(b) (6)

Cc: "Shapiro, Elizabeth (CIV)" <(b) (6)   >, "Begian, Lernik (OLP)" <(b) (6)   "Smith, David L (USAEO)" (b) (6)   gov , "Ibrahim, Anitha (CRM)" (b) (6)

Date: Fri, 25 Aug 2017 09:48:57 -0400

Works for me.

From: Antell, Kira M. (OLP)
Sent Friday, August 25, 2017 8 59 AM
To: Young, Cynthia (USAMA) <(b) (6)   Goldsmith, Andrew (ODAG) <(b) (6)
Hafer, Zachary (USAMA) (b) (6)   ; Hunt, Ted (ODAG) (b) (6)
Cc: Shapiro, Elizabeth (CIV <(b) (6)   GOV>; Begian, Lernik (OLP) <(b) (6)   Smith, David L. (USAEO) (b) (6)   ; Ibrahim, Anitha (CRM) (b) (6)
Subject: RE: Proposed Presentations for FRE Conference

I am writing to follow up and make sure no one has any issues with the proposed presentations. Betsy will need to email Dan Capra with this information no later than this afternoon so we can verify that we will have three representatives. If you do have issues, please let us know by noon today so we can make any needed edits.

Thanks,
Kira

From: Antell, Kira M. (OLP)
Sent Wednesday, August 23, 2017 10 45 AM
To: Young, Cynthia (USAMA) <   (b) (6) Goldsmith, Andrew (ODAG) <   (b) (6)
Hafer, Zachary (USAMA) (b) (6)   Hunt, Ted (ODAG)   (b) (6)
Cc: Shapiro, Elizabeth (CIV <   (b) (6) >; Begian. Lernik (OLP) <   (b) (6) Smith, David L (USAEO)   Ibrahim, Anitha (CRM)   (b) (6)
Subject: Proposed Presentations for FRE Conference

Hello all,

Yesterday, Andrew Goldsmith, Betsy Shapiro, Ted Hunt, and I met to discuss presentations at the October FRE Mini-Conference on Forensics Dan Capra, the reporter, has suggested that we have too many Department representatives on the first panel and proposed that we reduce our representatives from three to two. The four of us agreed that each representative has a separate message and we believe that it is important for each representative to stay on the panel especially given the number of PCAST proponents that have been invited to participate.

I propose the following general topics for each presenter. I welcome your input. After we make any necessary edits to the below, (b) (5)
          After we confirm continued participation for all presenters, I will draft proposed talking points for each individual that you may find helpful

Zach do you have any information from Judge Grimm as to his expectations for your panel or who the other participants will be? I'm happy to assist in your preparation including drafting bios, summaries, or providing other information.

PROPOSAL FOR CAPRA:
(b) (5)

Kira Antell
Senior Counsel
Office of Legal Policy
U.S. Department of Justice
950 Pennsylvania Avenue, NW
Washington. DC 20530

RE: Proposed Presentations for FRE Conference

From: "Hunt, Ted (ODAG)" <(b) (6)
To: "Shapiro, Elizabeth (CIV)" (b) (6)   >, "Antell, Kira M. (OLP)" <(b) (6)
Cc: "Begian, Lernik (OLP)" <(b) (6)   "Smith, David L. (USAEO)" <(b) (6)   >,
"Ibrahim, Anitha (CRM)" <(b) (6)   >, "Young, Cynthia (USAMA)"
<(b) (6)   >, "Goldsmith, Andrew (ODAG)" <(b) (6)   "Hafer, Zachary
(USAMA)" (b) (6)
Date: Mon, 28 Aug 2017 13:18:21 -0400

I agree.

From: Shapiro, Elizabeth (CIV)
Sent Monday, August 28, 2017 11 17 AM
To: Antell, Kira M. (OLP) <(b) (6)
Cc Begian, Lernik (OLP) (b) (6)   Smith, David L (USAEO) (b) (6)   ; Ibrahim,
Anitha (CRM) <(b) (6)   >; Young, Cynthia (USAMA) <(b) (6)   >; Hunt, Ted
(ODAG) (b) (6)   Goldsmith, Andrew (ODAG) (b) (6)   Hafer, Zachary (USAMA)
<(b) (6)   >
Subject RE Proposed Presentations for FRE Conference

Here's a suggestion: (b) (5)

How
does that sound?

From Antell, Kira M (OLP)
Sent: Monday, August 28, 2017 9:29 AM
To Shapiro, Elizabeth (CIV) (b) (6)
Cc: Begian, Lernik (OLP) <   (b) (6)   Smith, David L. (USAEO)   >; Ibrahim,
Anitha (CRM)   ; Young, Cynthia (USAMA)   ; Hunt, Ted
(ODAG) <   Goldsmith, Andrew (ODAG) <   Hafer, Zachary (USAMA)
Subject: RE: Proposed Presentations for FRE Conference

Hi Betsy,

(b)(5)

Thanks,
Kira

From Antell, Kira M (OLP)
Sent: Friday, August 25, 2017 10:18 AM
To Shapiro, Elizabeth (CIV)   (b) (6)
Cc: Begian, Lernik (OLP) <   Smith, David L. (USAEO)   ; Ibrahim,
Anitha (CRM)   'Young, Cynthia (USAMA)'   ; Hunt, Ted
(ODAG) <   Goldsmith, Andrew (ODAG) <   Hafer, Zachary (USAMA)
Subject: RE: Proposed Presentations for FRE Conference

Thanks everyone. Betsy, please go ahead and share our proposal with Capra.

From: Young, Cynthia (USAMA) [mailto   (b) (6)
Sent Friday, August 25, 2017 9 39 AM
To: Goldsmith, Andrew (ODAG) <   (b) (6)   Antell. Kira M. (OLP) <   (b) (6)   Hafer,
Zachary (USAMA)   : Hunt, Ted (ODAG)
Cc: Shapiro, Elizabeth (CIV) <   (b) (6)   >; Begian. Lernik (OLP) <   (b) (6)   Smith, David
L (USAEO)   ; Ibrahim, Anitha (CRM)
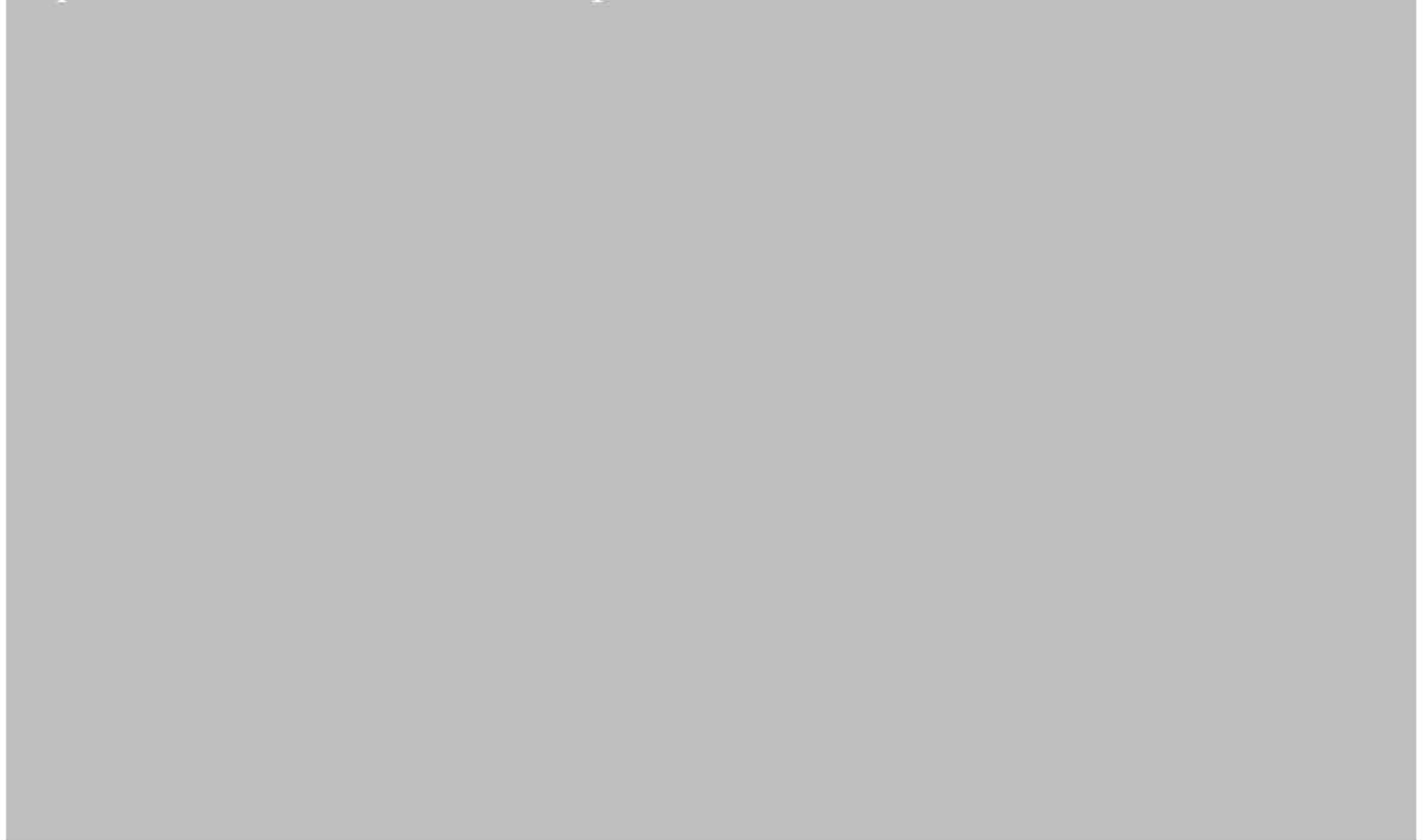**Subject:** RE: Proposed Presentations for FRE Conference

Thanks, Kira. Fine with me.

From: Goldsmith, Andrew (ODAG) [mailto ███████████ (b) (6)
Sent: Friday, August 25, 2017 9:27 AM
To: Antell, Kira M. (OLP) (JMD) <████████ (b) (6) Young, Cynthia (USAMA) ████████ (b) (6)>; Hafer, Zachary (USAMA) <████████ (b) (6)>; Hunt, Ted G) (JMD) <████████ (b) (6)>
Cc: Shapiro, Elizabeth (CIV) <████████ (b) (6)>; Begian, Lernik (OLP) (JMD) <████████ (b) (6)>; Smith, David L. (USAEO) <████████ (b) (6)>; b ahi Anitha (CRM) <████████ (b) (6)>
Subject: RE: Proposed Presentations for FRE Conference

Looks good to me.

From: Antell, Kira M. (OLP)
Sent: Friday, August 25, 2017 8:59 AM
To: Young, Cynthia (USAMA) <████████ (b) (6) Goldsmith, Andrew (ODAG) <████████ (b) (6) Hafer, Zachary (USAMA) ████████ >, Ted (ODAG) <████████ (b) (6)
Cc: Shapiro, Elizabeth (C   ) <████████ ; Begian, Lernik (OLP) <████████ (b) (6) Smith, David L. (USAEO) <████████ ; Ibrahim, Anitha (CRM) <████████
Subject: RE Proposed Presentatio   FRE Conference

I am writing to follow up and make sure no one has any issues with the proposed presentations. Betsy will need to email Dan Capra with this information no later than this afternoon so we can verify that we will have three representatives. If you do have i ue , plea e let u know by noon today o we can make any needed edit

Thanks,
Kira

From: Antell, Kira M. (OLP)
Sent: Wednesday, August 23, 2017 10:45 AM
To: Young, Cynthia (USAMA) <████████ (b) (6) Goldsmith, Andrew (ODAG) <████████ (b) (6) Hafer, Zachary (USAMA) ████████ , Ted (ODAG) <████████ (b) (6)
Cc: Shapiro, Elizabeth (C   ) <████████ ; Begian, Lernik OLP) <████████ (b) (6) Smith, David L. (USAEO) <████████ ; Ibrahim, Anitha (CRM) <████████
Subject: Pr posed Presentations to   E Conference

RE: Proposed Talkers for Call with Judge Livingston on 702

From:	"Hunt, Ted (ODAG)" <(b) (6) ████████████
To:	"Antell, Kira M. (OLP)" <(b) (6) ████████████
Date:	Mon, 18 Sep 2017 19:27:47 -0400
Attachment	Hunt Edit Propo ed Talker for Call with Judge Living ton on 702 doc (24 04 kB)

Here' my edit of the talker

From: Antell, Kira M. (OLP)
Sent: Monday, September 18 2017 5:06 PM
To: Hunt, Ted (ODAG) <(b) (6) ████████████ >
Subject: Proposed Talke   Livingston on 702

Hi Ted,

See what you think about these talkers and background. If you like, you can drop your section in and then we can send to Betsy and Andrew.

-K

Kira Antell
Senior Counsel
Office of Legal Policy
U.S. Department of Justice
950 Pennsylvania Avenue, NW
Washington. DC 20530
(b) (6)
(b) (6)

Talkers-My 2nd Edit

From: "Hunt, Ted (ODAG)" < (b) (6)

To: "Antell, Kira M. (OLP)" < (b) (6)

Date: Tue, 19 Sep 2017 10:10:30 -0400

Attachment: Hunt Second Edit Propo ed Talker for Call with Judge Living ton on 702 doc (24 3 kB)

I added (b)(5)

Ted R. Hunt
Senior Advisor to the Attorney General on Forensic Science
Office of the Deputy Attorney General
United States Department of Justice
950 Pennsylvania Ave, NW
Washington D C. 20530
(b) (6)
(b) (6)

RE: Talkers on PCAST for FRE

From:	"Hunt, Ted (ODAG)" < (b) (6)
To:	"Antell, Kira M. (OLP)" < (b) (6)
Date:	Thu, 28 Sep 2017 18:16:44 -0400

Yes, that's my weekend project. I'm trying to finish editing the NAS speech, and will get to that shortly.

From Antell, Kira M (OLP)
Sent: Thursday, September 28. 2017 5:52 PM
To Hunt, Ted (ODAG) (b) (6)
Subject: FW: Talkers on PCAST for FRE

Hi Ted,

Wanted to follow up on this. I don't think you've responded. If you have it all under control and do not need anything from me, that' great but want to be helpful

-K

From: Antell, Kira M. (OLP)
Sent: Monday, September 18, 2017 6:11
To: Hunt, Ted (ODAG) < (b) (6)
Subject: Talkers on PCAST

Hi Ted,

→(b) (5)

Let me know what you think.

Thanks,
K

Kira Antell
Senior Counsel
Office of Legal Policy
U.S. Department of Justice
950 Pennsylvania Avenue, NW
Washington. DC 20530
(b) (6)
(b) (6)

PCAST

Ted R. Hunt
Senior Advisor to the Attorney General on Forensic Science
Office of the Deputy Attorney General
United States Department of Justice
950 Pennsylvania Ave, NW
Washington, D.C. 20530
(b) (6) ██████████
(b) (6) ██████████

PCAST

Ted R. Hunt
Senior Advisor to the Attorney General on Forensic Science
Office of the Deputy Attorney General
United States Department of Justice
950 Pennsylvania Ave, NW
Washington D C. 20530
(b) (6)
(b) (6)

Re: Call on FRE Conference: Panel II

From: "Hafer, Zachary (USAMA)" <(b) (6) ███████████████>
To: "Goldsmith, Andrew (ODAG)" <(b) (6) █████████████
Cc: "Antell, Kira M. (OLP)" <(b) (6) ███████████ "Young, Cynthia (USAMA)" <(b) (6) █████████████>,
"Hunt, Ted (ODAG)" <(b) (6) ███ "Shapiro, Elizabeth (CIV)" <(b) (6) ████████████>, "Ibrahim,
Anitha (CRM)" <(b) (6) ████████████ >, "Smith, David L. (USAEO)" <(b) (6) ████████ >
Date: Fri, 01 Sep 2017 06:55:34 -0400

Thanks Andrew and Kira. I will definitely be there during the first panel and would love to see Andrew's materials whenever they're ready.

I will also circulate my outline as it develops and would welcome any input.

Hope you all have a nice long weekend.

Zach

Sent from my iPhone

On Aug 31, 2017, at 7:30 PM, Goldsmith, Andrew (ODAG) <(b) (6) ███████████ wrote:

> Kira - thanks for the update. I agree with your assessment, and will gladly share my materials with Zach so he knows where I'm going And pre umably Zach will be in the audience during the fir t panel o he can hear e actly what i said. - Andrew
>
> On Aug 31, 2017, at 3:08 PM, Antell, Kira M. (OLP) <(b) (6) ██████████ wrote:

Hello all,

This afternoon I joined a call with Zach, Judge Grimm, and Dan Capra. The purpose of the call was to better understand the contours of Zach's panel on 702 and Daubert in criminal cases and to find out the topic the judge and Capra wanted Zach to addre

(b) (5)
████████████████████████████
████████████████████████████

████████████████████████████
████████████████████████████

████████████████████████████
████████████████████████████

████████████████████████████
████████

(b) (5)

I have more complete notes and am happy to speak with you and the call in greater detail.

Thank ,
Kira

Kira Antell

Senior Counsel

Office of Legal Policy

U.S. Department of Justice

950 Pennsylvania Avenue, NW

Washington, DC 20530

(b) (6)

(b) (6)

Re: Call on FRE Conference: Panel II

From: "Goldsmith, Andrew (ODAG)" <(b) (6)
To: "Antell, Kira M. (OLP)" <(b) (6)
Cc: "Young, Cynthia (USAMA)" <(b) (6) >, "Hunt, Ted (ODAG)" <(b) (6)
"Shapiro, Elizabeth (CIV)" <(b) (6) >, "Hafer, Zachary (USAMA)" <(b) (6) >,
"Ibrahim, Anitha (CRM)" <(b) (6) >, "Smith, David L. (USAEO)"
(b) (6)
Date: Thu, 31 Aug 2017 19:29:49 -0400

Kira - thanks for the update. I agree with your assessment, and will gladly share my materials with Zach so he knows where I'm going. And presumably Zach will be in the audience during the first panel so he can hear exactly what is said. - Andrew

On Aug 31, 2017, at 3:08 PM, Antell, Kira M. (OLP) <(b) (6) wrote:

Hello all,

Duplicative Material see bates stamp numbers 20220314-10810 and 20220314-10811

PCAST report

(See attached file: PCAST Addendum.pdf)(See attached file: PCAST.pdf)

Ellen Leonida
Assistant Federal Public Defender
450 Golden Gate Avenue, Room 19-6884
San Franci co, CA 94102
Phone (415) 436-7700

RE: PCAST report

All,

Below is my take on framing the PCAST debate proposition

(b)(5) per EOUSA

Ted

(See attached file: PCAST Addendum.pdf)(See attached file: PCAST.pdf)

Ellen Leonida
Assistant Federal Public Defender
450 Golden Gate Avenue, Room 19 6884
San Francisco, CA 94102
Phone (415) 436-7700

RE: Justice Dept response

From: "Hunt, Ted (ODAG)" <(b) (6)

To: "Kaye, David" <(b) (6) @dsl.psu.edu>

Date: Fri, 03 Nov 2017 15:17:29 -0400

Hi David,

Sorry for the late response. I've been out of the office over the last few days and am just catching up.

I didn't have a formal prepared text for my remarks -- only a set of notes -- so there is no written statement to send. As I under tand it, the ympo ium wa recorded, and material that document the comment of each pre enter will be prepared. Typically, the Department only posts statements made by the Attorney General, Deputy Attorney General, and AAGs.

Good to ee you at the Sympo ium

Take care,

Ted

From: Kaye, David (b) (6) @dsl.psu.edu]
Sent: Tuesday, Octo 7 9:59 PM
To: Hunt, Ted (ODAG) <(b) (6)
Subject: Justice Dept re

Dear Ted,
I looked to see if the statement you read at the Boston meeting last week is on the DOJ web site and did not see it. Could you send me a copy? I am about to prepare a talk on the PCAST Report for the Northeast Association of Forensic Scientists meeting next week.
Thanks,
David

Re: Federal Rules of Evidence

From: "Isenberg, Alice R. (LD) (FBI)" (b)(6), (b)(7)(C), (b)(7)(E) per FBI
To: "Antell, Kira M. (OLP)" <(b) (6)
Cc: "Hunt, Ted (ODAG) <(b) (6)   (b)(6), (b)(7)(C), (b)(7)(E) per FBI (OGC) (FBI)" (b)(6), (b)(7)(C), (b)(7)(E) per FBI
Date Tue, 03 Oct 2017 10 15 51 0400

Ye   My a   i tant will   end it   hortly
Alice

       Original me   age
From: "Antell, Kira M. (OLP)" <(b) (6)
Date: 10/3/17 9:44 AM (GMT-05:00)
To: "Isenberg, Alice R. (LD) (FBI)" (b)(6), (b)(7)(C), (b)(7)(E) per FBI
Cc "Hunt, Ted (ODAG) (JMD)" (b) (6)   (b)(6), (b)(7)(C), (b)(7)(E) per FBI (OGC) (FBI)" (b)(6), (b)(7)(C), (b)(7)(E) per FBI
Subject: RE: Federal Rules of Evidence

Alice,

Do you have a short bio I can send to Capra in the event we get the invite?

Thanks,
Kira

From: (b)(6), (b)(7)(C), (b)(7)(E) per FBI (OGC) (FBI) [mailto: (b)(6), (b)(7)(C), (b)(7)(E) per FBI]
Sent: M   er 2, 2017 8:41 PM
To: Antell, Kira M. (OLP) (b) (6)   Isenberg, Alice R. (LD) (FBI) (b)(6), (b)(7)(C), (b)(7)(E) per FBI
Cc: Hunt, Ted (ODAG) <(b) (6)
Subject: Re: Federal Rules of Evidence

Thanks and just hope the moderator can keep Lander to his alloted time.

       Original message
From: "Antell, Kira M. (OLP)" <   (b) (6)
Date 10/2/17 5 47 PM (GMT 07 00)
To: "Isenberg, Alice R. (LD) (FBI)" (b)(6), (b)(7)(C), (b)(7)(E) per FBI
Cc "Hunt, Ted (ODAG) (JMD)"   (b) (6)   (b)(6), (b)(7)(C), (b)(7)(E) per FBI (OGC) (FBI)" (b)(6), (b)(7)(C), (b)(7)(E) per FBI
Subject: Re: Federal Rules of Evidence

Thanks Alice! I really appreciate it. I'll let you know ASAP.

+ (b)(6), (b)(7)(C) (initial omission inadvertent)

Sent from my iPhone

On Oct 2, 2017, at 7:30 PM, Isenberg, Alice R. (LD) (FBI) (b)(6), (b)(7)(C), (b)(7)(E) per FBI wrote:

> Kira,
> Yes, I am available and am happy to pull out my talking points and sharpen them up. I'll even promise that
> the FBI will pay for my travel expenses! Just let me know if and when you get a green light.
> Thanks for thinking of us,
> Alice
>
>
> -------- Original message --------
> From: "Antell, Kira M. (OLP)" <   (b) (6)
> Date: 10/2/17 5:38 PM (GMT-05:00)

To: "Isenberg, Alice R. (LD) (FBI)" [REDACTED]
Cc: "Hunt, Ted (ODAG) (JMD)" < [REDACTED] (b) (6)
Subject: Federal Rules of Evidence

Hi Alice,

I've mentioned that the committee in charge of the federal rules of evidence is holding a conference on potential changes to rule 702 for forensic experts   The conference is on Friday, 10/27 in Boston   Bruce Budowle was planning to attend but is now unable to travel. This created an opportunity for another perspective. I looked at the agenda and was struck that there was no representation from the forensics community and suggested that the Department should provide a practitioner to speak about the modern practice of forensics, the kinds of validation test that are performed, and the rigorous competency and proficiency testing that is done

Unclear whether we can wrangle and invite but if we can, do you think you would be able to provide this position? Andrew Goldsmith and Ted are doing the actual PCAST rebuttal and legal arguments but I think there is a need to hear form someone who can actually describe how forensic analysis is done

Thanks,
Kira

Kira Antell
Senior Counsel
Office of Legal Policy
U.S. Department of Justice
950 Pennsylvania Avenue, NW
Washington. DC 20530
(b) (6)
(b) (6)

Re: Rules Committees Forensics Proposals

From: Nelson Bunn <(b) (6)　　　　　　　　>
To: "Antell, Kira M. (OLP)" <(b) (6)　　　　　>
Cc: "Hunt, Ted (ODAG)" <(b) (6)　　　　>
Date Tue, 17 Oct 2017 12 01 55 0400

I've got another meeting during that time this afternoon. Would 3pm work tomorrow afternoon?

Nelson

Nel on O Bunn, Jr
Acting Executive Director
National District Attorneys Association
1400 Crystal Drive, Suite 330
Arlington VA 22202
O: (b) (6)
C: (b) (6)
(b) (6)

From: Kira Antell <(b) (6)　　　　　　>
Date: Tuesday, October 17, 2017 at 11:44 AM
To: Nelson Bunn (b) (6)
Cc: "Hunt, Ted (ODAG)" <(b) (6)
Subject: RE: Rules Committees Forensics Proposals

Hi Nelson,

Are you free thi   afternoon between 2 00 and 3 00 or tomorrow after 1 00? If not, let me know when would be convenient for a quick call.

Thanks,
Kira

From Nelson Bunn [mailto (b) (6)
Sent: Sunday, October 15. 2017 12:27 PM
To Antell, Kira M (OLP) (b) (6)
Cc: Hunt, Ted (ODAG) <(b) (6)
Subject Re Rules Committees Forensics Proposals

Hey Kira,

Thanks for flagging both of these items for us I'm more than happy to chat this week as well if you could just let me know some good days and times for you.

Nelson

Nelson O. Bunn, Jr.
Acting Executive Director
National District Attorneys Association
1400 Cry tal Drive, Suite 330
Arlington. VA 22202
O: (b) (6)
C: (b) (6)
(b) (6)

From: Kira Antell <　　　　　　　(b) (6)
Date: Saturday, October 14, 2017 at 11:52 AM
To: Nelson Bunn <　　　　　　(b) (6)
Cc: "Hunt, Ted (ODAG)" <　　　　(b) (6)
Subject: Rules Committees Forensics Proposals

Hi Nelson,

I hope you are well. I know NDAA has many members interested in efforts to restrict the use of forensic evidence in court   As such, I thought you would be interested into two proposals pending before Federal Rules Committees   They are both at very early stages.

The first link is to materials available on the AO's website about the Advisory Committee on Evidence's upcoming meeting   That committee is holding a mini conference in Boston on 10/27 on forensic science in light of the PCAST report. The Reporter has made an initial suggestion that a new Federal Rule of Evidence to limit forensic expert testimony could be appropriate   I will be attending that meeting
http://www.uscourts.gov/rules-policies/archives/agenda-books/advisory-committee-rules-evidence-october-2017 (Tab 9)

The second link is to a related proposal before the Advisory Committee on Criminal Rules to change Rule 16 on disclosure of expert testimony.   While the proposed change would not be limited to forensic experts, it is certainly tied to forensics.
http://www.uscourts.gov/sites/default/files/2017-10-criminal-agenda-book_0.pdf (Tab 6)

I'm in the office next week and would be happy to discuss these proposals with you.

Thanks,
Kira

Kira Antell
Senior Counsel
Office of Legal Policy
U.S. Department of Justice
950 Pennsylvania Avenue, NW
Washington. DC 20530
(b) (6)
(b) (6)

RE: Justice Dept response

From:	"Kaye, David" <​(b) (6)​@dsl.psu.edu>
To:	"Hunt, Ted (ODAG)" <(b) (6)
Date:	Sun, 05 Nov 2017 09:31:00 -0500

Hi Ted,
No problem. Because your public presentation was so well organized and crisp, and was designated an official position of the Department, I thought it might be in written form already. I'll just refer to the general idea in my talk this week.
Best,
David

From: Hunt, Ted (ODAG) [mailto:(b) (6)
Sent: Friday, Nov___er 3, 2017 3:17 PM
To: Kaye, David <​(b) (6)​@dsl.psu.edu>
Subject: RE: Justice Dept response

Hi David,

Sorry for the late response. I've been out of the office over the last few days and am just catching up.

I didn't have a formal prepared tet for my remark only a et of note o there i no written tatement to end A I understand it, the symposium was recorded, and materials that document the comments of each presenter will be prepared. Typically, the Department only posts statements made by the Attorney General, Deputy Attorney General, and AAGs.

Good to see you at the Symposium.

Take care,

Ted

From: Kaye, David [mailto:(b) (6)@dsl.psu.edu]
Sent: Tuesday, October 31 2017 9:59 PM
To: Hunt, Ted (ODAG) <(b) (6)
Subject: Justice Dept response

Dear Ted,
I looked to see if the statement you read at the Boston meeting last week is on the DOJ web site and did not see it. Could you send me a copy? I am about to prepare a talk on the PCAST Report for the Northeast Association of Forensic Scientists meeting next week.
Thanks,
David

RE: Forensics Law Review Articles

Good by me – thanks to all!

From: Bolitho, Zachary (ODAG)
Sent: Monday, January 29, 2018 6:35 PM
To: Antell, Kira M. (OLP) (b) (6)   Terwilliger, Zachary (ODA   (b) (6)
Cc: Hunt, Ted (ODAG) <(b) (6)   smith, Andrew (ODAG) <(b) (6)   apiro,
Elizabeth (CIV) <(b) (6)   Hur, Robert (ODAG) <(b) (6)
Subject: RE: For

If the ethics folks have signed off, I see no issues. Of course, I defer to Rob and Zach T. if they see any issues.

Thank ,
Zac

From: Antell, Kira M. (OLP)
Sent: Monday, January 29, 2018 4:44 PM
To: Bolitho, Zachary (ODAG) <   (b) (6) Terwilliger, Zachary (ODAG) <   (b) (6)
Cc: Hunt, Ted (ODAG) <   (b) (6) Goldsmith, Andrew (ODAG) <   (b) (6) Shapiro,
Elizabeth (CIV) <   (b) (6) Hur, Robert (ODAG) <   (b) (6)
Subject: RE: Forensics Law Review Articles

Good afternoon,

Circling back on this. Any issue with moving forward? Ted has one minor edit to his article (addition of a cite) so I plan to submit tomorrow. Please do let me know as soon as possible if I need to hold submission.

Thanks,
Kira

From: Antell, Kira M. (OLP)
Sent: Friday, January 26, 2018 11 58 AM
To: Bolitho, Zachary (ODAG) <   (b) (6) Terwilliger, Zachary (ODAG) <   (b) (6)
Cc: Hunt, Ted (ODAG)   (b) (6) Goldsmith, Andrew (ODAG)   (b) (6) Shapiro,
Elizabeth (CIV) <   (b) (6) Hur, Robert (ODAG) <   (b) (6)
Subject: Forensics Law Review Articles

Good afternoon,

In October, the Department presented at a forensics evidence symposium at Boston College. The purpose of the symposium was to discuss whether it was appropriate to amend Rule 702 for cases involving forensic evidence. The transcript of the symposium will be published in an upcoming issue of the *Fordham Law Review*. Department speakers were invited to provide hort article to the *Fordham Law Review Online* for March publication   Ted, Andrew, and Alice Isenberg from FBI lab have written articles.

We expect to submit these articles to Fordham on Monday by COB. Rob attended the symposium and hoped to have a chance to review the article but ha a ked u to proceed with clearance given our relatively tight timeline   The ethic office indicates there are no issues on their end.

Ted's article is a direct written response to the PCAST report drawn from previous public remarks but it is more granular than previou ly provided tatement   Andrew' article i quite imilar to hi approved tatement from the ympo ium but provides more in depth legal arguments. Both of them contain Department legal and policy positions. Alice's article is a more technical review of lab procedures.

Draft articles by Andrew and Ted were reviewed by a small group of people from CIV (Betsy CC'd here), CRM, CRM-Appellate, and the Criminal Chiefs Working Group. FBI reviewed Alice's article. No one believed there were any issues with publication.

Please do let me know if you have questions or concerns prior to submission.

Thanks,
Kira

Kira Antell
Senior Counsel
Office of Legal Policy
U.S. Department of Justice
950 Pennsylvania Avenue, NW
Washington. DC 20530
(b) (6)
(b) (6)

Re: Forensics Law Review Articles

From: "Antell, Kira M. (OLP)" <(b) (6)

To: "Bolitho, Zachary (ODAG)" <(b) (6)

Cc: "Terwilliger, Zachary (ODAG)" <(b) (6)   "Hunt, Ted (ODAG)" <(b) (6)
"Goldsmith, Andrew (ODAG)" <(b) (6)   "Shapiro, Elizabeth (CIV)"
<(b) (6)   "Hur, Robert (ODAG)" <(b) (6)

Date: Mon, 29 Jan 2018 20:59:08 -0500

Thanks all.

Sent from my iPhone

On Jan 29, 2018, at 6:35 PM, Bolitho, Zachary (ODAG) <(b) (6)   wrote:

(b) (6)

(b) (6)

RE: FRE Spring Meeting Talkers_04202018_v2

From: "Antell, Kira M. (OLP)" <(b) (6)
To: "Goldsmith, Andrew (ODAG)" <(b) (6)   "Shapiro, Elizabeth (CIV)"
<(b) (6)
Cc: "Hunt, Ted (ODAG)" <(b) (6)
Date: Fri, 20 Apr 2018 16:43:20 -0400

Thanks

From Goldsmith, Andrew (ODAG)
Sent: Friday, April 20, 2018 4:43 PM
To Antell, Kira M (OLP) (b) (6)   Shapiro, Elizabeth (CIV) (b) (6)
Cc: Hunt, Ted (ODAG) <(b) (6)
Subject RE FRE Spring Meeting Talkers 04202018 v2

I think this looks very through. I changed the style of the note to Rob on page 3 to (b) (5)

.

From: Antell, Kira M. (OLP)
Sent: Friday, April 20, 2018 12:25 PM
To: Shapiro, Elizabeth (CIV) <   (b) (6)>; Goldsmith, Andrew (ODAG) <   (b) (6)
Cc: Hunt, Ted (ODAG) <   (b) (6)
Subject: FRE Spring Meeting Talkers_04202018_v2

Attached are revised talkers for 702. I am pulling post-PCAST cases and will include in distro to Rob. Please let me know if you have thoughts or additions prior to about 5:00 today. If I don't hear anything, I send COB

Thanks,
Kira

RE: Talkers for Spring Advisory Committee Meeting

From:	"Antell, Kira M. (OLP)" <(b) (6)
To:	"Shapiro, Elizabeth (CIV)" <(b) (6) >
Cc:	"Hunt, Ted (ODAG)" <(b) (6) "Goldsmith, Andrew (ODAG)" <(b) (6)
Date:	Thu, 12 Apr 2018 09:17:56 -0400
Attachment	agenda book advi ory committee on rule of evidence EDITED FOR FORENSICS_COMMENTS.docx (111.43 kB); FRE Spring Meeting Talkers_04122018.docx (28.34 kB); Hunt's Additions-FRC Meeting Talkers.docx (15.82 kB)

Good morning,

Attached are
    (1) Revised talkers
    (2) Notes from Ted
    (3) Annotated forensics materials with comment bubbles.

Please let me know how you'd like to proceed. I can send these materials directly to Rob or you are welcome to send them to him. Also, happy to take edits form you or Andrew. Finally, were you able to get a VTC for Friday? I think it would be extremely helpful. Ted or Andrew's assistant may be able to help facilitate that.

Thanks,
Kira

From: Antell, Kira M. (OLP)
Sent: Wednesday, April 11, 2
To: Shapiro, Elizabeth (CIV) <(b) (6)   Hunt, Ted (ODAG) <(b) (6)   Goldsmith, Andrew (ODAG) <(b) (6)
Subject: ommittee Meeting

Attached are proposed high level talkers for Rob for the Spring meeting. Ted has been designated the responsibility to talk about the Department forthcoming projects and commitment so this reflects just responses to the memo.

I have also attached a version of the forensics portion of the materials with comment bubbles. Most of it is included in the talkers but thought you might find it helpful.

I am happy to take your comments and edits – I am free this afternoon or tomorrow morning if anyone would like to provide them by phone. Otherwise, please email me edits and then I'll ask Betsy to share with Rob tomorrow.

Thanks,
K

Kira Antell
Senior Counsel
Office of Legal Policy
U.S. Department of Justice
950 Pennsylvania Avenue, NW
Washington. DC 20530
(b) (6)
(b) (6)

RE: Probability Paper

Ted,

1. I actually read this paper last week and found it very compelling. I assume you've seen the paper out of UNIL (Biederman's colleagues) that criticized PCAST, but if not, I am attaching as well. We could try to invite Biederman to IPTES to discuss this more? On page 19, the authors address subjectivity, but unfortunately don't e pand too much

In the US environment, subjectivity has been associated with bias and sloppy thinking, and objectivity with an absence of bias and rigorous thinking. It is worthwhile examining whence the fear of subjectivity arises. There is considerable proof that humans are susceptible to quite a number of cognitive effects many of which can affect judgement <u>We suspect that the fear is that these effects bias the decisions in ways that are detrimental to justice. Hence, it is bias arising from cognitive effects that is the enemy, not subjectivity.</u>

They later tate in their conclu ion

That framework is provided by probability theory coupled with the <u>recognition that probability is necessarily subjective and conditioned by knowledge and judgement</u>. It follows that our view of the forensic scientist is a knowledgeable, logical and reasonable person Whereas data collectio s are valuable they should be viewed within the context of reliable knowledge

2. Also, Jon let me know that you also stumbled across the concept of 'convergent validation' in the AAAS Report. I've been intrigued with this concept and I've been asking my social scientists a little more. I was thinking it would be nice to get an e pert in thi area to come to IPTES, but I've yet to identify a true SME

3. Finally, I am attaching the Excel worksheets for the abstracts submitted to IPTES. I haven't had a chance to review since they came in last night, but I'll probably browse the titles this evening.

Best Regards,

Gerry

Gerry LaPorte
Director
National Institute of Justice
Office of Inve tigative & Foren ic Science
810 7th Street NW
Washington, DC 20531
Office Phone (b) (6)
Mobile Phone (b) (6)
Email: (b) (6)

For a li t of R&D NEEDS, plea e vi it http //nij gov/topic /foren ic /Document /fy15 foren ic twg table pdf

For a comprehensive list of Forensic Science R&D Projects, please visit
http://www.nij.gov/topics/forensics/Pages/research-development-projects.aspx.

For information related to Forensic DNA and DNA grant programs, please visit
http://www.nij.gov/topics/forensics/evidence/dna/Pages/welcome.aspx

For information relating to SE UAL ASSAULT, plea e vi it
http://www.nij.gov/topics/crime/rape-sexual-violence/pages/welcome.aspx

---

Gerry,

Take a look at the attached paper, and we'll follow up later. It addresses one of the areas that forensic feature comparison methods have taken a beating from critics   probabilistic statements of weight that utilize no explicit reference set of empirical data upon which inferences are drawn and statistical statements of weight can be offered.

Alex Biederman is probably the most active proponent (Taroni a close second) of what is described in this new paper as "justified subjectivism" in the field of forensics   Biederman, Taroni, and Aitken have written a number of papers about how subjective, or epistemic probability is every bit as legitimate in the field of probability theory as is empirical probability and/or frequentist statistical theory, and is actually more useful in forensics than is empirical probability

Biederman, Taroni, and Aitken, as you probably know, are big proponents of the use of LRs in forensic science, but the LR is simply one means by which a probability (empirical or epistemic) can be expressed.

(b) (5)
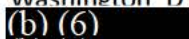
One of the topics for the pattern evidence symposium might relate to this topic, and provide some much-needed diversity of opinion on this issue would be   the foundational grounding in probability theory for qualitative statements of weight for forensic feature comparison disciplines. I have personally never heard a presentation on this topic at a mainstream forensic science conference; but it would certainly be a timely topic given the current climate of forensic critique, as well as the opportunities offered by this large symposium stage.

Ted

Ted R. Hunt
Senior Advisor to the Attorney General on Forensic Science
Office of the Deputy Attorney General
United States Department of Justice
950 Pennsylvania Ave, NW
Washington D C 20530
(b) (6)
(b) (6)

Review Article

Finding the way forward for forensic science in the US—A commentary on the PCAST report

I.W. Evett[*,a], C.E.H. Berger[b], J.S. Buckleton[c,d], C. Champod[e], G. Jackson[f]

[a] Principal Forensic Services Ltd., 34 Southborough Road, Bickley, Bromley, Kent, BR1 2EB, United Kingdom
[b] Institute for Criminal Law and Criminology, Faculty of Law, Leiden University, PO Box 9520, 2300 RA Leiden, The Netherlands
[c] Environmental Science & Research Ltd, Private Bag 92021, Auckland 1142, New Zealand
[d] Department of Statistical Genetics, University of Washington, Box 357232 Seattle, WA 98195-7232, United States
[e] Ecole des Sciences Criminelles, Faculty of Law, Criminal Justice and Public Administration, Université de Lausanne, Batochime — quartier Sorge, CH-1015 Lausanne-Dorigny, Switzerland
[f] Abertay University, Dundee, DD1 1HG, United Kingdom

ABSTRACT

A recent report by the US President's Council of Advisors on Science and Technology (PCAST), (2016) has made a number of recommendations for the future development of forensic science. Whereas we all agree that there is much need for change, we find that the PCAST report recommendations are founded on serious misunderstandings. We explain the traditional forensic paradigms of *match* and *identification* and the more recent foundation of the logical approach to evidence evaluation. This forms the groundwork for exposing many sources of confusion in the PCAST report. We explain how the notion of treating the scientist as a black box and the assignment of evidential weight through error rates is overly restrictive and misconceived. Our own view sees inferential logic, the development of calibrated knowledge and understanding of scientists as the core of the advance of the profession.

© 2017 Elsevier B.V. All rights reserved.

Contents

* Corresponding author.
  E-mail address: ianevett@btinternet.com (I.W. Evett).

In Memoriam

This paper is dedicated to the memory of Bryan Found who did so much to advance the profession of forensic scientist through his work on calibrating and enhancing the performance of experts under controlled conditions. He will be sorely missed.

1. Introduction

This paper is written in response to a recent report on forensic science of the US President's Council of Advisors on Science and Technology (PCAST) [1]. There have already been several responses to the report from the forensic community [2 7] which have resulted in an addendum to the report [8]. Our main concern is that the report (and its addendum) fails to recognise the advances in the logic of forensic inference that have taken place over the last 50 years or so. This is a serious omission which has led PCAST to a narrowly focussed and unhelpful view of the future of forensic science.

The structure of our paper is as follows. In Section 2 we briefly outline our view of the requirements imposed by logic on the assessment of the probative value of evidence. This allows us to set up a framework against which we can contrast some of the suggestions of the report. In Sections 3 and 4 we briefly explain the notions of "match" and "identification" paradigms that have underpinned much of forensic inference over the last century or so. Section 5 will point out misconceptions, fallacies, sources of confusion and improper terminology in the PCAST report. Our contrasting view of the future path for forensic science follows in Section 6.

2. The logical approach

Much has been written over the past 40 years on inference in forensic science. The frequency of appearance of articles, papers and books on the topic has increased markedly in recent years. Practically all of this material is founded on a logical, probabilistic approach to the assessment of the probative value of scientific observations [9,10]. The PCAST report mentions this body of work only briefly and pays scant attention to its principles [11], which we list and explain briefly as follows.

2.1. Framework of circumstances

It is necessary to consider the evidence within a framework of circumstances.

A simple example will illustrate this. Imagine that a sample[1] has been obtained from a crime scene which yielded a DNA profile from which the genotype of the originator of the sample has been inferred. A suspect for the crime is known to have the same genotype. Because the alleles revealed by a DNA profile will be found in different proportions in different ethnic groups, it is relevant to the assessment of the probative value of this correspondence of genotypes that a credible eyewitness of the crime said that the offender was of a particular ethnic appearance.

It follows that, when presenting an evaluation, the scientist should clearly state the framework of circumstances that are relevant to their assessment of the probative value of the observations, with a caveat that, if details of the circumstances change, the evaluation must be revisited.

2.2. Propositions

The probative value of the observations cannot be assessed unless two propositions are addressed.

In a criminal trial, these will represent what the scientist believes the prosecution may allege and a sensible alternative that represents the defence position.[2] In taking account of both sides of the argument, the scientist is able to assess the evidence in a balanced, justifiable way and display to the court an unbiased approach, irrespective of which side calls the witness.

Propositions may be formed at any of at least four levels in a hierarchy of propositions [12 14]. These levels are termed offence, activity, source and sub source. We do not discuss these in any depth here. Most of the PCAST report appears to address questions at the source or sub source level. Examples of these would be:

1. Sub source: The DNA came from the person of interest (POI),[3] or
2. Source: This fingermark was made by the POI.

2.3. Probability of the observations

It is necessary for the scientist to consider the probability[4] of the observations given the truth of each of the two propositions in turn.

The ratio of these two probabilities is widely known as the *likelihood ratio* (LR) and this is a measure of the weight of evidence that the observations provide in addressing the issue of which of the propositions is true. A likelihood ratio greater than one provides support for the truth of the prosecution proposition. A likelihood ratio less than one provides support for the truth of the defence proposition.

It cannot be sufficiently emphasized that it is the scientist's role to provide expert opinion on the probability of the *observations* given the proposition. The role of assigning a value to the probability of the *proposition* given the observations is that of the jury in a criminal trial. This probability will take account, not just of the scientific observations, but also of all of the other evidence presented at court.

---

[1] The term "sample" is used generically to describe what is available for forensic examination. The term is not used here to suggest any statistical sampling process.

[2] We recognise that the scientist, particularly at an early stage of proceedings, may not know the position that defence will take. It is common practice for the scientist to adopt what appears to be a reasonable proposition, given what is known of the circumstances—making it clear that this is provisional and subject to change at any time.

[3] A source level DNA proposition would specify the nature of the recovered material, e.g. "the semen came from the POI".

[4] This could be a probability density, depending on the nature of the observations. But the principle remains unchanged.

3. The match paradigm

In most forensic comparisons, one of the items will be from a known origin (such as: a reference sample for DNA profiling from a particular individual; a pair of shoes from a suspect; a set of control fragments of glass from a broken window). The other will be from an unknown, or disputed origin (such as: DNA recovered from a crime scene; a footwear mark from the point of entry at a burglary; or a few small fragments of glass recovered from the clothing of a suspect). It is convenient to refer to these as the *reference* and *questioned* samples, respectively. The matter of interest to the court relates to the origin of the questioned sample. This question will be addressed scientifically by carrying out observations on both samples. These observations may be purely qualitative: such as, for example, the shapes of the loops of letters such as "y" and "g" in a passage of handwriting. They may be quantitative and discrete, such as the alleles in a DNA STR profile. Or they may be quantitative and continuous, such as the refractive index of glass fragments. The match paradigm calls for a judgement, by the scientist, as to whether or not the two sets of observations agree within the range of what would be expected if the questioned sample had come from the same origin as the reference sample. The basis for that judgement may, in the case of quantitative observations, be based on a set of pre determined criteria; but where the observations are qualitative such criteria may be vague or purely judgemental.

If the two sets of observations are considered to be outside the range of what may have been expected if the two samples had come from the same source then the result may be reported as a "non match". Depending on the nature of the observations, this provides the basis for a strong implication that the questioned and reference samples came from different sources. In many instances this conclusion will be non controversial in the sense that prosecution and defence will be content to accept it.

However, when the result of the comparison is a "match" it does not logically follow that the two samples do share the same source or even that they are likely to be from the same source. It is possible that the two samples came from two different sources that, by coincidence, have similar properties. Throughout the history of forensic science there has been the notion often imperfectly expressed that the smaller the probability of such a coincidence, the greater the evidential value to be associated with the observed match. In DNA profiling, for example, we encounter the notion of a "match probability". The implication of this approach is that the jury should assign an evidential weight that is related to the inverse of the match probability.

The logical approach has done much to clarify the rather woolly inference that historically has been associated with the match paradigm but it has also demonstrated the considerable advan tages of the single stage approach implied by the assignment of weight through the calculation of the likelihood ratio, over the rather clumsy and inefficient two stage approach implied by the match paradigm. This has already been pointed out by Morrison et al. [4].

4. The identification paradigm

Historically, fingerprint comparison was seen to be the gold standard by which the power of any other forensic technique could be judged. The paradigm here was the notion of "identification"[5] or

"individualization" (the terms are used synonymously here). Provided that sufficient corresponding detail was observed, the outcome of a comparison between a fingermark of questioned origin and a print taken from a known person would be reported as a categorical opinion: the two were definitely made by the same person.

So, the match and identification paradigms are related with the difference that in the latter the scientist is allowed to state that the match probability is so infinitesimally small that it is reasonable to conclude that the two items came from the same source. Historically, many examiners would have claimed that the source was established with certainty to the exclusion of all others.

The identification paradigm went largely unchallenged for many years until later in the 20th century when its logical basis was questioned (see, for example, [16] or more recently [17,18]) and also when, in a number of high profile cases, misidentifications with serious consequences were exposed.

An example of the paradigm is given in box 6, p. 137 of the PCAST report (DOJ proposed uniform language) (emphasis added).

The examiner may state that it is his/her opinion that the shoe/tire *is the source of the impression* because there is sufficient quality and quantity of corresponding features such that the examiner would not expect to find that same combination of features repeated in another source. This is the highest degree of association between a questioned impression and a known source.

The PCAST report rightly indicates that the conclusions conveying "100 percent certainty" or "zero or negligible error rates" are not scientifically defensible. Such conclusions tend to overestimate the weight to be assigned to the forensic observations.

5. Misconceptions, fallacies and confusions in the PCAST report

The most serious weakness in the PCAST report is their flawed paradigm for forensic evaluation. Unfortunately, the report contains more misconceptions, fallacies, confusions and improper wording. In this section we will discuss the main problems with the report.

5.1. Confusion between the match and identification paradigms

This is the first source of confusion in the report. For example, from p. 90 of the report (emphasis added):

An FBI examiner concluded with "100 percent certainty" that the fingerprint *matched* Brandon Mayfield . . . even though Spanish authorities were unable to confirm the *identification*.

On p. 48 we find (emphasis added):

To meet the scientific criteria of foundational validity, two key elements are required:
(1) 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*").

We have seen that declaring a match and declaring an identification are not the same thing. Declaring a match implies nothing about evidential weight whereas declaring an identifica tion implies evidential weight amounting to complete certainty.

The PCAST report proposes an approach that is fusion of the match and identification paradigms. See, from p. 45/46:

---

[5] Kirk [15] defined the term identification as only placing an object in a restricted class. The criminalist would, for example, identify a particular mark as a fingerprint. Individualization was defined by Kirk as establishing which finger left the mark. An opinion of the kind "this latent mark was made by the finger which made this reference print" is an individualization.

Because the term "match" is likely to imply an inappropriately high probative value, a more neutral term should be used for an examiner's belief that two samples came from the same source. We suggest the term "proposed identification" to appropriately convey the examiner's conclusion, along with the possibility that it might be wrong. We will use this term throughout the report.

If a scientist says that the questioned and reference samples match, the immediate inference to be drawn from this (as we have explained) is that they might have come from the same source but it is also true that they might not have come from the same source. These two statements make no implication with regard to evidential weight. Weight only comes from the second stage of the paradigm which entails coming up with some impression of rarity. The identification paradigm, on the other hand, is different in that implies a statement of certainty: the two samples certainly came from the same source.

The PCAST paradigm requires that the scientist should make a categorical statement (an identification) that cannot be justified on logical grounds as we have already explained. Most scientists would be comfortable with the notion of observing that two samples *matched* but would, rightly, refuse to take the logically unsupportable step of inferring that this observation amounts to an *identification*.

5.2. Judgement

The report emphasises the value of empirical data (emphasis added):

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.* ([1], p. 6)

This denial of the importance of judgement betrays a poor understanding of the nature of forensic science. We offer a simple example.

Mr POI is the suspect for a crime who was arrested at time *T* in location *Z*. Some questioned material has been found on the clothing of Mr POI which is to be compared with reference material taken from the crime scene. Denote the observations on the two samples by y and x respectively. Whichever paradigm we follow, we are interested in the probability of finding material with observations y on the clothing of Mr POI if he had nothing to do with the crime. Ideally, of course, we would like a survey carried out near to time T and in the general region of Z and of people of a socio economic group Q that would include Mr POI. But this is, of course unrealistic. What we do have is a survey of materials on clothing carried out at some earlier time T' and at another location Z' and of a slightly different socio economic group Q'. Who is to make a judgement on the relevance of this survey data to the case at hand? We would argue that this is where the knowledge and understanding of the forensic scientist is of crucial importance.

The reality is, of course, that the perfect database never exists. The council is wrong: it is most certainly *not* the case that "only empirical evidence" is relevant. Without downplaying the impor tance of data collections, they can only inform judgement it is judgement that is paramount and informed judgement is founded in reliable knowledge.

5.3. Subjective versus Objective

PCAST give their definition of the distinction between "objectivity" and "subjectivity" p. 5 footnote 3.

Feature comparison methods may be classified as either objective or subjective. By objective feature comparison methods, we mean methods consisting of procedures that are each defined with enough standardized and quantifiable detail that they can be performed by either an automated system or human examiners exercising little or no judgment. By subjective methods, we mean methods including key proce dures that involve significant human judgment . . .

What is suggested is that many of the decisions be moved from the examiner to the procedure and/or software. The procedure or software will have been written by one or more people and the decisions about what models are used or how decisions are made are now enshrined in paper or code. Hence all the subjective judgements are now made by this person or group of people via the paper or code. Whereas this approach could be viewed as repeatable and reproducible, the objectivity is illusory.

In the US environment, subjectivity has been associated with bias and sloppy thinking, and objectivity with an absence of bias and rigorous thinking. It is worthwhile examining whence the fear of subjectivity arises. There is considerable proof that humans are susceptible to quite a number of cognitive effects many of which can affect judgement. We suspect that the fear is that these effects bias the decisions in ways that are detrimental to justice. Hence, it is bias arising from cognitive effects that is the enemy, not subjectivity.

If we return to the concept of enforced precision, we could assume that trials could be conducted on such a system and that the outputs could be calibrated. Such a system could be of low susceptibility to bias arising from cognitive effects. We suspect that these are the goals sought by PCAST. We certainly could support calibrating subjective judgements but we see little value in pretending that writing them down or coding them makes them objective.

5.4. Transposed conditional

We are concerned by the report's poor use of the notion of probability. In particular we note in the report many instances where the fallacy of the transposed conditional either occurs explicitly or is implied. We have seen that the logic of forensic inference directs us to assign a value to the probability of the observations given the truth of a proposition. The probability of the truth of a proposition is for the jury *not* the scientist. Confusion between these two different probabilities has been called the "prosecutor's fallacy" [19]. We prefer the term *transposed conditional* because, in our experience, the fallacy is regularly committed by prosecutors, defence attorneys, the judiciary and the media alike.

The fallacy is widespread, even though it can be grounds for a retrial if given in testimony by an expert witness. The document [20] that attempts to explain DNA statistics to defence attorneys in the US describes incorrectly a likelihood ratio for a mixture profile as:

4.73 quadrillion times more likely[6] to have originated from [suspect] and [victim/complainant] than from an unknown individual in the U.S. Caucasian population and [victim/complainant]." ([20], p. 52)

---

6 We are fully aware of the distinction made in statistical theory between "likelihood" and "probability". We believe that attempting to explain that distinction in this paper would cause more confusion than the worth of it. It is our experience that in courts of law the two terms are taken to be synonymous.

This is a classic example of the transposed conditional. It is a transposition of the likelihood ratio, which would be more correctly presented as follows:

> The DNA profile is 4.73 quadrillion times more likely to be obtained if the DNA had originated from the suspect and the victim/complainant rather than if it had originated from an unknown individual in the U.S. Caucasian population and the victim/complainant.

The contrast between these two statements, though apparently subtle, is profound. The first is an expression of the probability (or odds) that a particular proposition is true this, we have seen, is the probability that the jury must address, not the scientist.[7] The second considers the probability of the *observations*, given the truth of one proposition then the other, which is the appropriate domain for the expertise of the scientist. It is important to realise that the first statement is not a simple rephrasing of the second statement. Whereas the second may be a valid representation of the scientist's evaluation in a given case, the first most definitely cannot be.

Consider the following quote from the first paragraph on footwear methodology in the PCAST report ([1], p. 114):

> Footwear analysis is a process that typically involves comparing a known object, such as a shoe, to a complete or partial impression found at a crime scene, to assess whether the object is likely to be the source of the impression.

This is wrong. We state again that it is not for the scientist to present a probability for the truth of the proposition that the object was the source of the impression. The scientist addresses the probability of the outcome of the comparison *if* the object were the source of the impression: this probability forms the numerator of the likelihood ratio. Just as important, of course, is the probability of the outcome of the comparison *if* some other object were the source of the impression. The latter forms the denominator of the likelihood ratio. It is the two probabilities, taken together, that determine the evidential weight in relation to the two propositions of interest to the court.

The PCAST report sentence clearly states that the objective of the footwear analysis is to present a probability for the proposition given the observations, and not for the observations given the proposition. This is clearly a transposition of the conditional.

Similarly, the scientist is not in a position to consider the probability addressed in the following ([1], p. 65 and repeated on p. 146):

> . . . determining, based on the similarity between the features in two sets of features, whether the samples should be declared to be likely to come from the same source . . .

We have seen that is not for the scientist to consider the probability that the samples came from the same source given the observation of a "match". It is another example of the fallacy of the transposed conditional.

This confusion is systematic in the original report and we note that it continues into the addendum ([8], p. 1) (emphasis added):

> These methods seek to determine whether a questioned sample *is likely to come* from a known source based on shared features in certain types of evidence.

We have seen that this is most certainly *not* what a feature comparison should aspire to. It is not the role of the forensic scie

---

[7] In Bayesian terms, the first statement is one of posterior odds. This can be derived from the second statement either by assigning prior odds of one (which would be highly prejudicial in most criminal trials) or by making the mistake of transposing the conditional. Neither is acceptable behaviour for a scientist.

ntist to offer a probability for the proposition that a questioned sample came from a given source since this would require the scientist to take account of all of the non scientific information which properly lies within the domain of the jury.

The need for precision of language when presenting probabilities is exemplified by two quotations from the report. First, from p. 8 when talking about the interpretation of a DNA profile:

> Could a suspect's DNA profile be present within the mixture profile? And, what is the probability that such an observation might occur by chance?

As we read it, this second sentence can be taken to mean:

> What is the probability that such an observation would be made if the suspect's DNA were not present in the mixture?

Within the logical paradigm, this is a legitimate question to ask it is the probability of the observations given that one of the propositions were true.

However, later in the report we find (p. 52):

> the random match probability that is, the probability that the match occurred by chance".

There is an economy of phrasing here that obscures meaning and the reader could be forgiven for believing that the question implied by the second phrase is:

> What is the probability that the two samples had come from different sources and matched by chance?

This is a probability of a proposition (the two samples came from different sources) given the observation (a match) and would imply a transposed conditional. We are aware that the council may respond that this is not at all what they meant to which we would respond that the council should have been far more careful in its phraseology.

5.5. "Probable match"

In giving their definition of the distinction between "objectivity" and "subjectivity" p. 5 see footnote 3 the report states:

> how *to determine whether the features are sufficiently similar to be called a probable match.*

The council do not say what they mean by a "probable match" but it seems to us that it is another example of confusion between the match and identification paradigms. Following the match paradigm there is no such thing as a probable match the two samples either match or they do not.

5.6. Foundational validity and accuracy

The report distinguishes two types of scientific validity: "foundational validity" and "validity as applied". We confine ourselves to the first of these (p. 4):

> *Foundational validity* for a forensic science method requires that it be shown based on empirical studies to be *repeatable, reproducible, and accurate*, at levels that have been measured and are appropriate to the intended application. Foundational validity, then, means that a method can, *in principle*, be reliable.

Repeatability refers to the ability of the same operator with the same equipment to obtain the same (or closely similar) results when repeating analysis of the same material. Reproducibility refers to the ability of the equipment to obtain the same (or closely similar) results with different operators. As such, both are

expressions of precision, which is how close each measurement or result is to the others.

Accuracy is a measure of how close one or a set of measurements is to the true answer. This has an obvious meaning when we know or could know the true answer. We could imagine some measurement such as the weight of an object where that object has been weighed by some very advanced technique and we can accept that as the "true" weight. We wish then to consider the accuracy of some other, perhaps cheaper, technique. We could assess the accuracy of this second technique by using it to weigh the object multiple times and observing the deviation of the results from the "true" weight of the object.

For some questions in forensic science, such as "How much heroin is in this seized sample?" or "How much ethanol is in this blood sample?", the notion of the accuracy of an applied analytical technique is relevant because it is possible to assess a technique's accuracy using trials with known quantities of heroin or ethanol. However, when it comes to answering a question such as "What is the probability that there would have been a match with a suspect's shoe if it did not make the mark at the scene of crime?", then there is no sense in which there is a "true answer". The values that experts assign for such probabilities will vary depending on the specific knowledge of the experts and the nature of any databases that experts may use to inform their probabilities.

We could use a weather forecaster as an illustration. If she says that there is a 0.8 probability of a sunny day tomorrow, there can be no sense in which this is a "true" statement. Equally, if tomorrow brings rain, she is not "wrong" in any sense. Nor is she "inaccurate". A probabilistic statement of this nature may be unhelpful or misleading, in the sense that it may lead us to make a poor decision, but it cannot be either true or false.

Once we abandon the idea of a true answer for probabilities, we are left with the difficult question of what we mean by accuracy. We suggest that the report does a disservice to the important task of calibrating probabilities by a simplistic allusion to accuracy. The PCAST report says (p. 46):

> Without appropriate estimates of accuracy, an examiner's statement that two samples are similar — or even indistinguishable — is scientifically meaningless; it has no probative value, and considerable potential for prejudicial impact. Nothing — not training, personal experience nor professional practices — can substitute for adequate empirical demonstration of accuracy.

We have seen that the report is wrong here — it is not a matter of "accuracy" but of evidential weight.

5.7. The PCAST paradigm

The PCAST report proposes an approach that is fusion of the match and identification paradigms. See, from p. 45/46:

> Because the term "match" is likely to imply an inappropriately high probative value, a more neutral term should be used for an examiner's belief that two samples came from the same source. We suggest the term "proposed identification" to appropriately convey the examiner's conclusion, along with the possibility that it might be wrong. We will use this term throughout the report.

First, we have seen that the term "match", if used properly, makes no implication of probative value: it implies that the two samples might have come from the same source but also might have come from different sources. This is evidentially neutral. Second, we have seen that there is no place for the "examiner's

belief that two samples came from the same source": it is not for the scientist to assign a probability to the proposition that the two samples came from the same source.

Next we must consider what the council understand the phrase "proposed identification" to mean. Do they mean that, because it is an identification, it is a categorical opinion? Note that the qualifier "proposed" does not make the identification less than categorical — if it were probabilistic it could not be "wrong".[8] If it is not probabilistic then the scientist is to provide a categorical opinion while telling the court that he/she might be wrong! It is difficult to believe that any professional forensic scientist would be happy to be put in this position.

5.8. The scientist as a "black box"

On page 49 we find:

> For subjective methods, procedures must still be carefully defined but they involve substantial human judgment. For example, different examiners may recognize or focus on different features, may attach different importance to the same features, and may have different criteria for declaring proposed identifications. Because the procedures for feature identification, the matching rule, and frequency determinations about features are not objectively specified, the overall procedure must be treated as a kind of "black box" inside the examiner's head.

The report justifiably emphasises weaknesses of qualitative opinions. The intuitive "black box" view of the scientist will certainly have been true in many instances in the past and, indeed, in certain quarters in the present day. But for us the solution is emphatically not to continue to treat this as an acceptable state of affairs for the future. The PCAST view appears to be "it's a black box, so let's treat it like a black box". Our approach has been, and will continue, to break down intuitive mental barriers by expanding transparency, knowledge and understanding. We do not see the future forensic scientist as an *ipse dixit* machine — whatever the opinion, we expect the scientist to be able to explain it in whatever detail is necessary for the jury to comprehend the mental processes that led to it.

5.9. Black box studies

That the council intend the proposed identification to be categorical is clarified in the following from page 49 (emphasis added):

> In black box studies, many examiners are presented with many independent comparison problems — typically, involving "questioned" samples and one or more "known" samples — and *asked to declare whether the questioned samples came from the same source as one of the known samples.*[9] The researchers then determine how often examiners reach erroneous conclusions.

PCAST proposes that the error rates from such experiments would be used to assign evidential value at court.

We are strongly against the notion that the scientist should be forced into the position of giving categorical opinions in this way. Whereas, we are strongly in favour of the notion of calibrating the

---

[8] Though, of course, it would be logically incorrect because it would imply a transposed conditional.

[9] In footnote 111 the report says: "Answers may be expressed in such terms as "match/no match/inconclusive" or "identification/exclusion/inconclusive". This strengthens our belief that the council see match and identification as interchangeable".

opinions of forensic scientists under controlled conditions we see those opinions expressed in terms of statements of evidential weight. We return to the subject of calibration later.

5.10. Governance

PCAST suggests that forensic science should be governed by those, such as metrologists, from outside the profession. This speaks to the view, reinforced by a very selective reference list, that the forensic science discipline is not to be trusted with developing procedures, testing them, and self governance. We do not reject input from outside the profession: we welcome it. But our own observations are that those outside may be engaged to different extents, varying from a passing interest to years of study. They may be unduly influenced by headlines in newspapers highlighting or exaggerating deficiencies. On occasion, these same commentators from outside the profession may not recognise the limitations in their own knowledge base where it concerns specifically forensic aspects, may be reticent to consult subject matter experts from amongst practising scientists and may give well intentioned, but erroneous, advice [1,21].

6. Our view of the future

6.1. Logical inference

The recommendations of the PCAST report are founded on a conflation of two classical forensic paradigms: match and identification. These paradigms are as old as forensic science but their inadequacies and illogicalities have been comprehensively exposed over the last 50 years or so. All of us maintain, and have done so in our writings, that the future of forensic science should be founded first on the notion of logical inference and second on the notion of calibrated knowledge. The former leads to a framework of principles (which have been adopted by ENFSI) and we are disappointed that PCAST has apparently chosen to ignore, or at most pay lip service to, this fundamental change. The second is a deeper and far richer concept than the profoundly limited notion of false positive and false negative error rates: this is the notion of *calibration.*

6.2. Calibration

We are most definitely in favour of the studying of expert opinion under controlled circumstances, see for example Evett [22] but proficiency testing is far more than the counting of errors. The PCAST black box approach calls for a categorical opinion that is recorded as right or wrong but we have seen that forensic interpretation is far richer and more informative than simple yes/ no answers. In a source level proficiency test we expect the participants to respond with a statement of evidential weight in relation to one of two clearly stated propositions. Support thus expressed for a proposition that is, in fact, false is undesirable because it is misleading not "wrong". Obviously, the desirable outcome of the proficiency test is a small value for the expected weight of evidence in relation to a false proposition. But whatever the outcome, the study must be seen as a learning exercise for all participants: the pool of knowledge has grown. The notion of an error rate to be presented to courts is misconceived because it fails to recognise that the science moves on as a result of proficiency tests. The work led by Found and Rogers [23] has shown how the profession of handwriting comparison in Australia and New Zealand has grown in stature because of the culture of advancing knowledge through repeated study under controlled conditions. To repeat then, our vision is not of the black box/error rate but of continuous development through calibration and feedback of opinions.

A striking example of forensic calibration is the evolution of fingerprints evidence from the identification paradigm to the logical paradigm via mathematical modelling [24,25]. Instead of the categorical identification, we have a mathematical approach that leads to a likelihood ratio. The validation of such approaches is founded on two desiderata: we require large likelihood ratios in cases in which the prosecution proposition is true; and small likelihood ratios in cases in which the defence proposition is true. Investigation of performance in relation to these two desiderata is undertaken by considering two sets of comparisons: one set in which it is known that the two samples came from the same source; and one set in which it is known that the two samples came from different sources. There have been major advances over recent years in how the likelihood ratio distributions from such experiments may be compared and evaluated (Ramos [26], Brümmer [27] see also Robertson et al. [28] for a layman's introduction to calibration). The elegance and performance of such methods far transcends the crude PCAST notion of "false positive" and "false negative" error rates.

6.3. Knowledge and data

The PCAST report focuses on "feature comparison" methods and, as we have explained, this has meant that it is concerned with inference relating to source level propositions. At this level, the report sees data as the sole means for assigning probabilities. An important part of the role of the forensic scientist is concerned with inference with regard to activity level propositions. Consider, for example, a question of the form "what is the probability of finding this number of fragments of glass on Mr POI's jacket if he is the person who smashed the window at the crime scene?" The answer is heavily dependent on circumstantial information (how large is the window? where was the person who smashed the window standing? was any implement used? how much time elapsed between the breaking of the window and the seizure of the jacket from Mr POI? etc.) and the variation in this between cases is vast. There is no single database to inform such probabilities. The scientist will, it is hoped, be thoroughly familiar with all of the published literature on glass transfer in crime cases [29] and may, if resources permit, carry out experiments that reproduce the current case circumstances. The knowledge and judgement of other scientists who have encountered similar questions is also relevant. We agree with PCAST that length of experience is not a measure of reliability of scientific opinion: the foundation is reliable knowledge. Too little effort has been devoted within the forensic sphere thus far to the harnessing of knowledge through knowledge based systems but see [29] for examples of how such a system was created for glass evidence interpretation.

We do not deny the importance of data collections but the view that data may replace judgement is misconceived. A data collection should be used to inform reliable knowledge not replace it.

We have explained that our view of the scientist is the antithesis of the PCAST "black box" automaton. Although there is a need for data, PCAST are mistaken in seeing it as the be all and end all: qualitative judgement will always be at the centre of forensic science evidence evaluation. We reject the PCAST vision of the scientist who gives a categorical opinion and a statement about the probability that the opinion is wrong. We see the model scientist as deeply knowledgeable about her domain of expertise and able to rationalise the opinion in terms that the jury will understand. The principles have been expressed elsewhere [11] as balance, logic, robustness and transparency. There is no place for the black box. We agree that the scientist should be able to provide the court with evidence of performance under controlled conditions. Found and Rogers [23] have provided a model for handwriting comparison

and we see such approaches as extending into other areas: the emphasis is on calibration of probabilistic assessments.

7. Conclusion

The 44th US president's request was "to consider whether there are additional steps that could usefully be taken on the scientific side to strengthen the forensic science disciplines and ensure the validity of forensic evidence used in the Nation's legal system" ([1], p. 1). We suggest that the report has very little emphasis on positive steps and does much to reinforce poor thinking and terminology.

Our own view of the future of forensic science is based on the principle that forensic inference should be founded on a logical framework for reasoning in the face of uncertainty. That framework is provided by probability theory coupled with the recognition that probability is necessarily subjective and conditioned by knowledge and judgement. It follows that our view of the forensic scientist is a knowledgeable, logical and reasonable person. Whereas data collections are valuable they should be viewed within the context of reliable knowledge. The overarching paradigm of reliable knowledge should be founded on the notion of knowledge management, including comprehensive systems for the calibration of expert opinion.

References

[1] President's Council of Advisors on Science and Technology, Report to the president Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods, Washington DC, 2016. https://obamawhite-house.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_foren-sic_science_report_final.pdf.

[2] Federal Bureau of Investigation—FBI, Comments on: President's Council of Advisors on Science and Technology Report to the President on Forensic Science in Federal Criminal Courts: Ensuring Scientific Validity of Pattern Comparison Methods. September 20, 2016. www.fbi.gov/file-repository/fbi-pcast-response.pdf/view.

[3] National District Attorneys Association—NDAA, Report Entitled Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods. November 16, 2016. http://www.ciclt.net/ul/ndaajustice/PCAST/NDAAPCASTResponseFINAL.pdf.

[4] G.S. Morrison, D.H. Kaye, D.J. Balding, D. Taylor, P. Dawid, C.G.G. Aitken, S. Gittelson, G. Zadora, B. Robertson, S. Willis, S. Pope, M. Neil, K.A. Martire, A. Hepler, R.D. Gill, A. Jamieson, J. de Zoete, R.B. Ostrum, A. Caliebe, A comment on the PCAST report: skip the match/non-match stage, Forensic Sci. Int. 272 (2017) e7–e9, doi:http://dx.doi.org/10.1016/j.forsciint.2016.10.018.

[5] Association of Firearm and Tool Mark Examiners AFTE, Response to PCAST Report on Forensic Science. October 31, 2016. https://afte.org/uploads/documents/AFTE_PCAST_Response.pdf.

[6] Bureau of Alcohol Tobacco Firearms and Explosives—ATF, ATF Response to the President's Council of Advisors on Science and Technology Report. September 21, 2016. https://www.theiai.org/president/20160921_ATF_PCAST_Response.pdf.

[7] The International Association for Identification (IAI), IAI Response to the President's Council of Advisors on Science and Technology Report, 2016. https://www.theiai.org/president/IAI_PCAST_Response.pdf.

[8] President's Council of Advisors on Science and Technology, An addendum to the PCAST report on forensic science in criminal courts, Washington DC, 2017. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensics_addendum_finalv2.pdf.

[9] C.G.G. Aitken, F. Taroni, Statistics and the Evaluation of Evidence for Forensic Scientists, 2nd ed, John Wiley & Sons Ltd., Chichester, 2004.

[10] C. Aitken, P., Roberts, G. Jackson, Fundamentals of Probability and Statistical Evidence in Criminal Proceedings, London, 2011. http://www.rss.org.uk/uploadedfiles/userfiles/files/Aitken-Roberts-Jackson-Practitioner-Guide-1-WEB.pdf.

[11] Expressing evaluative opinions: a position statement, Sci. Justice 51 (1) (2011) 1–2, doi:http://dx.doi.org/10.1016/j.scijus.2011.01.002.

[12] R. Cook, I.W. Evett, G. Jackson, P.J. Jones, J.A. Lambert, A model for case assessment and interpretation, Sci. Justice 38 (3) (1998) 151–156, doi:http://dx.doi.org/10.1016/S1355-0306(98)72099-4.

[13] R. Cook, I.W. Evett, G. Jackson, P.J. Jones, J.A. Lambert, A hierarchy of propositions: deciding which level to address in casework, Sci. Justice 38 (4) (1998) 231–240, doi:http://dx.doi.org/10.1016/S1355-0306(98)72117-3.

[14] R. Cook, I.W. Evett, G. Jackson, P.J. Jones, J.A. Lambert, Case pre-assessment and review in a two-way transfer case, Sci. Justice 39 (2) (1999) 103–111, doi:http://dx.doi.org/10.1016/S1355-0306(99)72028-9.

[15] P.L. Kirk, The ontogeny of criminalistics, J. Crim. Law Criminol. Police Sci. 54 (1963) 235–238.

[16] D.A. Stoney, What made us ever think we could individualize using statistics, J. Forensic Sci. Soc. 31 (2) (1991) 197–199, doi:http://dx.doi.org/10.1016/S0015-7368(91)73138-1.

[17] A. Biedermann, S. Bozza, F. Taroni, Decision theoretic properties of forensic identification: underlying logic and argumentative implications, Forensic Sci. Int. 177 (2–3) (2008) 120–132, doi:http://dx.doi.org/10.1016/j.forsciint.2007.11.008.

[18] A. Biedermann, S. Bozza, F. Taroni, The decisionalization of individualization, Forensic Sci. Int. 266 (2016) 29–38, doi:http://dx.doi.org/10.1016/j.forsciint.2016.04.029.

[19] W.C. Thompson, E.L. Schumann, Interpretation of statistical evidence in criminal trials: the prosecutor's fallacy and the defence attorney's fallacy, Law Hum. Behav. 11 (3) (1987) 167–187, doi:http://dx.doi.org/10.1007/BF01044641.

[20] E.H. Holder, M.L. Leary, J.H. Laub, DNA for the Defense Bar, U.S. Department of Justice Office of Justice Programs, Washington, DC, 2012.

[21] National Research Council - Committee on DNA Technology in Forensic Science, DNA Technology in Forensic Science, National Academy Press, Washington, D.C. 1992.

[22] I. Evett, The logical foundations of forensic science: towards reliable knowledge, Philos. Trans. R. Soc. Lond. B Biol. Sci. 370 (1674) (2015), doi:http://dx.doi.org/10.1098/rstb.2014.0263.

[23] B. Found, D. Rogers, The initial profiling trial of a program to characterize forensic handwriting examiners' skill, J. Am. Society of Questioned Document Examiners 6 (2) (2003) 72–81.

[24] C. Champod, C.J. Lennard, P.A. Margot, M. Stoilovic, Fingerprints and other Ridge Skin Impressions, CRC Press, Boca Raton, 2016.

[25] C. Neumann, I.W. Evett, J. Skerrett, Quantifying the weight of evidence from a forensic fingerprint comparison: a new paradigm, J. Roy. Stat. Soc. Ser. A. (Stat. Soc.) 175 (Part 2) (2012).

[26] D. Ramos, J. Gonzalez-Rodriguez, G. Zadora, C. Aitken, Information-theoretical assessment of the performance of likelihood ratio computation methods, J. Forensic Sci. 58 (6) (2013) 1503–1518, doi:http://dx.doi.org/10.1111/1556-4029.12233.

[27] N. Brümmer, J. du Preez, Application-independant evaluation of speaker detection, Comput. Speech Language 20 (2006) 230–275, doi:http://dx.doi.org/10.1016/j.csl.2005.08.001.

[28] G.A. Robertson, C.E.H. Vignaux, Interpreting Evidence—Evaluating Forensic Science in the Courtroom, 2nd ed., John Wiley & Sons, Ltd., Chichester, 2016.

[29] J.M. Curran, T.N. Hicks, J.S. Buckleton, Forensic Interpretation of Glass Evidence, CRC Press LLC, Boca Raton, 2000.

Re: Probability Paper

Great, thanks!

On Oct 18, 2017, at 7:29 AM, Laporte, Gerald (OJP) <(b) (6) > wrote:

As an FYI, I contacted Alex and asked if he would be interested in speaking at IPTES. I'll keep you posted.

Best Regards,

Gerry

Gerry LaPorte

Director

National In titute of Ju tice

Office of Investigative & Forensic Sciences

810 7th Street NW

Wa hington, DC 20531

Office Phone: (b) (6)

Mobile Phone: (b) (6)

Email: (b) (6)

For a li t of R&D NEEDS, plea e vi it http //nij.gov/topic /foren ic /Document /fy15 foren ic twg table pdf

For a comprehen ive li t of Foren ic Science R&D Project , plea e vi it http://www.nij.gov/topics/forensics/Pages/research-development-projects.aspx.

For information related to Forensic DNA and DNA grant programs, please visit

http //www nij.gov/topic /foren ic /evidence/dna/Page /welcome a p

For information relating to SEXUAL ASSAULT, please visit

http://www.nij.gov/topics/crime/rape-sexual-violence/pages/welcome.aspx

From: Hunt, Ted (ODAG)
Sent: Tuesday, October 17, 2017 3:49 PM
To: Laporte, Gerald (OJP) <(b) (6)
Subject: RE: Probability Paper

Yes, have read this one too, and have actually already quoted parts of it in some remarks I've given.

I think having Biederman present on this topic would be great.

I spoke about convergent validity briefly at the NAS as a more flexible and realistic way to use ALL relevant and available information   more broadly con i tent with normative principle of cience   and didn't get a lot of pu hback Jennifer Mnookin even backed that concept up a bit and provided some support for my point in front of an otherwise semi-hostile crowd. I've been talking to others about it as well and think that such an approach (b) (5)

Thanks for the worksheet-will review.

Ted

243f4028-9c81-400d-888e-92780c0f23ab
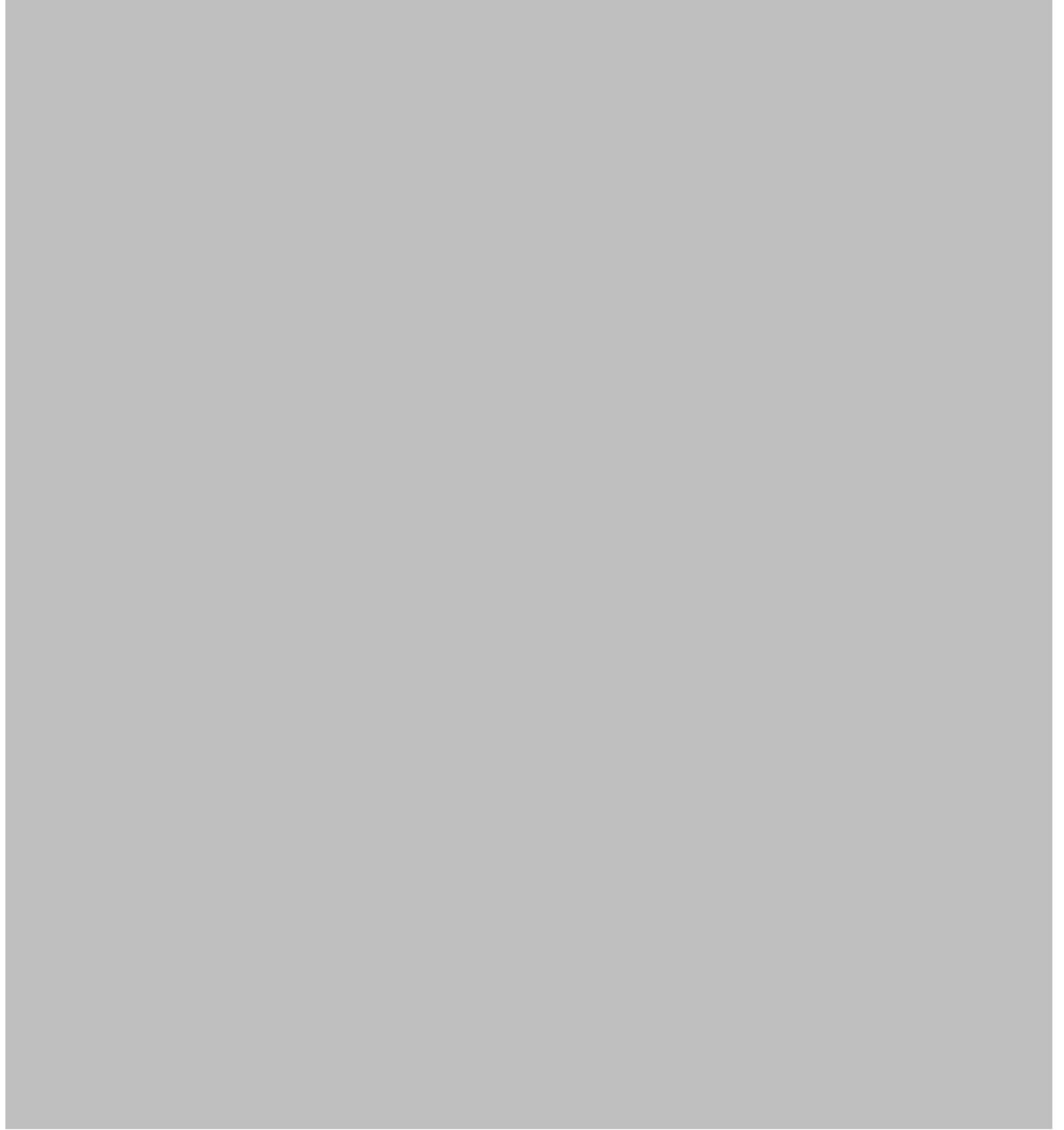
DC DFS Science Advisory Board Statement in Response to PCAST Report - DNA

From: "McGrath, Jonathan (OJP)" <(b) (6) >
To: "Hunt, Ted (ODAG)" <(b) (6)   "Antell, Kira M. (OLP)" <(b) (6)
Cc "Laporte, Gerald (OJP)" (b) (6)
Date: Tue, 15 May 2018 09:33:55 -0400
Attachments: DC DFS - SAB Statement in Response to PCAST Report_DNA_0.pdf (421.87 kB)

FYI, From 05/11/2018
https://dfs.dc.gov/release/sab-statement-response-pcast-report-dna

Jonathan G. McGrath, Ph.D., M.S.F.S.
Senior Policy Analyst
National Institute of Justice
Office of Investigative and Forensic Sciences
810 Seventh Street, NW
Washington. DC 20531
Cell: (b) (6)
Office: (b) (6)
(b) (6)

Department of Forensic Sciences Science Advisory Board's Statement with regard to the PCAST Report

Introduction

On September 20, 2016, the US President's Council of Advisors on Science and Technology (PCAST) published a report on *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods* in response to the President's question as to whether there are additional steps that could help ensure the validity of forensic evidence in the Nation's legal system.

As appropriate to the disciplines offered by the Department of Forensic Sciences, the Advisory Board will address the disciplines of Forensic Biology (DNA), Latent Fingerprint Analysis, and Firearms Analysis.   The Board has decided to address these disciplines separately, beginning with Forensic Biology. The other disciplines will be addressed in the next few meetings.

DNA

According to published reviews of this report (e.g., [1-4]), the PCAST report presents a flawed paradigm for forensic evaluation, misapplies statistics and the notion of probability, ignores existing data and literature in forensic science, and, as a result, state that the PCAST report is scientifically unsound.

The PCAST report concludes that the DNA analysis of single-source specimen and simple mixtures of two contributors is a foundationally valid and reliable method, yet raises several concerns about the interpretation of complex DNA mixtures (pp. 75-83). Regarding the latter, the report concludes (page 82):[1]

> *Objective analysis of complex DNA mixtures with probabilistic genotyping software is relatively new and promising approach. Empirical evidence is required to establish the foundational validity of each such method within specified ranges. 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. The range in which foundational validity has been established is likely to grow as adequate evidence for more complex mixtures is obtained and published.*

We, the Science Advisory Board, state that at the time of this writing, the range in which foundational validity has been established for the interpretation of complex mixtures at DFS using

---

[1] Note that an addendum to the report that appeared in January 2017 reached a slightly different conclusion (page 8):

> *PCAST found that empirical testing of PG* [probabilistic genotyping] *had largely been limited to a narrow range of parameters (number and ratios of contributors). We judged that the available literature supported the validity and reliability of PG for samples with three contributors where the person of interest comprises at least 20% of the sample. Beyond this approximate range (i.e. with a larger number of contributors or where the person of interest makes a lower than 20% contribution to the sample), however, there has been little empirical validation.*

938d4f34-ef30-4b78-ae9a-f566c8cb2d66   20220314-12831

probabilistic genotyping[2] extends from DNA mixtures of 2 individuals up to DNA mixtures of 5 individuals. The PCAST notion of a lower limit percentage of the minor contributor as a criterion for deciding whether a DNA profile is interpretable or uninterpretable is scientifically unsound. The scientific criterion for making this decision is the quantity of information in the electropherogram(s) for a particular contributor.[3] DFS has a valid pre-evaluation phase in place for making this decision.

More specifically, an internal validation study conducted by the DNA analysts at DFS[4] consisting of over 10,000 comparisons to 100 DNA mixtures ranging from 2 contributors to 5 contributors has addressed each of the PCAST concerns listed below (PCAST, pp. 79-80).

> These probabilistic genotyping software programs clearly represent a major improvement over purely subjective interpretation. However, they still require careful scrutiny to determine (1) whether the methods are scientifically valid, including defining the limitations on their reliability (that is, the circumstances in which they may yield unreliable results) and (2) whether the software correctly implements the methods. This is particularly important because the programs employ different mathematical algorithms and can yield different results for the same mixture profile. (PCAST, page 79)

The internal validation study conducted at DFS demonstrates that the interpretation of complex mixtures using STRmix™ version 2.4 in conjunction with GlobalFiler™ PCR Amplification Kit and 3500/3500xL Genetic Analyzer is scientifically valid for mixtures of 2 to 5 individuals.

To test the correctness of the software's implementation of the method, the DFS internal validation study reproduced the likelihood ratio values for each locus of a single-source profile in quadruple, once for each of four allele frequency databases. These results confirm that the software correctly implements the method.

> Appropriate evaluation of the proposed methods should consist of studies by multiple groups, not associated with the software developers, that investigate the performance and define the limitations of programs by testing them on a wide range of mixtures with different properties. In particular, it is important to address the following issues:
>
> (1) How well does the method perform as a function of the number of contributors to the mixture? How well does it perform when the number of contributors to the mixture is unknown? (PCAST, page 79)

---

[2] Note that probabilistic genotyping does not identify contributors with 100% certainty. Instead it applies mathematical models and probability theory to assign probabilities to the observed peak heights given different sets of potential contributors. The conclusion is therefore probabilistic, taking the form of a likelihood ratio.

[3] The quantity of information in the electropherogram(s) for a particular contributor depends on the quantity of data and the information known about the mixture.

[4] The DFS internal validation study strictly follows the FBI approved SWGDAM Guidelines for the Validation of Probabilistic Genotyping Systems available at
https://docs.wixstatic.com/ugd/4344b0_22776006b67c4a32a5ffc04fe3b56515.pdf (accessed January 2, 2018).
It was approved by the Technical Leader on 1/7/2016 for the Idenfiler Plus PCR Amplification kit and on 2/24/2017 for the GlobalFiler PCR Amplification kit. A summary of the results is available at
https://dfs.dc.gov/page/fbu-validation-studiesperformance-checks (accessed January 5, 2018), and these results have been published in a peer-reviewed journal as part of a larger compilation of results from STRmix™ internal validation studies [5].

The DFS internal validation study tested the performance of the method for 40 mixtures with 2 contributors, and 20 mixtures each for 3, 4 and 5 contributors. These mixtures varied in DNA quantity and mixture proportions to represent the typical profiles[5] encountered by the laboratory. The method correctly and reliably produced the expected results for each of the different number of contributors tested.

In addition, the results of the FBI internal validation study on the performance of STRmix™ version 2.3.06 contains a total of 290 mixtures with 2, 3, 4, and 5 contributors, for each of which the software proved to be appropriately sensitive and specific [6].

In casework, the number of contributors is always unknown (e.g., [7]). The DNA analyst assigns the number of contributors based on the number of peaks and the peak height information in the electropherogram.

To test the effect of an incorrect assignment of the number of contributors, the DFS internal validation study included the following tests:
- 10 mixtures each with 1, 2, 3 and 4 contributors were incorrectly interpreted as having 2, 3, 4 and 5 contributors, respectively; and
- 3 mixtures each with 2 and 3 contributors, and 4 mixtures each with 4 and 5 contributors were incorrectly interpreted as having 1, 2, 3 and 4 contributors, respectively

Each mixture was then evaluated against each of the known contributors and against 134 known non-contributors.

Overestimation of the number of contributors correctly produced likelihood ratios greater than 1 for the known contributors. It produced a few likelihood ratios greater than 1 for known non-contributors, but their order of magnitude is much lower than the likelihood ratios produced for the known contributors.[6]

Underestimation of the number of contributors did not have any influence on the likelihood ratios for the known major and minor contributors. It correctly produced lower likelihood ratios for the known trace contributors.

The FBI internal validation study included similar tests on an additional 30 mixtures which produced the same expected trends as the DFS internal validation results [6].

> (2) *How does the method perform as a function of the number of alleles shared among individuals in the mixture? Relatedly, how does it perform when the mixtures include related individuals?* (PCAST, page 79)

The DFS internal validation study performed sensitivity and specificity studies on mixtures with different amounts of alleles shared among the contributors across the loci. These tests correctly and reliably produced the expected results. Given that continuous probabilistic genotyping models take allele sharing into account in their peak height models, this method can handle the entire range of possible allele sharing among the DNA's contributors.

---

[5] This includes partial profiles.
[6] Note that DFS has defined likelihood ratios between 1 and 100 as being "uninformative" based on the results of their internal validation study.

3

With regard to related individuals, the FBI internal validation study tested the method on mixtures with 3 contributors that consisted of 2 parents and 1 child. This type of mixture entails a risk of an underestimation of the number of contributors if only the number of peaks is counted and peak height information is disregarded. An underestimation of the number of contributors has no impact on the likelihood ratios of the known major and minor contributors, yet lowers the likelihood ratio for the known trace contributor.

> *(3) How well does the method perform—and how does accuracy degrade—as a function of the absolute and relative amounts of DNA from the various contributors? For example, it can be difficult to determine whether a small peak in the mixture profile represents a true allele from a minor contributor or a stutter peak from a nearby allele from a different contributor. (Notably, this issue underlies a current case that has received considerable attention.)* (PCAST, page 79)

The DFS internal validation study included sensitivity and specificity studies on DNA mixtures of varying amounts of DNA. These ranged from an average peak height of about 20 rfu to >25,000 rfu (saturation). The mixture ratios ranged from 25:1 to 1:1 for two person mixtures, with the full range in between for three, four and five person mixtures. As expected for all methods, this method correctly and reliably produced uninformative results for contributors with very low template. For contributors with higher template, this method correctly and reliably produced high likelihood ratios greater than 1 for known contributors, and low likelihood ratios less than 1 for known non-contributors, which clearly separated the results of the known contributors from the results of the known non-contributors. On the high-template end, the method correctly interprets the profile qualitatively for saturated profiles.

Probabilistic genotyping does not determine whether a small peak in the mixture profile represents a true allele from a minor contributor or a stutter peak from a nearby allele from a different contributor. It takes all reasonable possibilities into account, and assigns probabilities to the observations given each of the possibilities. In other words, it assigns weights to the different possibilities, and must therefore not choose between the category of a true allele and the category of a stutter peak.

> *(4) Under what circumstances—and why—does the method produce results (random inclusion probabilities) that differ substantially from those produced by other methods?* (PCAST, page 80)

The method used by DFS uses a fully continuous probabilistic genotyping model to produce likelihood ratios which express the relative support the DNA typing results provide for one proposition with regard to an alternative proposition. A likelihood ratio is a different statistical quantity from a random match probability or a combined probability of inclusion, and will therefore produce different numerical results than either of the latter quantities. In addition, a fully continuous model can produce likelihood ratios that are different from likelihood ratios obtained from a binary model or a semi-continuous model: the reason for these differences is that a fully continuous model takes into account all of the available peak height information above the analytical threshold in the electropherogram, whereas binary and semi-continuous models only take a very limited amount of this information into account (e.g., comparing observed peak heights to a stochastic threshold), if at all. Hence a fully continuous model will produce results different from those produced by binary and semi-continuous models in circumstances where the electropherogram contains peak height information that is taken into account by the fully

continuous model and not taken into account by the binary and semi-continuous models. Taking into account more information makes this method produce higher likelihood ratios in support of the DNA contribution of known contributors and lower likelihood ratios (or exclusions) in support of no DNA contribution of known non-contributors (e.g., [8-11]). This is the expected performance for all likelihood ratio methods.

> *Most importantly, current 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 TrueAllele) appear to be reliable within a certain range, based on the available evidence and the inherent difficulty of the problem. Specifically, these methods appear to be reliable for three-person mixtures 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 level required for the method.* (PCAST, page 80)

The DFS internal validation study has shown that STRmix™ version 2.4 is reliable for DNA mixtures with 2, 3, 4 and 5 contributors. Independently, the FBI internal validation study has shown that STRmix™ version 2.3.06 is reliable for DNA mixtures with 2, 3, 4, and 5 contributors [6]. The results of additional internal validation studies of STRmix™ conducted by other laboratories can be found at **https://johnbuckleton.wordpress.com/strmix/strmix-validations/** (accessed October 24, 2017).

Again, we note that <u>the PCAST notion of a lower limit percentage of the minor contributor as a criterion for deciding whether a DNA profile is interpretable or uninterpretable is scientifically unsound. The scientific criterion for making this decision is the quantity of information in the electropherogram(s) for a particular contributor (e.g.</u> [12]).

References:

[1] I.W. Evett, C.E.H. Berger, J. Buckleton, C. Champod, G. Jackson, Finding the way forward for forensic science in the US - A commentary on the PCAST report, Forensic Science International 278 (2017) 16-23.
[2] G.S. Morrison, D.H. Kaye, D.J. Balding, D. Taylor, A.P. Dawid, C.G.G. Aitken, S. Gittelson, G. Zadora, B. Robertson, S. Willis, S. Pope, M. Neil, K.A. Martire, A. Hepler, R.D. Gill, A. Jamieson, J.d. Zoete, R.B. Ostrum, A. Caliebe, A comment on the PCAST report: Skip the 'match'/'non-match' stage, Forensic Science International 272 (2017) e7-e9.
[3] B. Budowle, Response to the PCAST report, 2017.
[4] J. Buckleton, U.S. v. Benito Valdez, 2017.
[5] J.-A. Bright, R. Richards, M. Kruijver, H. Kelly, C. McGovern, A. Magee, A. McWhorter, A. Ciecko, B. Peck, C. Baumgartner, C. Buettner, S. McWilliams, C. McKenna, C. Gallacher, B. Mallinder, D. Wright, D. Johnson, D. Catella, E. Lien, C. O'Connor, G. Duncan, J. Bundy, J. Echard, J. Lowe, J. Stewart, K. Corrado, S. Gentile, M. Kaplan, M. Hassler, N. McDonald, P. Hulme, R.H. Oefelein, S. Montpetit, M. Strong, S. Noël, S. Malsom, S. Myers, S. Welti, T. Moretti, T. McMahon, T. Grill, T. Kalafut, M.M.
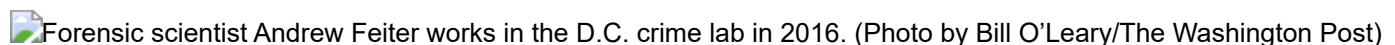
Greer-Ritzheimer, V. Beamer, D. Taylor, J. Buckleton, Internal validation of STRmix™ - A multi laboratory response to PCAST, Forensic Science International: Genetics, in press (2018).

[6] T.R. Moretti, R.S. Just, S.C. Kehl, L.E. Willis, J. Buckleton, J.-A. Bright, D. Taylor, A.J. Onorato, Internal validation of STRmix™ for the interpretation of single source and mixed DNA profiles, Forensic Science International: Genetics 29 (2017) 126-144.

[7] J.-A. Bright, D. Taylor, C. McGovern, S. Cooper, L. Russell, D. Abarno, J. Buckleton, Developmental validation of STRmix™, expert software for the interpretation of forensic DNA profiles, Forensic Science International: Genetics 23 (2016) 226-239.

[8] T.W. Bille, S.M. Weitz, M.D. Coble, J. Buckleton, J.-A. Bright, Comparison of the performance of different models for the interpretation of low level mixed DNA profiles, Electrophoresis 35 (2014) 3125-3133.

[9] H. Kelly, J.-A. Bright, J. Buckleton, J.M. Curran, A comparison of statistical models for the analysis of complex forensic DNA profiles, Science & Justice 54 (2014) 66-70.

[10] J.-A. Bright, I.W. Evett, D. Taylor, J.M. Curran, J. Buckleton, A series of recommended tests when validating probabilistic DNA profile interpretation software, Forensic Science International: Genetics 14 (2015) 125-131.

[11] D. Taylor, Using continuous DNA interpretation methods to revisit likelihood ratio behaviour, Forensic Science International: Genetics 11 (2014) 144-153.

[12] J.-A. Bright, D. Taylor, S. Gittelson, J. Buckleton, The paradigm shift in DNA profile interpretation, Forensic Science International: Genetics 31 (2017) e24-e32.

Two FBI officials say the state of forensics is fine. Here's why they're wrong. - The Washington Post

From: "Laporte, Gerald (OJP)" <(b) (6)   >
To: "Hunt, Ted (ODAG)" <(b) (6)   "Antell, Kira M. (OLP)" <(b) (6)
Cc: "McGrath, Jonathan (OJP)" (b) (6)
Date: Thu, 07 Jun 2018 08:20:23 -0400

https://www.washingtonpost.com/news/the-watch/wp/2018/06/06/two-fbi-officials-say-the-state-of-forensics-is-fine-heres-why-theyre-wrong/?utm_term=.98d15248ddf8

Two FBI officials say the state of forensics is fine. Here's why they're wrong.

Forensic scientist Andrew Feiter works in the D.C. crime lab in 2016. (Photo by Bill O'Leary/The Washington Post)

Last October, the Fordham University Law School hosted a symposium on forensics and how they're used in the courtyard. Presenters then wrote up their presentations for publication in the March 2018 Fordham Law Review.

I want to addre an e ay by Alice R I enberg and Cary T Oien, titled "Scientific E cellence in the Foren ic Science Community" I enberg i the deputy a i tant director of the Laboratory Divi ion at the FBI Oien i a enior foren ic cienti t at the FBI A you might imagine, I have ome i ue with their e ay

Generally, the piece makes the case that the alarm about forensics is much ado about very little. It's also a defense of how forensics is used by the federal government, and by the FBI specifically.

So let's jump right in:

> The practice of foren ic cience ha e i ted for centurie   Each year, hundred of thou and of ca e are clo ed, u pect cleared, and offender convicted through routine, accurate, and reliable foren ic te ting Foren ic te ting include chemical analy i to determine the nature of eized drug ; e amination performed on phy ical material uch a fiber , gla , and pent bullet ca ing ; and e amination of biological material uch a DNA Te t performed for each of the e e amination , regardle   of the material e amined, are trictly pre cribed by laboratory policie , upported by peer reviewed re earch, and lead to accurate and reliable re ult

The first fingerprint case in the United States was in 1911. Some early forebears of what we call "forensics" today were used in some European courts going back to the 19th century, but most fields of modern forensics were developed and first used in the Progressive Era. So while the U.S. court system has been around for centuries (just long enough to use the plural), forensics as we know it today is about 100 years old.

As for "peer-reviewed research," it depends on how you define the term. Some forensics journals claim to be peer-reviewed. But that review is generally done by other forensic specialists. The truth is, most fields of forensics weren't subjected to rigorous scientific testing until the last 10 to 15 years. That testing has shown that, in many fields, expert

witnesses have been giving testimony that is either completely unsupported by science or they've been significantly exaggerating the significance of their findings.

"Accurate and reliable results?" Well, that depends. The most problematic fields of forensics are those known as the pattern matching fields. This includes any specialty that requires an analyst to look at one sample and "match" it to another. Think hair and carpet-fiber analysis, bite-mark analysis, shoe-print and tire-tread analysis, blood-spatter analysis and fingerprint matching. The degree to which these fields are problematic vary quite a bit (bite-mark matching is probably on the least reliable end of the spectrum, with fingerprint matching at the other end), but all at their core are subjective. (Fingerprint matching breaks down the moment you start looking at partial prints.) That means they cannot calculate a margin for error. It means analysts will often disagree about conclusions, sometimes in ways that directly contradict one another. And by definition, any method of analysis that results in experts coming to contradictory conclusions about the same piece of evidence can't possibly be accurate (one of them is obviously wrong) or reliable.

This means that these fields aren't science. That doesn't mean they have no evidentiary value at all. But it does mean that analysts need to be extremely careful about how they present this sort of evidence to juries. The language they use needs to be standardized and then explained to juries, so that the amount of emphasis the jury puts on it is based on the evidence's actual significance and not other factors, such as the charisma or persuasiveness of the analyst. This hasn't been happening.

> A casual reader of recent media reports might be led to believe that forensic science lacks any scientific credibility. However, this narrative is completely inaccurate and at odds with the scientific excellence that exists throughout the forensic science community. Forensic disciplines are grounded in diverse sciences such as chemistry, biology, and physics, and every forensic discipline practiced in an accredited forensic laboratory must demonstrate that it is reliable, accurate, and fit for its intended use.

There are a lot of subjective and undefined terms in this paragraph. To date, a number of commissions, panels and other bodies staffed with actual scientists have reviewed some of the most common fields of forensics and found them lacking in scientific merit. These include the [National Academy of Sciences](#), the [Texas Forensic Science Commission](#), the [President's Council of Advisors on Science and Technology](#) and the [National Commission on Forensic Science](#).

To say that several forensic disciplines are "grounded in diverse sciences" tells us little about whether the actual methods of analysis that the practitioners use to reach their conclusions are scientifically sound and reliable. One important test of scientific merit, for example, is repeatability. If a method of forensic analysis is scientifically reliable, two trained analysts using the same experiment with the same evidence should reliably get the same result. In a scientifically proven field such as DNA testing, two trained scientists testing the same blood, hair, semen, skin cells or other biological material will reliably produce the same DNA profile. The same can't always be said of blood-spatter analysts, tool-mark analysts or bite-mark analysts.

One could say, for example, that blood-spatter analysis is "grounded in" physics — the physics of what a liquid such as blood does as it is forced from the body, travels through the air and strikes a wall or floor. But simply referencing some bit of scientific knowledge doesn't mean that the end product of your analysis will be scientifically reliable. Astrologers, for

example, utilize some principles and knowledge from astronomy, which is a legitimate field of science. This doesn't mean that astrologers' predictions are "grounded in science." Just last week, the New York Times ran an editorial on this, based on a two-part investigation into a conviction based on blood-spatter evidence, written by Pamela Colloff.

> Joe Bryan was convicted on the word of a detective named Robert Thorman, who testified before the jury as an expert in what is known as bloodstain-pattern analysis . . . People like Detective Thorman got certified as bloodstain-pattern analysts after taking a weeklong course that now costs as little as a few hundred dollars.
>
> Pamela Colloff, who wrote the articles on the Times Magazine/ProPublica investigation of Joe Bryan's case, enrolled in one of these courses, where the instructor told her, "We're not really going to focus on the math and physics; it just kind of bogs things down." Ms. Colloff passed the final exam, as did everyone in the class.
>
> Thanks in part to such dubious standards, the interpretation of bloodstain evidence has become notoriously ambiguous. The same patterns can, like a Rorschach test, be read in very different ways; some trials feature two bloodstain "experts," one on each side, who testify to opposite conclusions. A 2009 report by the National Academy of Sciences found that "the opinions of bloodstain-pattern analysts are more subjective than scientific," and, "The uncertainties associated with bloodstain pattern analysis are enormous."
>
> And yet judges in many states have accepted these experts' testimony as scientifically valid — not because of any concrete evidence that it is, but because other courts have accepted it before.

This of course is how courts routinely adjudicate challenges to the scientific validity of expert testimony. They look to see what other courts have done. This means that once a scientifically dubious field enters the criminal justice system, it becomes really difficult to reverse the process.

Let's get back to Isenberg and Oien.

> Accreditation and quality assurance systems assure the public that accredited organizations are competent and their results can be relied upon. Many groups — such as the National Commission on Forensic Science, the National Academy of Sciences, the President's Council of Advisors on Science and Technology (PCAST), and the Department of Justice (DOJ)— recognize that accreditation is critically important. In fact, in December 2015, the Attorney General directed that all DOJ forensic laboratories must obtain or maintain accreditation.

About that 2015 directive. It has a pretty big loophole. It only states that federal prosecutors use accredited crime labs "when practicable." As Frontline reported at the time, if finding an accredited lab would result in too much of a delay or too great an expense, federal prosecutors can continue to use unaccredited labs.

> Accreditation is an external assessment of a laboratory's technical competence to perform specific types of testing. Accreditation demonstrates that a laboratory is performing its work correctly and consistent with appropriate standards. To maintain this recognition, a laboratory is periodically reevaluated to ensure its ongoing compliance

All else being equal, some accreditation is certainly better than no accreditation. But accreditation is merely a baseline. It doesn't ensure competency. It doesn't ensure that a crime lab is operating within the parameters of sound science. In the end, an accreditation is only as valuable as the rigor, reputation and thoroughness of the accrediting organization. And as I'll get to in a moment, the history of these groups leaves a lot to be desired.

Accrediting bodies also typically exist within the existing culture of forensics. Typically, when we talk about accreditation, we aren't talking about outside scientists who are assessing the scientific credibility and rigor of a crime lab and its analysts. These accrediting organizations usually begin their evaluations of crime labs from the perspective that even the more subjective and controversial fields of forensics, when applied according to the prevailing standards of those fields, are legitimate and reliable. In other words, accrediting organizations evaluate subjective fields such as blood-spatter, shoe-print, or hair-fiber analysis on their own respective terms. If you're meeting the minimum standards put forth by the blood-spatter, shoe-print, or hair-fiber analysis community, you're probably going to be accredited. Accreditation does *not* mean that these fields *as a whole* are scientific, credible or reliable.

According to the Bureau of Justice Statistics, 88 percent of the 409 publicly-funded forensic crime labs in the United States are accredited. Unaccredited labs are often very small — less than ten people — and offer services in a limited number of disciplines. In addition to forensic laboratories, laboratories performing other types of tests are accredited according to the same international standard. This includes environmental labs checking for levels of lead in groundwater, chemistry labs preparing chemicals for consumer use, or food labs ensuring the safety of our food supply.

This is a strange comparison. The fact that a crime lab is subjected to similar standards as labs that perform other types of tests says nothing about the validity of the crime lab's analysis.

But let's look at that 88 percent figure. It *seems* impressive. But it seems impressive only until you start to look at the seemingly endless parade of crime lab scandals we've seen over the past 15 to 20 years, and you begin to notice how many of those labs . . . *were already accredited*.

Here's a quick list of such cases, which isn't remotely comprehensive:

- Last October, Massachusetts officials fired the head of the Office of Alcohol Testing within the Massachusetts State Police Crime Laboratory after discovering that analysts at the lab routinely withheld exculpatory information about blood-alcohol tests from defense attorneys. The revelation could affect thousands of drunk-driving tests dating back to 2011. On its website, the Massachusetts State Police Crime Laboratory describes itself as "fully accredited."

- In 2016, police officials in Austin shut down the DNA division of the city's police-run crime lab indefinitely after discovering that analysts were using analytical tools that were badly outdated. The lab also had problems with cross-contamination. Despite multiple red flags about the lab dating back to 2008, the lab was repeatedly accredited by the appropriate bodies within the field of forensics. It took an inspection by the Texas Forensic Science Commission — an oversight body from outside the forensics world — to finally bring these problems to light.
- In 2013, the New York City Medical Examiner's Office underwent a review of 800 rape cases after investigators found 26 instances in which DNA evidence was mishandled or overlooked by an analyst at the lab. The cases spanned 2001 to 2011. In fact, the same year that the city's Medical Examiner's Office announced this review, it also boasted in a news release that its DNA lab received "perfect scores" from an international accrediting organization.
- In their various incarnations, the crime labs in Houston have been the scene of multiple scandals, including 2014 revelations that an analyst had tampered with evidence; a 2005 report finding incompetence, cheating on proficiency tests, and even possible perjury and faking of test results; and the shuttering of the fingerprint lab in 2009 after an audit showed exceptionally high error rates. (This is not at all a comprehensive list of the crime lab problems in Houston.) Through all of this, the Harris County and Houston labs have boasted of accreditation from both national and international forensic organizations. The American Society of Crime Laboratory Directors, for example, accredited the Houston PD lab in 2006. Incredibly, when a 2008 investigation into cheating on proficiency tests led to the closing of the lab's DNA testing unit, the resignation of the head of that unit and the suspension of two other analysts, the head of the organization that accredited the lab said his group *had no plans to revoke the lab's accreditation*.
- A 2013 audit of the crime lab in St. Paul, Minn., found widespread ineptitude, including "sloppy documentation, dirty equipment, faulty techniques and ignorance of basic scientific procedures." The report recommended that the lab cease operations until it could be overhauled from the ground up. Yet in its 2007, 2008, 2009 and even 2012 annual reports, the Minnesota Department of Public Safety touted the St. Paul crime lab's accreditation by the American Society of Crime Lab Directors. In 2009, three years before the scandal broke, crime lab director Frank Dolejsi boasted that the accreditation "is an affirmation of the quality of all aspects of the [crime lab's] forensic science services." After the audit, state officials vowed to win back public trust by promising that the St. Paul crime lab would — you guessed it — get itself accredited.
- In 2014, a crime lab analyst in Florida was arrested for stealing seized drugs and tampering with evidence, possibly tainting up to 2,600 cases from at least 80 different law enforcement agencies. His lab was accredited.
- In 2012, an audit of the North Carolina crime lab done after an exoneration found that analysts in the serology unit had been withholding exculpatory evidence from defense attorneys for at least 16 years. The audit found that analysts routinely overstated their findings and that training manuals were adamantly pro-prosecution, to the point of referring to defense experts as "whores." The manuals were likely pro-prosecution because the analysts themselves reported to prosecutors, who wrote their year-end reviews and determined raises and promotions. The North Carolina lab had been accredited since 1988.

These are just a few examples. In 2013, the ABA Journal reviewed dozens of crime lab scandals across the country. The report found a lot of problems in both accredited and non-accredited labs. Of course, there's Isenberg and Oien's own FBI crime lab, which is arguably the most accredited and respected crime lab in the world. But the FBI lab, too, has been no

stranger to scandal. Just a few years ago, the agency admitted that it's hair-fiber analysts had overstated their findings nearly every time they testified. This occurred over a period spanning decades and implicated thousands of convictions. Those analysts then trained analysts at state and local labs, potentially corrupting untold thousands of more cases. And that's just one scandal of several. There's the Brandon Mayfield debacle. There's the FBI's long-held and since disproved claim about "compositional bullet lead analysis," which also affected thousands of criminal cases. That's quite a record from one of the most respected labs in the world.

Back to Isenberg and Oien.

> Forensic examiners must complete extensive training to be qualified to perform casework in accredited laboratories. Training programs can be one to two years, or longer, and require examiners to perform analyses on samples with a known correct answer. The examiner must also demonstrate a thorough understanding of the science behind the method employed and an understanding of lab policies, procedures, legal rules, evidence handling, etc. The examiner must undergo oral examinations, mock trials, and competency tests for which the correct answer is known. In addition, all examiners must demonstrate competency to apply the processes accurately and reliably before they are assigned actual cases. Once qualified to conduct casework in an accredited laboratory, every examiner undergoes continual competency monitoring through proficiency tests administered at least once per year.

This is difficult to square with the examples above, or with the other scandals at accredited labs far too numerous to list in their entirety.

The most common credentialing agency for crime labs is the aforementioned American Society of Crime Laboratory Directors (ASCLD). Isenberg and Oien even cite this organization's website in laying out the allegedly rigorous process to becoming an accredited forensic examiner. But while ASCLD may claim that these are the criteria crime labs need to meet to earn and keep their credentials, those requirements don't appear to be all that tightly enforced. Labs that don't meet them still get accredited. And previously accredited labs that don't meet them don't seem to lose their status.

From that 2013 report by the ABA Journal:

> If the accreditation process is so rigorous and demanding, critics wonder, then why have so few labs been sanctioned? ASCLD/LAB's website lists the status of all accredited labs and shows that no lab's accreditation is currently revoked or suspended; there are also no labs on probation. And [ASCLD Executive Director Ralph] Keaton says he can count on one hand the number of labs whose accreditation has ever been revoked or suspended, though he says it would probably take two hands to count the number of labs that have ever been placed on probation.
>
> Keaton says that has a lot to do with the overall quality of accredited labs. But critics say it has more to do with the chummy nature of the inspection process, which creates a tendency to "go along to get along" among inspectors, and the agency's own interest in keeping labs accredited. . . .

The links to the pages listing revoked/suspended labs and labs on probation now go to a site called the ANSI-ANQ National Accreditation Board, which apparently merged with ASLD's accrediting arm. But neither website appears to lists any crime labs on probation, currently in suspension, or whose accreditation has been revoked. If even egregious incompetence and misconduct rarely if ever results in a revocation or suspension, what does accreditation really mean?

More from Isenberg and Oien:

Testimony monitoring is also a requirement for accredited laboratories. The Federal Bureau of Investigation (FBI) Laboratory requires that examiners request a transcript for each testimony provided. FBI examiners also must follow approved standards for scientific testimony and reports, which document the acceptable range of conclusions expressed in both laboratory reports and testimony. The DOJ is developing similar documents called Uniform Language for Testimony and Reports, as well as a testimony-monitoring framework, which will apply to all DOJ laboratories. The purpose of these testimony-monitoring activities is to prevent examiner testimony from exceeding scientific limitations.

This sounds great at first. But it's all dependent on the answers to several important questions. First, who determines the "acceptable range of conclusions" for lab reports and testimony? Is it someone who already practices in that particular field? What does "approved standards for scientific testimony" mean? Who is setting those standards? Who is approving them? Who decides if an analyst has "exceeded scientific limitations?"

The entire point of the National Commission on Forensic Science empaneled by President Barack Obama was to bring in actual scientists to make such determinations. Attorney General Jeff Sessions allowed that commission's charter to expire shortly after taking office. (It's worth noting that while Obama deserves credit for setting up the commission, his record in this area was mixed at best.) Last August, Deputy Attorney General Rod J. Rosenstein announced a new forensics review and monitoring process at DOJ. He also revealed that the new system would be overseen by a man named Ted Hunt.

Ted Hunt is not a scientist. He is a longtime prosecutor. His supporters point out that he was part of the very Forensic Science Commission that Obama empaneled and that reformers lament was allowed to expire. But Hunt was often a voice of dissent on the commission. Here's a look at Hunt's record on the commission from Mother Jones:

> In March 2016, the commission [recommended](#) that then-Attorney General Loretta Lynch direct forensic experts and attorneys working on behalf of the Justice Department to stop using the phrase "to a reasonable degree of scientific certainty." The phrase is commonly used on witness stands and in lab reports and gives juries and judges a sense of factuality, but it is subjective and lacks any agreed-upon meaning across the sciences. Hunt was one of two commission members who [opposed](#) the recommendation, which Lynch [adopted](#) last September.
>
> Lynch also adopted a [recommendation](#) by the commission requiring forensic testing labs that work with the department and its attorneys to publicize their internal procedures, from equipment maintenance to estimations of uncertainty, in order to foster transparency, trust, and best practices in the industry. Hunt as one of four commissioners who [opposed](#) it.
>
> Last September, when the commission released a [document](#) supporting stricter accreditation standards for forensic labs, Hunt [voted](#) against it. And when the commission [recommended](#) that the National Institute of Standards and Technology conduct scientific evaluations of the "technical merit of test methods and practices used in forensic science disciplines," he [opposed](#) that, too. At its final meeting, when members had already been informed that the group would be coming to an end, several commissioners pushed for a resolution encouraging experts to use more quantitative language to convey the accuracy of forensic testimony. The resolution narrowly failed, with Hunt among the nays.

So DOJ did away with a transparent, external commission mostly populated and led by scientists, and replaced it with an internal commission led by a career prosecutor who has opposed efforts to increase transparency, increase accountability, increase scientific accuracy and strengthen the requirements for accreditation. This is likely why Hunt's appointment was [hailed by ASCLD](#) and [forensics groups](#) such as the International Association for Identification and the American Academy of Forensic Sciences, and why the new DOJ venture has been viewed with skepticism by [groups such as the Innocence Project](#).

In the remainder of their essay, Isenberg and Oien take aim at the report by the President's Council of Advisors on Science and Technology (PCAST) issued in the waning months of the Obama administration. That report was perhaps the most strident and explicitly critical report on forensics to date. The authors of and researchers for the report were scientists, not forensics practitioners or lawyers (though a panel of judges and attorneys did contribute in an advisory role).

The essay's discussion of the PCAST report is a bit technical to delve into here with too much detail, but in general, the PCAST authors posited that because most pattern-matching fields of forensics are so subjective — an analyst basically just "eyeballs" two samples and comes up with language to describe how similar or different they are — the only way to determine if these fields are reliable is what's known as a "black box" test. We can't evaluate a particular analyst's process for determining matches because so much of it is done in his or her head. Instead, we have to look at results. That means administering competency tests. Give analysts a sample from a case in which the culprit is known, give them samples from multiple suspects, and record who gets it right and who doesn't. As you might guess, many forensic analysts and the professional groups they belong to are reluctant to submit to such tests. Most of these fields have already been accepted by the courts. They have little to gain, and a lot to lose.

In their essay, Isenberg and Oien dispute the notion that black-box tests are the only real way to measure the reliability of a given field. They argue that relying too much on such tests "ignores much peer-reviewed research, overlooks critical aspects of many studies, and fails to acknowledge the empirical value of these studies." But these studies are, again, typically reviewed only by other practitioners from the same field, using the same procedures and standards within the field. Here again, you're evaluating the legitimacy of an entire field based on criteria that already accepts that field's legitimacy.

In conclusion, Isenberg and Oien write:

> Science continuously evolves and is built upon observation, study, and experience that spans hundreds of years. The justice system would not be well served by the exclusion of reliable forensic methods and techniques that provide valuable information to a wide range of stakeholders.

Of course, whether some of these fields are reliable is precisely what's in dispute.

But at issue here is how we evaluate new or questionable expertise for use in the courtroom. What do we do going forward? I think the answer is that we only allow this sort of analysis and expertise after it has been subjected to the rigors of scientific inquiry. Before we let a new drug hit the market, we require its manufacturer to prove that the drug is safe and effective. Those standards are enforced by scientists with expertise in pharmaceuticals. When a drug slips through and causes harm, we pull it from the market. It isn't a perfect system, and there are interesting debates at its parameters. But in general, we test new technology before we put it to everyday use.

Contrast this to forensics. The decision whether to allow a new field of forensics into court is made by a judge, not a scientist, or even a fellow practitioner. Judges typically look for guidance on these questions not from scientists, but from other judges. The briefs in such challenges are written by lawyers. Judges then tend to err on the side of letting evidence in, on the assumption that our adversarial system will sort it out. (In his speech last year, Rosenstein touted this flaw as a feature.) Even once we discover that a field is scientifically suspect, it's difficult to get the courts to even acknowledge it, much less stop it from being used again, much less correct the cases that may have already been tainted.

Like the other defenders of these fields, Isenberg and Oien say we should simply trust the internal procedures at FBI and DOJ to get forensics right. Even at face value, that's a difficult argument to accept. FBI agents and federal prosecutors aren't evaluated for their allegiance to scientific principles. They're evaluated on their ability to close cases and win convictions. Therefore, there's a strong incentive for them to sacrifice sound science for expedience. That incentive will affect even the most conscientious and fair-minded at DOJ.

But the case Isenberg and Oien are making is even harder to accept given the history of these disciplines and the ongoing onslaught of forensics and crime lab scandals. "Just trust us" hasn't worked in the past. Why should we think it would work now?

Gerry LaPorte
Director
Office of Investigative and Forensic Sciences
National Institute of Justice

Office: (b) (6)
Mobile: (b) (6)

FW: Bite Mark Thinkshop

FYI

Best Regards,

Gerry LaPorte
Director
Office of Investigative and Forensic Sciences
National Institute of Justice
810 7th Street NW
Washington, DC 20531
Office: (b) (6) ██████
Mobile (b) (6) ███████

From: Cavanagh, Richard R Dr. (Fed) [mailto:(b) (6) ████████████]
Sent: Friday, April 27, 2018 2:51 PM
To: Laporte, Gerald (OJP) <(b) (6) ████████████>
Subject: Bite Mark Thinksh

Good afternoon,

Attached please find an invite for the Steering Committee, I hope you can join us in the planning of a Thinkshop on Bite Mark.

Thank you.

Rich

April 26, 2018

MEMORANDUM FOR Gerry LaPorte, Director, NIJ Office of Investigative and
Forensic Sciences

From: Richard Cavanagh, Director, NIST Special Programs Office

Subject: Steering Committee for Bitemark Thinkshop

NIST is moving forward with our plan to host a Bite Mark Thinkshop. This activity falls
under our Congressionally-appropriated mandate to establish *Foundational Reviews of
Forensic Methods* (see appended document). As part of our planning process for the Bite
Mark Thinkshop, we would like to invite you to be part of the Steering Committee for
this meeting.

NIST members of the Steering Committee currently include:

> Richard Cavanagh – Director, NIST Special Programs Office
> Willie May – NIST Associate
> William (Bill) MacCrehan – NIST Research Chemist

We see this as a one-and-a-half-day meeting at NIST, with somewhere near 50 attendees.
The format that we currently envision for the meeting would incorporate plenary
presentations on the overall technical needs of the forensic enterprise as well as 5 to 7
topic areas for more focused discussions related to Bite Marks. "Break Out" discussions
of these select topic areas would result in a summary that would be presented on the
second day by each team. A professional facilitator would preside over the meeting,
encourage fruitful discussions, and help assemble a final report. Our intent is to publish
our findings as a publicly-available NIST Special Publication which would be publicly
available (with a DOI).

We are considering several dates this fall/winter for the Thinkshop, so we are hoping to
(virtually) convene the steering committee soon in order to gather/share ideas on how the
Thinkshop can be most effective.

Please let me know if you will be able steer this effort.

NIST Scientific Foundation Reviews of Forensic Methods

Motivation: The scientific foundation of methods used by forensic practitioners has been challenged in recent years. In September 2016, the National Commission on Forensic Science (NCFS) requested that NIST perform technical merit evaluations of forensic methods[1]. About the same time, the President's Council of Advisors on Science and Technology (PCAST) published a report that recommended NIST examine "foundational validity" of key forensic methods:

> PCAST report[2], page 124: "PCAST recommends that NIST be tasked with responsibility for **preparing an annual report evaluating the foundational validity of key forensic feature-comparison methods**, **based on available, published empirical studies**. These evaluations should be conducted under the auspices of NIST, with **input from additional expertise as deemed necessary from experts outside forensic science, and overseen by an appropriate review panel**. The reports should, as a minimum, produce assessments along the lines of those in this report, updated as appropriate. Our intention is not that NIST have a formal regulatory role with respect to forensic science, but rather that **NIST's evaluations help inform courts, the DOJ, and the forensic science community**."

Expectations: Based on the PCAST recommendations to NIST shown above, expectations for these scientific foundation reviews include: (1) *evaluating* key forensic feature-comparison methods to assess foundational validity, (2) *examining* available, published empirical studies, (3) *conducting* work with input from outside experts, (4) *reporting* obtained results via a publicly-available annual report, and (5) *informing* stakeholders including courts, the Department of Justice, and the forensic science community.

Budget: The FY2018 Congressional budget provides $1 million per year to NIST to conduct "technical merit evaluations" of forensic disciplines.

Proposed Approach: The following approach is being pursued to meet expectations for these studies: (1) select forensic discipline and method to evaluate, (2) assemble a NIST review team, (3) seek input from outside experts on issues to consider, (4) gather and examine the scientific literature to develop a bibliography of foundational literature, (5) summarize capabilities and limitations of methods used, (6) conduct inter-laboratory studies as appropriate to address specific questions, (7) identify knowledge gaps for future research, (8) prepare and publish report, and (9) inform stakeholders of findings.

Pilot Studies: NIST launched pilot studies in September 2017 involving (1) DNA mixture interpretation and (2) bitemark evidence. From these pilot studies, we hope to examine challenges involved and efforts required for these reviews. Various methods for outside expert input are being explored. The DNA mixture study report is planned for release in the summer of 2018 as a NIST Special Publication.

Plans for Future Work: Several forensic disciplines are under consideration for future NIST scientific foundation reviews pending assembly of appropriate review teams. These include (1) firearms and tool marks, (2) latent fingerprints (building on a September 2017 AAAS report[3]), (3) digital evidence, (4) toxicology, (5) footwear & tire treads, and (6) gunshot residue analysis. With the FY2018 funds provided, two or three studies can probably be conducted simultaneously. We anticipate that each study will have different challenges and timelines for completion. The pilot studies should help inform future efforts. In some cases, there may be connections between scientific foundation reviews and NIST-sponsored working groups developing process maps and examining human factors issues in a specific forensic discipline, such as has been done previously[4] with latent fingerprints and handwriting analysis.

---

[1] https://www.justice.gov/archives/ncfs/page/file/905541/download

[2] https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf

[3] https://www.aaas.org/report/latent-fingerprint-examination

[4] https://www.nist.gov/topics/forensic-science/working-groups/nistnij-working-groups

RE: Revised FRE 702 Talkers

From: "Hur, Robert (USAMD)" ▓(b) (6)▓
To: "Antell, Kira M. (OLP)" <▓(b) (6)▓
Cc: "Hunt, Ted (ODAG)" <▓(b) (6)▓ "Goldsmith, Andrew (ODAG)" <▓(b) (6)▓
"Shapiro, Elizabeth (CIV)" <▓(b) (6)▓
Date: Mon, 23 Apr 2018 19:14:30 -0400

Just saw these – these may address some of the points I made earlier. Will review tonight.
Thanks!

From: Antell, Kira M. (OLP) <▓(b) (6)▓
Sent Friday, April 20, 2018 5:21 PM
To: Hur, Robert (USAMD) <▓(b) (6)▓ >
C Hunt, Ted (ODAG) (JMD) ▓(b) (6)▓ Goldsmith, Andrew (ODAG) (JMD)
<▓(b) (6)▓ Shapiro, Elizabeth (CIV) <▓(b) (6)▓ >
Subject Revised FRE 702 Talkers

Good evening Rob,

Attached is a revised version of the talkers we discussed last week It incorporates your edits and highlights two cases from the digest you may find interesting. I am in the process of pulling together a few post-PCAST cases with brief summaries There won't be many but I want to make sure give you the most relevant ones

Thanks,
Kira

Kira Antell
Senior Counsel
Office of Legal Policy
U.S. Department of Justice
950 Pennsylvania Avenue, NW
Washington. DC 20530
▓(b) (6)▓
▓(b) (6)▓

efbafecf-c8b4-4f74-b3de-9dc86a2d44a9

RE: OJP and PCAST

Hi Ted,

Given that none of us have heard of this, I think it is appropriate to advise Ken that we are not aware of this policy.

-K

From: Laporte, Gerald (OJP)
Sent: Monday, Septemb
To: Hunt, Ted (ODAG) <(b) (6)   >
Cc: McGrath, Jonathan (OJP) <(b) (6)   >; Antell, Kira M. (OLP) <(b) (6)
Subject: Re: OJP and PCAST

I have no idea where this is coming from? I assume he is referring to BJA - and not BJS?

Best Regards,

Gerry LaPorte
Director
Office of Investigative and Forensic Sciences
National Institute of Justice
810 Seventh Street NW
Washin
Office: (b) (6)
Cellphone (b) (6)

On Sep 2, 2018, at 10 52 AM, Hunt, Ted (ODAG)   (b) (6) wrote

> All,
>
> Ken Melson asked me the question set forth below at the recent Symposium we co sponsored at NAAG I didn't know what he was talking about, so told him I'd try to run this to ground. Have any of you heard of anything like this coming out of OJP/BJS?
>
> Thanks,
>
> Ted
>
> Begin forwarded message
>
> > **From:** Kenneth Melson <   (b) (6)
> > **Date:** September 1, 2018 at 8:08:09 PM C
> > **To:** <   (b) (6)
> > **Subject:** OJP and PCAST
> >
> > Ted, you may remember that at the NAAG conference I mentioned that BJS grantees doing training for prosecutors and defense attorneys on issues of admissibility of scientific evidence are forbidden by OJP/BJS from mentioning the PCAST report. You indicated that you could look into that for me. I will me speaking to such a group this week and would appreciate your guidance. Thanks. Ken.

Re: OJP and PCAST

From:	Ken <(b) (6) >
To:	"Hunt, Ted (ODAG)" <(b) (6)
Date:	Tue, 04 Sep 2018 17:08:37 -0400

Thanks. Then I will go ahead and mention PCAST. Ken

On Sep 4, 2018, at 11:24 AM, Hunt, Ted (ODAG) <(b) (6)   wrote:

Ken,

I've reached out to a few people who said they are unaware of any such directive. And as I've said, i haven't either.

Ted

On Sep 4, 2018, at 9:51 AM, Ken <(b) (6)   > wrote:

Thanks. It is actually BJA, not BJS. Ken

On Sep 2, 2018, at 6:14 PM, Hunt, Ted (ODAG) <(b) (6)   wrote:

Ken, I'm checking with others and will let you know if there's anything to this as soon as I hear back.

Ted

On Sep 1, 2018, at 8:08 PM, Kenneth Melson <(b) (6)   > wrote: