

A Call to Arms: Part II

Summary of the NAS assessment of frequently used forensic evidence.

BY MARK W. PROTHERO

DNA Evidence: The committee gave DNA evidence a thumbs up, the model for all forensic disciplines:

Unlike many forensic techniques that were developed empirically within the forensic science community, with limited foundation in scientific theory or analysis, DNA analysis is a fortuitous by-product of cutting-edge science. Eminent scientists contributed their expertise to ensuring that DNA evidence

offered in a courtroom would be valid and reliable.... and by 1996 the National Academy of Sciences had convened two committees that issued influential recommendations on handling DNA forensic science....

DNA analysis has also been subjected to more scrutiny than any other forensic science discipline, with rigorous experimentation and validation performed prior to its use in forensic

investigations.¹

Controlled Substance Analysis:

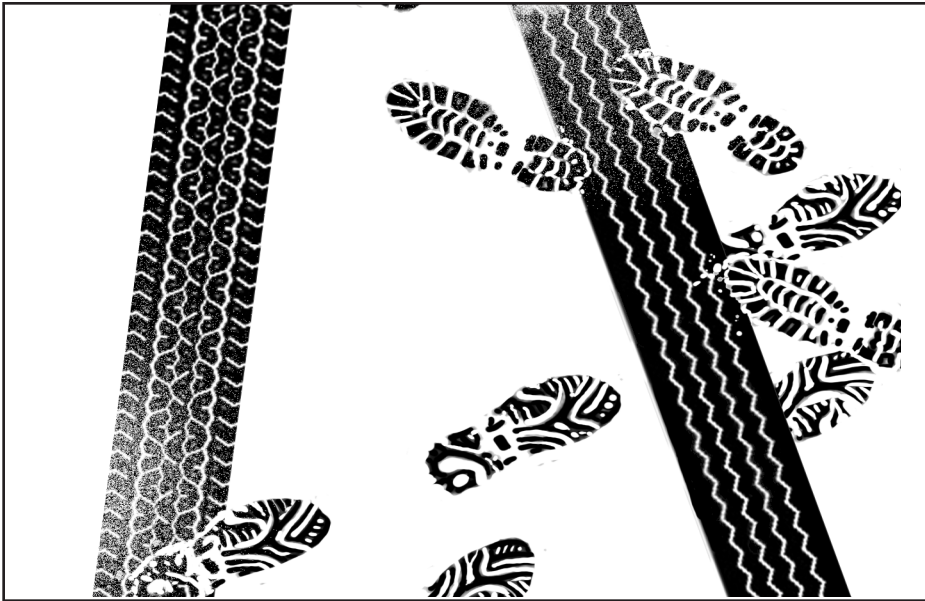
The committee recognized that “[T]he analysis of controlled substances is a mature forensic science discipline and one of the areas with a strong scientific underpinning.”² The report concluded: “The chemical foundations for the analysis of controlled substances are sound, and there exists an adequate understanding of the uncertainties and potential errors.”³

The committee was critical of the summary “terse reports” produced by the labs for the attorneys and courts. While experienced chemists may know what tests to perform in a given situation, the reports typically do not include details of the analysis. Whether the right tests were done and proper protocols followed is often omitted absent supplemental discovery requests. “This ambiguity would be a less significant issue if the reports presented in court contained sufficient detail about the methods of analysis.”⁴

Friction-Ridge Analysis — Fingerprints:

Formerly the gold-standard of identification (e.g., pre-forensic DNA), friction-ridge analysis applies to fingerprints, palm prints, and sole prints. It is “an example of what the forensic science community uses as a method for assessing ‘individualization’ — the conclusion that a piece of evidence....comes from a single unambiguous source. Friction ridge analysis shares similarities with other experienced-based methods of pattern recognition, such as those for footwear





and tire impressions, toolmarks, and handwriting analysis....”⁵

The report discusses data collection, analysis, comparison, evaluation, verification (ACE-V), methods of interpretation, and issues regarding the conclusions in the reporting of the results. Regarding the “ACE-V” method, the report states:

ACE-V.... is not specific enough to qualify as a validated method for this type of analysis. ACE-V does not guard against bias; is too broad to ensure repeatability and transparency; and does not guarantee that two analysts following it will obtain the same results. For these reasons, merely following the steps of ACE-V does not imply that one is proceeding in a scientific manner or producing reliable results.⁶

Problems with lack of documentation of the analysis are recognized, as well as the claims by the latent print community that the method and analysis has a zero error rate.

Errors can occur with any judgment-based method, especially

when the factors that lead to the ultimate judgment are not documented.... Clearly, this assertion is unrealistic, and moreover, it does not lead to a process of method improvement. The method, and those who use it, are inextricably linked, and both involve multiple sources of error (e.g., errors in executing the process steps, as well as errors in human judgment).⁷

The report calls for much additional research regarding the presumption of uniqueness, different sources vs. distortions caused by variations in pressure, ridge flow and crease pattern distribution, skin conditions, residue, mechanics of touch, and other factors impacting the quality and characteristics of latent prints.

Formal research could provide examiners with additional tools to support or refute distortion explanations. Currently, distortion and quality issues are typically based on “common sense” explanations or on information that is passed down through oral tradition from examiner to examiner. A criticism

of the latent print community is that the examiners can too easily explain a “difference” as an “acceptable distortion” in order to make an identification. (footnote omitted)⁸

Pattern/Impression Evidence

— **Shoeprints, Tire Tracks:** Like fingerprints, other types of pattern or impression evidence — such as a shoeprint impression left on the assault victim or in the ground outside the window used for entry by a burglar — are used in criminal investigations and prosecutions. Other types of impression evidence include bite marks, toolmarks, and some types of bloodstain evidence, which are examined in separate sections. Other less-common types of impression evidence include ear prints, lip prints, and glove prints. This section of the NAS Report deals specifically with shoeprints and tire tracks.

The analyses of these types of evidence generally follow the same accepted methods:

- First, determining the general class of the source of the impression (for example, identifying a shoeprint as having been made by a size 10 Adidas Superstar or a tire track left by a Michelin Steel-belted Radial 110x65).
- Second, identifying individual characteristics caused by use and wear and tear then comparing the individual characteristics of the crime scene impression evidence with the individual characteristics of the suspect evidence.

According to the report, such “Identifications are largely subjective and are based on the examiner’s experience and on the number of individual, identifying characteristics in common with a known standard.”⁹

The report is critical of the lack of a solid scientific basis for making a positive identification: “But there is no defined threshold that must be surpassed, nor are there any studies that associate the number of matching characteristics with the probability that the impressions were made by a common source.”¹⁰

The committee further noted that:

there is no consensus regarding the number of individual characteristics needed to make a positive identification, and the committee is not aware of any data about the variability of class or individual characteristics or about the validity or reliability of the method. Without such population studies, it is impossible to assess the number of characteristics that must match in order to have any particular degree of confidence about the source of the impression.¹¹

Toolmark & Firearms Identification:

Toolmarks are another type of impression evidence commonly seen in criminal investigations and prosecutions. These include any number of tools such as hammers, screwdrivers, bolt-cutters, or crowbars. Firearm examination is essentially a sub-set of toolmark examination in that the “tools” are the inner workings of a firearm, such as the firing pin or the extractor and ejector mechanisms or the barrel that the bullet passes through after the trigger is pulled.

Examiners initially look for class and sub-class¹² characteristics. If evidentiary items cannot be excluded on that basis, examiners then look for microscopic individual characteristics. The examiner looks at the evidence through use of “comparison microscopes” that allow the examiner to look at two samples at the same time,

side by side. The examiner makes a determination of whether or not a “match” exists through this visual comparison. A “match” is called if the examiner determines there is “significant agreement” between the samples. Agreement is “significant” as defined by the Association of Firearm and Toolmark Examiners (AFTE) “when it exceeds the best agreement demonstrated between toolmarks known to have been produced by different tools and is consistent with the agreement demonstrated by toolmarks known to have been produced by the same tool.”¹³ In other words, like obscenity, “I know it when I see it.”

AFTE standards acknowledge that these decisions involve subjective qualitative judgments by examiners and that the accuracy of examiners’ assessments is *highly dependent on their skill and training*.... even with more training and experience using newer techniques, the decision of the toolmark examiner remains a *subjective* decision based on unarticulated standards and no statistical foundation for estimation of error rates.¹⁴

The need to substantiate claims of uniqueness in firearms identification (e.g., identifying a cartridge casing as having come from one gun to the exclusion of all other guns in the world) had led to an earlier NAS study, *Ballistics Imaging*, published in 2008, which concluded:

The validity of the fundamental assumptions of uniqueness and reproducibility of firearms-related toolmarks has not yet been fully demonstrated.... Although they are subject to numerous sources of variability, firearms-related toolmarks are not completely random and volatile; one can find similar marks on bullets and cartridge cases from the

same gun...A significant amount of research would be needed to scientifically determine the degree to which firearms-related toolmarks are unique or even to quantitatively characterize the probability of uniqueness.¹⁵ (*emphasis added*)

The NAS report is critical of the lack of scientifically-sound research:

Because not enough is known about the variabilities among the individual tools and guns, we are not able to specify how many points of similarity are necessary for a given level of confidence in the result. Sufficient studies have not been done to understand the reliability and repeatability of the methods...additional studies should be performed to make the process of individualization more precise and repeatable.... the scientific knowledge base for toolmark and firearms analysis is fairly limited.”¹⁶

The NAS Report is also critical of the few studies that have been done in this area:

...capsule summaries [of these studies] suggest a heavy reliance on the subjective findings of examiners rather than the rigorous quantification and analysis of the sources of variability... This is not to say that toolmark analysis needs to be as objective as DNA analysis in order to provide value... [B]ut the protocols for DNA analysis do represent a precisely specified, and scientifically justified, series of steps that lead to results with well-characterized confidence limits, and that is the goal for all the methods of forensic science.¹⁷

Hair Evidence Analysis: Hair analysis may be useful in excluding suspects or narrowing the pool of po-



tential suspects based on class characteristics of hair found at a crime scene. “The results of analyses from hair comparisons typically are accepted as class associations; that is, a conclusion of a ‘match’ means only that the hair could have come from any person whose hair exhibited — within some levels of measurement uncertainties — the same microscopic characteristics, but it cannot uniquely identify one person.”¹⁸

The NAS Report notes that “no scientifically accepted statistics exist about the frequency with which particular characteristics of hair are distributed in the population. There appear to be no uniform standards on the number of features on which hairs must agree before an examiner may declare a ‘match.’”¹⁹

Although it may continue to serve a role in investigation of crimes, because of advancements in forensic Mitochondrial DNA (mtDNA) analysis, the traditional hair analysis we use to see in the courtroom is a thing of the past. And that’s a good thing: Incorrect or unvalidated hair analysis was determined to be involved in 48 of the first 241 Innocence Project exonerations.²⁰

From our standpoint, if the prosecution seeks to introduce hair analysis, the defense must argue against its admissibility. There simply is not adequate empirical data on the frequency of various class characteristics in human hair. Assertions by the prosecution’s analyst that hairs are “consistent” or “similar” are prejudicial, lack probative value, and should not be presented to the jury.

Fiber Evidence Analysis: Analysis of fiber evidence — from 100% wool to cotton, nylon, acrylic, rope, etc. — involves the science of analytical chemistry, with sound, proven scientific credentials. Associating a fiber with a known class of fibers is accepted in the scientific community. In some cases, fibers have been subjected to “relatively distinctive environmental conditions (e.g., sunlight exposure or laundering agents) that impart characteristics that can distinguish particular items from others from the same manufacturing lot. *Fiber examiners agree, however, that none of these characteristics is suitable for individualizing fibers (associating a fiber from a crime scene with one, and only one, source).*”²¹

The report was critical that there were no set standards “for the number and quality of characteristics that must correspond in order to conclude that the two fibers came from the same manufacturing batch. There have been no studies of fibers (e.g., the variability of their characteristics during and after manufacturing) on which to base such a threshold.”²²

The report also criticized the lack of scientifically-based research:

Similarly, there have been no studies to inform judgments about whether environmentally related changes discerned in particular fibers are distinctive enough to reliably individualize their source, and there have been no studies that characterize either reliability or error rates in the procedures. Thus, a “match” means only that the fibers could have come from the same type of garment, carpet, or furniture; it can provide only class evidence.²³

Questioned Document Examination: This section of the report deals

primarily with handwriting comparison.²⁴ This typically involves comparison of handwriting or a signature on a questioned item with handwriting or signatures from an item or items of known origin to determine forgery. Sometimes, it involves someone trying to disguise their own writing.

The NAS committee again noted the paucity of scientifically-based research:

The scientific basis for handwriting comparisons need to be strengthened. Recent studies.... suggest that there may be a scientific basis for handwriting comparison.... Although there has been only limited research to quantify the reliability and replicability of the practices used by trained document examiners, the committee agrees that there may be some value in handwriting analysis.²⁵

The clear implication with this forensic discipline, and others, is that there “may” be a scientific basis, and there “may” be value in this type of evidence, but not yet. Not without further study and research which follows the scientific method.

Paint Evidence Analysis: Have you ever had the situation where your client is a truck painter and the prosecution finds molecules of paint “consistent” with the paint your client works with in the ligature around the neck of a woman he’s suspected of strangling? I hate it when that happens. In and of itself, depending on how common or uncommon that particular paint might be, it might not be enough to convict, or even prosecute. However, if they have your client’s DNA in three or four other similarly situated victims, it could be a real problem.

More commonly associated with car accidents (sometimes referred

to as vehicular homicide or vehicular assault), the transfer of paint from one item to another can indeed provide valuable evidence. Depending on the number of layers of paint on a given item, the value of the evidence increases. Given sufficient distinct layers, an examiner may conclude that “it is unlikely that the questioned paint originated from any source other than that of the known paint.”²⁶ Just where that threshold lies is up to the subjective determination of the examiner. There are no standards or statistical studies.

The report summarizes: “As is the case with fiber evidence, analysis of paints and coatings is based on a solid foundation of chemistry to enable class identification.... However, the community has not defined precise

criteria for determining whether two samples comes from a common source class.”²⁷

Explosives and Fire Debris Evidence Analysis: Like fiber and paint analysis, this discipline is based on well-founded principles of chemistry and physics. However, such evidence must be scrutinized. The report states:

By contrast, much more research is needed on the natural variability of burn patterns and damage characteristics and how they are affected by the presence of various accelerants. Despite the paucity of research, some arson investigators continue to make determinations about whether or not a particular fire was set. However, according to



testimony presented to the committee, many of the rules of thumb that are typically assumed to indicate that an accelerant was used (e.g., “alligatoring” of wood, specific char patterns) have been shown not to be true. Experiments should be designed to put arson investigations on a more solid scientific footing.²⁸

Forensic Odontology — Bite-marks: On July 27, 2009, the State of Wisconsin announced it was dismissing all charges against Robert Lee Stinson. Stinson had been wrongfully convicted of murder based in large part on the testimony of an expert who said the bitemark on the victim matched Stinson’s bite. The expert testified that the process had “zero margin of error.” Forensic DNA evidence proved otherwise. Robert Lee Stinson had served over 23 years.

Sounding a bit like a broken record, the NAS Report again criticizes the discipline for lacking in scientifically-based studies:

there is considerable dispute about the value and reliability of the collected [bitemark] data for interpretation. Some of the key areas of dispute include the accuracy of human skin as a reliable registration material for bite marks, the uniqueness of human dentition, the techniques used for analysis, and the role of examiner bias.... Although the majority of forensic odontologists are satisfied that bite marks can demonstrate sufficient detail for positive identification, no scientific studies support this assessment, and no large population studies have been conducted.... *The committee received no evidence of an existing scientific basis for identifying an individual to the exclusion of all others.*²⁹

Bloodstain Pattern Analysis: The committee found that while bloodstain

pattern evidence can be informative as to what occurred at a crime scene, it cautions about the complexities of the analysis and the many sources of variability. “In general, the opinions of bloodstain pattern analysts are more subjective than scientific.... [e]xtra care must be given to the way in which the analyses are presented in court. The uncertainties associated with bloodstain pattern analysis are enormous.”³⁰

Digital & Multimedia Analysis: The report includes this section on the emerging, dynamic field of forensic computer and digital technology analysis. The rapid proliferation of multimedia devices and other forms of digital technology has given rise to all sorts of new ways to commit crimes and/or leave evidence of committing crimes. The committee states:

Over the past ten years, this process has become more routine and subject to the rigors and expectations of other fields of forensic science. Three holdover challenges remain: (1) the digital evidence community does not have an agreed certification program or list of qualifications for digital forensic examiners; (2) some agencies still treat the examination of digital evidence as an investigative rather than a forensic activity; and (3) there is wide variability in and uncertainty about the education, experience, and training of those practicing this discipline.³¹



Notes

- 1 Committee on Identifying the Needs of the Forensic Science Community, National Research Council, *Strengthening Forensic Science in the United States: A Path Forward* (Washington DC: National Academies Press, 2009), 133
2. *Path Forward*, 134.
3. *Path Forward*, 135.
4. *Path Forward*, 136.

5. *Path Forward*, 136.
6. *Path Forward*, 142.
7. *Path Forward*, 143.
8. *Path Forward*, 145.
9. *Path Forward*, 146.
10. *Path Forward*, 147.
11. *Path Forward*, 149.
12. Subclass characteristics are produced during the manufacturing process by a tool that leaves virtually identical markings on a number of firearms in the same production lot.
13. AFTE Criteria for Identification Committee, “Theory of Identification, Range of Striae Comparison Reports and Modified Glossary Definitions – An AFTE Criteria for Identification Committee Report.” *Journal of the Association of Firearm and Toolmark Examiners* 24, no. 2 (1992): 336-340.
14. *Path Forward*, 153
15. *Path Forward*, 154, citing National Research Council *Ballistics Imaging* (Washington D.C: The National Academies Press) 2008, 3.
16. *Path Forward*, 154.
17. *Path Forward*, 155.
18. *Path Forward*, 156.
19. *Path Forward*, 160.
20. Brandon Garrett and Peter Neufeld, *Invalid Forensic Science Testimony and Wrongful Convictions* 9 Va. L. Rev. 1 (2009). See <http://www.innocenceproject.org> for updated statistics.
21. *Path Forward*, 161
22. *Path Forward*, 162-63
23. *Path Forward*, 163
24. The committee did not research or take testimony on paper and ink analysis, the other common components of questioned document examination. While acknowledging a sound scientific basis in chemistry, the committee did not offer any definitive view on these disciplines.
25. *Path Forward*, 166-67
26. *Path Forward*, 169
27. *Path Forward*, 170
28. *Path Forward*, 173 (footnotes omitted)
29. *Path Forward*, 176 (footnotes omitted) (emphasis added)
30. *Path Forward*, 178-79
31. *Path Forward*, 181