INVESTIGATION · SCIENTIFIC AND TECHNICAL INVESTIGATION

Forensic Science: Strengths, Limits and Reform

Forensic science helps investigators and courts establish what happened, but some forensic methods have proved far less reliable than once claimed. This article explains the scientific basis of the main forensic disciplines, the evidence on their accuracy, the role of bias, and the reforms that followed major reviews.

ModQu research article · 1,732 words · 19 references (APA 7th edition)

Forensic science applies scientific methods to questions that arise in legal proceedings: whose DNA is on a weapon, whether a fingerprint at a crime scene matches a suspect, what caused a death, or whether a document has been altered. For most of the twentieth century, forensic evidence was presented in courts as objective and nearly infallible. Over the past two decades, a series of reviews, research studies and wrongful convictions has shown that the reliability of forensic methods varies greatly, and that even well-established methods can be affected by human error and bias.

This article explains the strengths of forensic science, the evidence about its limits and the reforms that have followed.

Strengths

The strongest forensic methods rest on well-understood science and have been tested extensively. DNA profiling, developed by Alec Jeffreys and colleagues in the 1980s (Jeffreys et al., 1985), compares variable regions of DNA between a sample and a person. Because the frequencies of these variations in populations are known, analysts can calculate how likely a match would be by chance. For a clear sample from one person, the probability of a coincidental match is usually extremely small. DNA evidence has convicted many offenders and has also exonerated many people who were wrongly convicted.

Other disciplines, such as toxicology and many areas of chemical analysis, also rest on established measurement methods with known error rates and quality controls. In these fields, the main risks lie in contamination, laboratory error and the interpretation of results, rather than in the underlying science.

The 2009 and 2016 reviews

In 2009, a committee of the United States National Research Council (NRC) published a landmark report on forensic science. It concluded that, with the exception of nuclear DNA analysis, no forensic method had been rigorously shown to be able to connect evidence to a specific individual or source consistently and with a high degree of certainty (National Research Council [NRC], 2009). It found that many disciplines lacked research on their validity and error rates, that standards and oversight were weak, and that many laboratories were part of police agencies rather than independent.

Saks and Koehler (2005) had earlier argued that forensic identification was undergoing a "paradigm shift" from claims of certainty based on experience towards empirical testing and probabilistic reasoning, following the model of DNA analysis.

In 2016, the President's Council of Advisors on Science and Technology (PCAST) assessed "feature-comparison" methods, in which an examiner compares marks or patterns, and asked whether each had been shown to be valid through appropriate empirical studies.

Scientific validity of feature-comparison methods

Summary of the 2016 assessment by the President's Council of Advisors on Science and Technology

	Foundational validity	Main concern
DNA, single-source samples	Established	Laboratory error and sample handling
DNA, complex mixtures	Limited	Interpretation of mixtures from many people
Latent fingerprints	Established	False positive rate higher than often assumed
Firearms and toolmarks	Not yet established	Too few appropriately designed studies
Footwear marks	Not established	No studies of identifying features
Bitemarks	Not established	Unlikely to be shown valid

Based on PCAST (2016) ModQu.com

Figure 1. PCAST's assessment of six forensic comparison methods. Foundational validity means that appropriate empirical studies show the method can be applied reliably. Diagram by ModQu.

PCAST concluded that DNA analysis of single-source samples and latent fingerprint analysis had foundational validity, though fingerprint analysis had a higher false positive rate than juries might assume. It found that firearms analysis had only one appropriately designed study at that time, and that footwear and bitemark analysis lacked validity (President's Council of Advisors on Science and Technology [PCAST], 2016).

Forensic science beyond criminal courts

Forensic methods are not used only in criminal cases. Civil litigation, insurance claims, regulatory investigations and corporate inquiries regularly rely on document examination, digital forensics, accounting analysis, engineering failure analysis and toxicology. The same questions about validity, error rates and bias apply in these settings, although the standards of scrutiny are often lower because the evidence is less frequently challenged by a well-resourced opponent. Businesses that commission forensic work should ask the same questions a careful court would ask.

Errors and their causes

Research on accuracy has produced important findings. Ulery et al. (2011) tested 169 fingerprint examiners on hundreds of print comparisons and found that false positive identifications were rare, about 0.1 per cent, but that false negatives were much more common, about 7.5 per cent, and that examiners sometimes disagreed with one another and with their own earlier decisions. Cole (2005) documented a series of cases of misidentification in fingerprint evidence, challenging claims that the method had a zero error rate.

The most widely discussed fingerprint error involved Brandon Mayfield, a lawyer in Oregon, whom FBI examiners wrongly identified in 2004 as the source of a fingerprint on a bag connected with the Madrid train bombings. Spanish police later identified the true source. A review by the Justice Department's Inspector General found that the examiners had been influenced by unusual similarities between the prints, by Mayfield's background and by the reluctance of later examiners to disagree with the first identification (Office of the Inspector General, 2006).

Cognitive bias is a major source of error. Dror et al. (2006) showed that experienced fingerprint examiners changed their own earlier conclusions about the same prints when given contextual information suggesting a different answer. Kassin et al. (2013) reviewed evidence of a "forensic confirmation bias", in which knowing about other evidence, such as a confession, influences examiners' interpretation. Murrie et al. (2013) found that forensic psychologists who believed they had been hired by the prosecution scored the same offenders as higher risk than those who believed they had been hired by the defence. Dror (2020) identified several sources of bias in expert decisions, including the case information, reference materials, base rates, organisational factors and the expert's own training and motivation.

Errors also arise in how results are presented. Thompson and Schumann (1987) described the "prosecutor's fallacy", in which the probability of a match by chance is mistakenly presented as the probability that the defendant is innocent. Gill (2014) analysed cases in which DNA evidence was misinterpreted, especially with mixtures and very small amounts of DNA, contributing to miscarriages of justice.

Wrongful convictions

Studies of wrongful convictions show the practical consequences. Garrett and Neufeld (2009) examined trial transcripts in cases of people later exonerated by DNA testing and found that in many of them, forensic analysts had given invalid testimony, overstating the significance of their findings or misstating the science. Garrett (2011) found that flawed forensic evidence was one of the most common factors in the first 250 DNA exonerations in the United States.

In Canada, the Goudge inquiry examined the work of a pathologist, Charles Smith, whose flawed opinions in child death cases contributed to wrongful convictions. It found failures in his expertise and in the oversight system, and made recommendations on the training, accountability and evidence of forensic pathologists (Goudge, 2008).

Reform

These findings have prompted reforms in several areas.

- **Validation and research.** Reviews have called for empirical studies of each method's accuracy, using realistic test materials and examiners who do not know they are being tested (NRC, 2009; PCAST, 2016).
- **Managing bias.** Laboratories increasingly use procedures such as "linear sequential unmasking", in which examiners analyse the crime scene evidence before seeing the suspect's sample and receive only the case information they need.
- **Independence and accreditation.** Reviews have recommended that laboratories be independent of police and prosecutors and accredited to international quality standards. In England and Wales, the Forensic Science Regulator Act 2021 gave a statutory regulator power to set and enforce standards.
- **Better reporting.** Guidelines encourage examiners to report how strongly the evidence supports one proposition over another, rather than asserting certainty (European Network of Forensic Science Institutes [ENFSI], 2015).
- **Stronger gatekeeping by courts.** In the United States, *Daubert v. Merrell Dow Pharmaceuticals, Inc.* (1993) requires judges to assess the reliability of expert scientific evidence, including whether its methods have been tested and have known error rates. In England and Wales, the Law Commission (2011) recommended a reliability test for expert evidence in criminal proceedings, and its proposals influenced later rules of criminal procedure.

Using forensic evidence wisely

For investigators, lawyers and others who rely on forensic results, several practical lessons follow from this research. The first is to ask what a result actually means. A statement that a fingerprint "matches" or that DNA "is consistent with" a suspect needs to be accompanied by an explanation of how often such a result would occur by chance, and of how reliable the method is. A report that cannot answer these questions should be treated with caution.

The second is to consider the whole chain of evidence. Contamination, mislabelling and transfer of material can occur at the scene, during transport or in the laboratory. Modern DNA methods are so sensitive that tiny amounts of material, transferred innocently through touch or shared surfaces, can be detected. The presence of a person's DNA on an object does not by itself show how or when it got there (Gill, 2014).

The third is to protect examiners from information they do not need. Investigators who tell an examiner that a suspect has confessed, or that "we know he did it", risk influencing the result. Passing on only the information needed for the analysis helps keep the result independent of the rest of the investigation (Kassin et al., 2013).

The fourth is to seek independent review when the stakes are high. A second examination by an analyst who does not know the first result is one of the most effective ways to catch errors.

A short example

The following hypothetical case illustrates these points. In a burglary investigation, a partial fingerprint is recovered from a window frame. The investigating officer sends it to the laboratory with a note that a local man with previous convictions has been arrested and is "almost certainly responsible". The examiner reports an identification.

On review, the laboratory manager notices that the request contained information the examiner did not need, and that the print was of low quality. The print is sent, without the case information, to a second examiner, who reports the comparison as inconclusive. The investigation then finds that the man was at work at the time of the burglary, confirmed by time records and video. The laboratory changes its procedures so that case information is filtered before it reaches examiners, and the case is reviewed as a near miss.

Conclusion

Forensic science can provide powerful evidence, particularly where its methods rest on established science with known error rates, as DNA analysis does. But reviews and research have shown that some forensic methods lack scientific validation, that even reliable methods are affected by human error and cognitive bias, and that results have often been overstated in court. Reforms emphasising validation, bias management, independence, careful reporting and stronger scrutiny by courts aim to ensure that forensic evidence is as reliable as it appears.

References

Cole, S. A. (2005). More than zero: Accounting for error in latent fingerprint identification. *Journal of Criminal Law and Criminology, 95*(3), 985–1078.

Daubert v. Merrell Dow Pharmaceuticals, Inc., 509 U.S. 579 (1993).

Dror, I. E., Charlton, D., & Péron, A. E. (2006). Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International, 156*(1), 74–78. https://doi.org/10.1016/j.forsciint.2005.10.017

Dror, I. E. (2020). Cognitive and human factors in expert decision making: Six fallacies and the eight sources of bias. *Analytical Chemistry, 92*(12), 7998–8004. https://doi.org/10.1021/acs.analchem.0c00704

European Network of Forensic Science Institutes. (2015). *ENFSI guideline for evaluative reporting in forensic science*. ENFSI.

Garrett, B. L. (2011). *Convicting the innocent: Where criminal prosecutions go wrong*. Harvard University Press.

Garrett, B. L., & Neufeld, P. J. (2009). Invalid forensic science testimony and wrongful convictions. *Virginia Law Review, 95*(1), 1–97.

Gill, P. (2014). *Misleading DNA evidence: Reasons for miscarriages of justice*. Academic Press.

Goudge, S. T. (2008). *Inquiry into pediatric forensic pathology in Ontario: Report*. Ontario Ministry of the Attorney General.

Jeffreys, A. J., Wilson, V., & Thein, S. L. (1985). Individual-specific 'fingerprints' of human DNA. *Nature, 316*(6023), 76–79. https://doi.org/10.1038/316076a0

Kassin, S. M., Dror, I. E., & Kukucka, J. (2013). The forensic confirmation bias: Problems, perspectives, and proposed solutions. *Journal of Applied Research in Memory and Cognition*, *2*(1), 42–52. https://doi.org/10.1016/j.jarmac.2013.01.001

Law Commission. (2011). *Expert evidence in criminal proceedings in England and Wales* (Law Com No 325). The Stationery Office.

Murrie, D. C., Boccaccini, M. T., Guarnera, L. A., & Rufino, K. A. (2013). Are forensic experts biased by the side that retained them? *Psychological Science*, *24*(10), 1889–1897. https://doi.org/10.1177/0956797613481812

National Research Council. (2009). *Strengthening forensic science in the United States: A path forward*. National Academies Press. https://doi.org/10.17226/12589

Office of the Inspector General. (2006). *A review of the FBI's handling of the Brandon Mayfield case*. U.S. Department of Justice.

President's Council of Advisors on Science and Technology. (2016). *Forensic science in criminal courts: Ensuring scientific validity of feature-comparison methods*. Executive Office of the President.

Saks, M. J., & Koehler, J. J. (2005). The coming paradigm shift in forensic identification science. *Science, 309*(5736), 892–895. https://doi.org/10.1126/science.1111565

Thompson, W. C., & Schumann, E. L. (1987). Interpretation of statistical evidence in criminal trials: The prosecutor's fallacy and the defense attorney's fallacy. *Law and Human Behavior, 11*(3), 167–187. https://doi.org/10.1007/BF01044641

Ulery, B. T., Hicklin, R. A., Buscaglia, J., & Roberts, M. A. (2011). Accuracy and reliability of forensic latent fingerprint decisions. *Proceedings of the National Academy of Sciences, 108*(19), 7733–7738. https://doi.org/10.1073/pnas.1018707108