NATIONAL COMMISSION ON FORENSIC SCIENCE

Adjudication of Public Comments on the Views of the Commission Statistical Statements in Forensic Testimony

---

Type of Work Product:
Views Document by the Subcommittee on Reporting And Testimony

Public Comment Summary:
There were eleven public comments received. Nine were from non-commission members. Two commissioners provided comments. None identified themselves as statisticians.

Adjudication Process Used by the Subcommittee:
Members of the working group received copies of all public comments. Responses were suggested by working group members and The public comments and responses were summarized by one of the subcommittee co-chairs (Redle). A vote of the subcommittee membership was then held electronically. Subcommittee members were advised that comments by subcommittee members could be forwarded to the co-chair (Redle) for inclusion with this record.

Itemized Issues and Adjudication Summary
Attached as Exhibit A are the responses to the public comment as received.

Comment #DOJ-LA-2017-0004-0006
Comment restricted to fingerprint matches only:
The commenter states that the theory underlying this analysis is that fingerprints are both unique and persistent and this theory has not been disproved. She therefore objects to Commission view #4 which suggests that examiners generally should avoid statements declaring that a given fingerprint comes from an identified "source." She suggests that her comment is limited to the best case scenario in terms of quality of print and that prohibiting a statement of source automatically and unjustifiably places doubt in the mind of a juror. In her view examination by an opposing expert is a sufficient safeguard for purposes of reliability. She says the world has come to accept fingerprints as a definitive form of identification and we need not complicate the matter any further for juries. People can and have made mistakes.

The commenter suggests that fingerprints have stood the test of over 100 years of experience and no change is necessary.

RESPONSE: Fingerprint identifications can be very strong evidence of identity. However, the issue that is of interest in making statements about the origin of latent fingerprints is not whether fingerprints are unique. It is how strongly the perceived similarities and differences between a latent print (from an unknown source) and an exemplar print (from a known finger) support specific hypotheses about the origin of the latent print. Claims of "unique" identifications — those that necessarily exclude all other individuals in the world — are not required. E.g., NIST Expert Working Group on Human Factors in Latent Print Analysis, Latent Print Examination and Human Factors: Improving the Practice Through a Systems Approach, Gaithersburg: National Institute of Standards and Technology, 2012.

There is no need to continue doing something one way if that same task can be accomplished better, with greater reliability and transparency. Progress and improvement is in the nature of humanity.

Comment #DOJ-LA-2017-0004-0007
The writer brings the attention of the working group and the subcommittee to information given at a NIST conference or conferences and published in the American Journal of Physical Anthropology suggesting that fingerprint probabilities may be influenced by factors such as age, gender or racial ancestry. The writer suggests that this may caution against the use of criminal fingerprinting databases as they may bias toward young black males. The writer argues that databases or likelihood ratios may need to be regionally specific.

RESPONSE: The appropriate database for computing likelihood ratios is an important issue for study. The views document encourages attention to this matter.

Comment #DOJ-LA-2017-0004-0009 The commenter makes five points with respect to the use of statistical models in latent print examination. Those points argue that criticism of the discipline fails to consider all features present in a print; make the "false assumption" that examiners "operate in a vacuum" rather than having access to additional information(presumably about the case and the print); attempts to quantify a random process that cannot be quantified; he asserts that current practice has not been shown to suffer from excessive false positives; and that

a proposal that examiners not identify the "source" of a latent print will prove confusing to jurors and is contrary to 100 years of experience.

RESPONSE: This comment suggests that objective, quantitative models of likelihoods for fingerprint comparison have limitations. The views document recognizes that all statistical analyses have limitations. So do subjective judgments of trained examiners. Therefore, the views document looks forward to the integration of both methods and is consistent with the conclusion expressed in Christophe Champod, Fingerprint Identification: Advances Since the 2009 National Research Council Report, 370 Phil. Trans. Royal Soc'y B 20140259, at 7 (2015), that "in the future, fingerprint evidence will not be a matter of individual opinion anymore, but will be constructed through the harmonious play of multiple experts (humans and computer systems), working according to specified procedures, to deliver a consensus probabilistic assignment of the weight to be attached to their forensic findings." Acknowledging the possibility that the particular features discerned in a latent print might be replicated in another individual is more appropriate than "the blunt assertion of the uniqueness of fingerprints or the opinion delivered *ispe dixit*." Id. at 1.

Comment #DOJ-LA-2017-0004-0010
This comment is supportive: "I agree with the views expressed by the commission." It notes that efforts to improve on the identification of compounds by mass spectrometry are underway and that "the forensic chemist should work closely with the investigators."

Comment #DOJ-LA-2017-0004-0012
The author suggests the use of proficiency testing with individual examiner testing results being discoverable as providing greater reliability.

RESPONSE: This outside of the scope of this views document?

Comment #DOJ-LA-2017-0004-0013 The author supports the concept that information on methodology and limitations (including lack of appropriate statistical model or databases) be made available. He suggests that such information be made available "upon request." In this regard the comment seems to be more aptly addressed in the "report content" document and so this comment has been referred to that working group.

RESPONSE: This views document agrees that information regarding limitations should be made available but that the lack of appropriate statistical models or databases as required for statistical statements of weight to be made or implied are so fundamental that they be required. The writer suggests that such information is not of interest to the typical consumer of the information and therefore is just adding another task to already over-worked examiners and analysts. **We suggest that while the average consumer (attorneys, judges, investigators and jurors) may not be interested, they should be and identification of such limitations should not prove to be unduly burdensome.**

The commenter suggests that example #4 should be deleted as it is a philosophical discussion on how we know something. **Instead, example #4 is an example of limitations in measurement or available information that may have an effect upon the accuracy or reliability of the**

conclusions reached. The limitations identified in the example have such an effect in cases where blood alcohol content at a particular point in time are being estimated through the use of retrograde extrapolation.

Comment #DOJ-LA-2017-0004-0014
The author argues that statistical information can be misused and misleading and are only as good as data and methods used to arrive at the estimation. She writes that the document implies that some pattern evidence disciplines such as latent prints and shoe prints are currently capable of being quantified.

RESPONSE:
We acknowledge that statistics, like any human endeavor, are subject to error and fraud. Appropriately structured statistics can provide valuable information regarding uncertainty, which may prove helpful to the trier of fact. On the question of the unavailability presently of tools necessary for quantification the document states: "The Commission supports continued research to provide the requisite scientific data." "Statistical statements and those that appear in connection with many other forms of evidence should be based on…." The document supports a research agenda in this area and the lack of such information necessary to permit quantification should not serve as a justification for the field to halt progress in this area in favor of *ipse dixit* pronouncements.

It recognizes that for many of these examinations such information may not exist. For other types of trace and pattern evidence, however, no widely accepted probabilistic models of the phenomena that give rise to the features are available. **It argues for greater research and study. The document argues that if you intend to make or imply statistical statements regarding your analysis or examination, you need to have data and models which support those statements. To the extent that you believe such statements to be necessary or beneficial and such data or model doesn't currently exist research is necessary. In the absence of such statistical data or statistical models, the examination or analysis may still provide probative information only without an accompanying statistical statement, expressed or implied.**

Comment #DOJ-LA-2017-0004-0015
The commenter cites the document to the effect that "[s]tatistical statements should be based on a statistical model that accurately assesses the strength of the inference in question or describes the process." He argues, "[e]very model is inaccurate to some degree." How accurate should it be to nevertheless be admissible? Do forensic associations have value if a valid model is unattainable?

RESPONSE: This comment suggests that objective, quantitative models of likelihoods for fingerprint comparison have limitations. The views document recognizes that all statistical analyses have limitations. So do subjective judgments of trained examiners. Therefore, the views document looks forward to the integration of both methods and is consistent with the conclusion expressed in Christophe Champod, Fingerprint Identification: Advances Since the 2009 National Research Council Report, 370 Phil. Trans. Royal Soc'y B 20140259, at 7 (2015), that "in the future, fingerprint evidence will not be a matter of individual opinion anymore, but will be constructed through the harmonious play of multiple experts (humans

and computer systems), working according to specified procedures, to deliver a consensus probabilistic assignment of the weight to be attached to their forensic findings." Acknowledging the possibility that the particular features discerned in a latent print might be replicated in another individual is more appropriate than "the blunt assertion of the uniqueness of fingerprints or the opinion delivered *ispe dixit.*" Id. at 1.

The position we have taken is that such associations may have value but the practitioner may not make or imply a statistical statement or analysis in support of the value or weight to be ascribed to that association.

The commenter continues by citing the document:

> "'A statistical foundation is necessary to enable forensic science to assess and express uncertainty.' I would agree for fields when statements to express uncertainty have been demonstrated. However, accurate techniques exist with no accepted way to accurately express uncertainty. Further, no data exists to suggest that expressions of uncertainty are even possible for some techniques. This document needs to explore these real possibilities instead of assuming that accurate expressions of uncertainty exist or soon will."

The document says:

> **"The presentation should include statements of limitations and uncertainties in measurements or observations. … If the forensic science practitioner has no information on sources of error in measurements and inferences, the forensic science practitioner must state this fact. … Forensic science practitioners should not state that a specific individual or object is the source of the forensic science evidence and should make it clear that, even in circumstances involving extremely strong statistical evidence, it is possible that other individuals or objects could possess or have left a similar set of observed features. … The absence of models and empirical evidence should be expressed both in testimony and written reports."**

The writer then takes issue with other statements in the report by pointing to current absences in research supporting some existing disciplines. **These are testing and methodology limitations and we have taken the position that these limitations should be acknowledged until advances in research or methodologies permit otherwise.**

Comment #DOJ-LA-2017-0004-0016 The commenter makes the point that fingerprint examiners are encouraged to participate in proficiency testing. IAI also requires testing for certification that includes history, chemistry, photography and "actual" latent print examinations.

RESPONSE: It is our position that proficiency testing and certification are not a sufficient basis by themselves upon which to make or imply statistical statements.

Comment #DOJ-LA-2017-0004-0018
Professor Cole's criticism of the document starts with the premise that a declaration of "source" by an expert is not a statistical statement. We have taken the position that such a declaration is a statistical statement.

RESPONSE:

Comment #DOJ-LA-2017-0004-0019
Commenter suggest that the following statement in the document is inconsistent with Recommendation #6:
"But a "positive association" is not probative unless it is more probable when the items have a common source than when they originate from different sources. Indicating the statistical weight of the positive association therefore requires a statement of how common or rare the association is, based on a database or empirical data linked to the case at hand." (Page 4, paragraph 2, underscore added)."

He suggests:
"In light of the former statement, what value is that? There is no value (relevance) to the fact finder learning that a comparison took place unless that comparison resulted in some type of expert opinion regarding the weight of the match. And if a positive association is not probative unless it is more probable when the items have a common source than when they originate from different sources (per the former statement), then how can a comparison that has resulted in a determination of source commonality, without an assessment of weight (per the latter statement), be probative? (Per the former statement)."

RESPONSE: In our view, there is no contradiction. The premise that "There is no value (relevance) to the fact finder learning that a comparison took place unless that comparison resulted in some type of expert opinion regarding the weight of the match" is false. Case law allows experts to point out similarities in features without opining on how probative the similarities in the features are. Logically, the similarities must be more probable for same-source pairs than for different-source pairs for the comparison to be relevant as proof of their identity, but the expert need not quantify these probabilities for such relevance to exist. It may be a jury noticeable fact. Consider a comparison of duct tape that appears to have been torn from a roll when the ends are inspected under a microscope. The corresponding features are relevant evidence. The expert can show the correspondence without opining on what it proves. Indeed, some courts have required this features-only mode of presentation for some forms of evidence.

Commenter next suggests the following (quoting initially from the document):

"'More generally, when dealing with features, such as the refractive index of glass or the peak heights in an electropherogram of DNA fragments, that have more values than 'absent' and 'present,' the classification of 'matching' and 'not matching' omits statistical information related to the degree of similarity. The weight ascribed to any degree of association depends on (1) the probability of the degree of correspondence in the features, given that the samples came from the same source, and (2) the probability for the same measurement, given that the samples came from different sources. When the former probability is much larger than the latter (i.e.,—when the observed degree of similarity occurs much more often for same-source samples than for different- source samples), —the evidence supports the hypothesis of a common source." (Page 4, paragraph 3).' Without explicitly stating, this paragraph advocates for the use of likelihood

ratios in an effort to make full use of continuously valued data. Despite the relative merits of that position and notwithstanding the document's assertion to the contrary, this statement seems to endorse a specific statistical model (LRs) for a particular application (continuously valued data)."

RESPONSE: We disagree. The statement is one of explanation and not endorsement. The document does not insist on an LR for continuous data. It may incidentally points out the value of the approach. A LR is not a statistical model. There are different models that give rise to different likelihood functions.

Commenter suggests:

"'Any recommendation on presenting explicit probabilities, however derived for specific forensic evidence, might distinguish between probabilities based on a statistical model and ones that characterize the forensic science practitioner's subjective sense of how probable the evidence is under alternative hypotheses.'" (Page 4, paragraph 4).
If such a recommendation "might" distinguish between statistical and subjective estimates then why doesn't this one do so in a more candid way by conceding and making more explicit the significant human (subjective choices and decisions that are precedent to the construction and use of a statistical model?'
To imply that "statistical models" are not also "subjective" in significant ways is disingenuous. It is true, as the document notes, that "statistical models are approximations." What is left unstated, however, is that not only is there "some uncertainty in the selection of a model," but also that there is human subjectivity in that selection; in the background assumptions upon which that model is based; and with the assumptions used to acquire and utilize that data that populates the model. Simply stated, as with experience-based source conclusions, statistical models are also the inexorable result of human subjectivity."

RESPONSE: Human judgment goes into statistical modeling, and the document does not hide this fact. But it is one thing to make numbers up out of thin air or "training and experience" and another to derive them from an explicit model and data that others can inspect. Greater "candor" (or clarity) would result from changing "might distinguish" to "should distinguish."

Commenter raises questions about the meaning of View # 3:

"This view presumably relates to the reporting of speculative case-specific sources of error and uncertainties generally inherent in a forensic method that could have occurred, but for which there is no evidence that they actually *did* occur in a specific case."

RESPONSE: The principle advocated by View #3 is that "[t]he forensic science practitioner should report the limitations and uncertainty associated with measurements and the inferences that could be drawn from them." Thereafter it expresses possible options that might be pursued to implement reporting of limitations and uncertainty.

Commenter writes:

"But a "positive association" is not probative unless it is more probable when the items have a common source than when they originate from different sources. Indicating the statistical weight of the positive association therefore requires a statement of how common or rare the association is, based on a database or empirical data linked to the case at hand." (Page 4, paragraph 2, underscore added).

This statement remains inconsistent with this document's Recommendation #6: "Not all forensic sub-disciplines currently can support a probabilistic or statistical statement. There may still be value to the factfinder in learning whatever comparisons the forensic science practitioner in those sub-disciplines has conducted." (Underscore added). In light of the former statement, what value is that? There is no value (relevance) to the fact finder learning that a comparison took place unless that comparison resulted in some type of expert opinion regarding the weight of the match. And if a positive association is not probative unless it is more probable when the items have a common source than when they originate from different sources (per the former statement), then how can a comparison that has resulted in a determination of source commonality, without an assessment of weight (per the latter statement), be probative? (Per the former statement).

RESPONSE: We disagree that there is an inconsistency. The premise that "There is no value (relevance) to the fact finder learning that a comparison took place unless that comparison resulted in some type of expert opinion regarding the weight of the match" is false. Case law allows experts to point out similarities in features without opining on how probative the similarities in the features are. Logically, the similarities must be more probable for same-source pairs than for different-source pairs for the comparison to be relevant as proof of their identity, but the expert need not quantify these probabilities for such relevance to exist. It may be a jury noticeable fact. Consider a comparison of duct tape that appears to have been torn from a roll when the ends are inspected under a microscope. The corresponding features are relevant evidence. The expert can show the correspondence without opining on what it proves. Indeed, some courts have required this features-only mode of presentation for some forms of evidence.