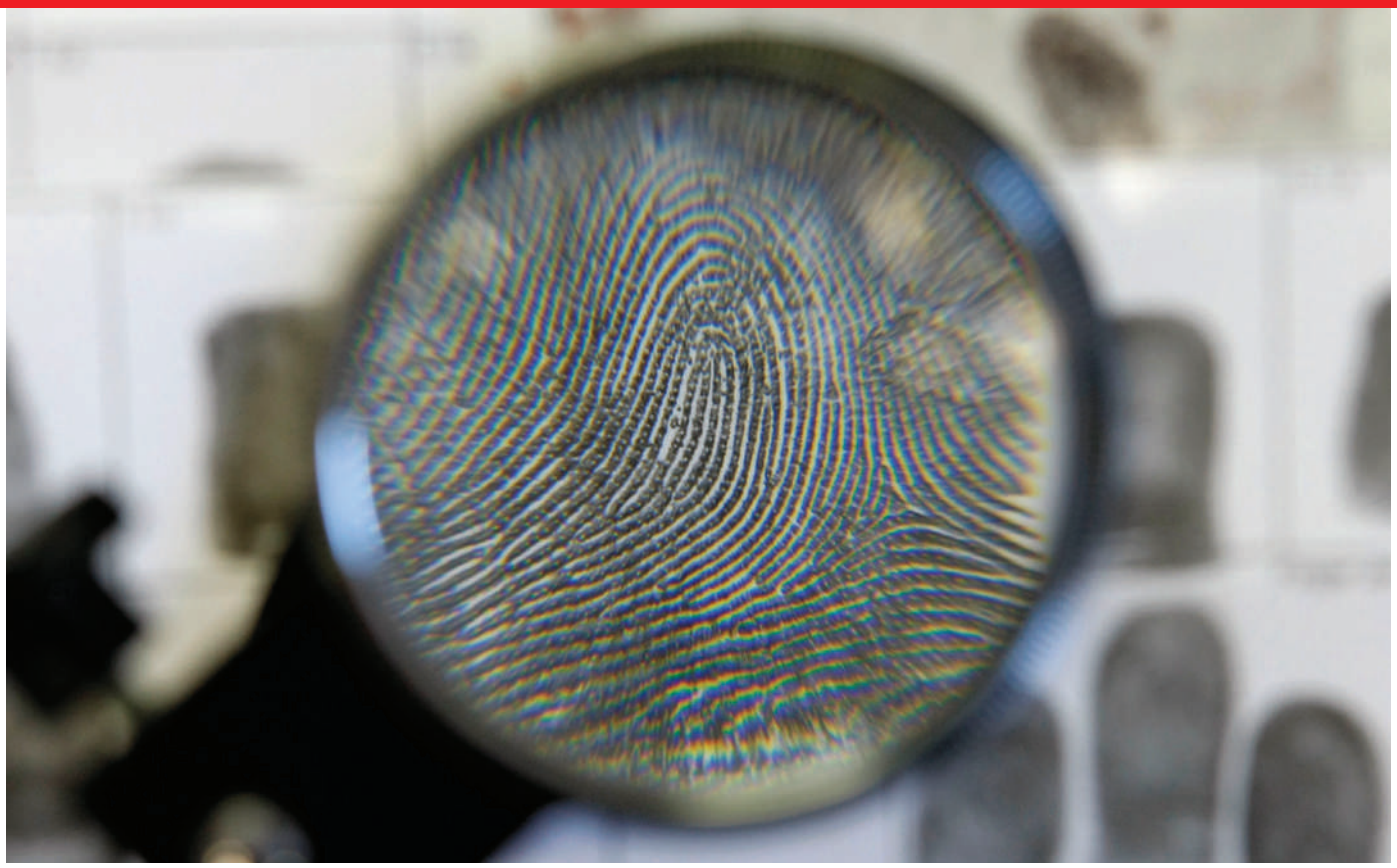


T. KIENZLE/AP

THE FINE PRINT

A single incriminating fingerprint can land someone in jail. But, **Laura Spinney** finds, there is little empirical basis for such decisions.

The terrorist explosions that ripped through Madrid's crowded commuter trains on the morning of 11 March 2004 killed 191 people, wounded some 2,000 more and prompted an international manhunt for the perpetrators. Soon after, Spanish investigators searching the area near one of the blasts discovered an abandoned set of detonator caps inside a plastic bag that bore a single, incomplete fingerprint. They immediately shared the clue with law-enforcement colleagues around the world. And on 6 May 2004, the US Federal Bureau of Investigation (FBI) arrested Oregon lawyer Brandon Mayfield, proclaiming that his print was a match.

Two and a half weeks later, a chagrined FBI was forced to release Mayfield after Spanish police arrested an Algerian national — one of several terrorists later charged with the crime — and found one of his fingerprints to be a much better match. The FBI eventually admitted that it had made multiple errors in its fingerprint analysis[1].

The Mayfield case is a textbook example of 'false positive' fingerprint identification, in which an innocent person is singled out erroneously. But the case is hardly unique. Psychologist Erin Morris, who works with the Los Angeles County Public Defender's Office, has compiled a list of 25 false positives, going back several decades, that is now being used to challenge fingerprint evidence in US courts.

Those challenges, in turn, are being fed by a growing unease among fingerprint examiners and researchers alike. They are beginning to recognize that the century-old fingerprint-identification process rests on assumptions that have never been tested empirically, and that it does little to safeguard against unconscious biases of the examiners.

That unease culminated last year in a stinging report by the US National Academy of Sciences (NAS)[2], which acknowledged that fingerprints contain valuable information — but found that long-standing claims of zero error rates were "not scientifically plausible".

Since then, fingerprint examiners have found themselves in an uncomfortable situation. "How do you explain to the court that what you've been saying for 100 years was exaggerated, but you still have something meaningful to say?" asks Simon Cole, a science historian at the University of California, Irvine.

The only way out of the dilemma is data, says Cole: do the research that will put fingerprinting on solid ground. And that is what researchers are starting to do. In January, for example, the US Department of Justice's research branch, the National Institute of Justice, launched the first large-scale research programme to classify fingerprints according to their visual complexity — including incomplete and unclear prints — and to determine how likely examiners are to make errors in each class. "The vast majority of fingerprints are not a problem," says Itiel Dror, a cognitive psychologist at University College London who is involved in the study. "But even if only 1% are, that's thousands of potential errors each year."

Leaving a mark

Even fingerprinting's harshest critics concede that the technique is probably more accurate than identification methods based on hair, blood type, ear prints or anything else except DNA. Granted, no one has ever tested its underlying premise, which is that every print on every finger is unique. But no one seriously doubts it, either. The ridges and furrows on any given fingertip develop in the womb, shaped by such a complex combination of genetic and environmental factors that not even identical twins share prints. Barring damage, moreover, the pattern is fixed for life. And thanks to the skin's natural oiliness, it will leave an impression on almost any surface the fingertip touches.

The concerns start with what happens after a fingerprint, or 'mark', is found at a crime scene

and sent to the examiners. The problem lies not so much with the individual examiners, most of whom have undergone several years of specialist training, but more with the ACE-V identification procedure they follow in most countries (see graphic). The acronym stands for the four sequential steps of analysis, comparison, evaluation and verification — the hyphen signifying that the last step is carried out by a different individual, who repeats the first three.

The analysis phase starts at the gross level, where there are three main patterns — loops, whorls and arches — that can be used to classify prints or to rapidly exclude suspects. Then comes a second level of analysis, which focuses on finer details, such as bifurcations and ridge endings (see graphic), which are highly discriminating between individuals. If necessary, the examiner can bore down to a third level of detail, related to the shape of ridge edges and the pattern of pores.

Having analysed a mark and noted its distinctive features, the examiner then goes to the comparison step: checking for similarities or differences with a reference fingerprint, or 'exemplar', retrieved from law-enforcement files or taken from a suspect. This part of the process has become increasingly automated, first with the development of automatic fingerprint identification systems (AFIS) in the 1980s, then with the advent of digital print-capture technology in the 1990s. Today's AFIS technology can scan though the vast fingerprint databases compiled by the FBI and other agencies and automatically filter out all but a handful of candidate matches to present to the examiner. The examiner will then winnow the candidates down by eye.

According to the ACE-V protocol, the third step, evaluation, can lead the examiner to one of three conclusions: 'identification', meaning that mark and exemplar came from the same finger; 'exclusion', meaning that they did not, as there is at least one significant difference that cannot be explained by factors such as smearing; and 'inconclusive', meaning that the mark is not clear enough for the examiner to be sure.

"The system as it's designed purposely produces false negatives."

"The system as it's designed purposely produces false negatives," says legal scholar Jennifer Mnookin of the University of California, Los Angeles. Because the protocol makes it possible to have one difference and exclude a match, but a lot of similarities and still not be sure, it builds in a preference for missing the identification of a criminal rather than risking the conviction of an innocent person.

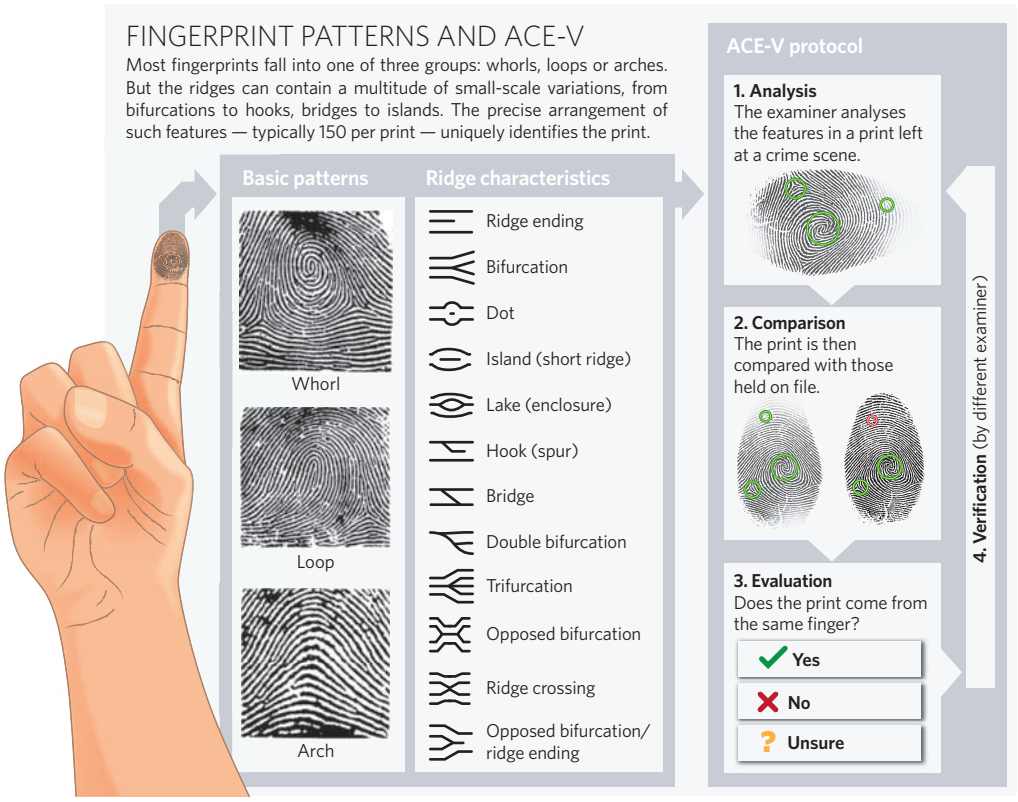
Yet, as the Mayfield case illustrates, false positives can slip through the net. In Scotland, for example, an ongoing inquiry is trying to understand how a fingerprint found at a murder scene was wrongly attributed to police officer Shirley McKie, leading her to be falsely accused of perjury. Such errors may not come to light until some other, incontrovertible piece of evidence trumps the fingerprint, or until the prints are reanalysed in an internal review. But for examiners and researchers alike, the urgent question is why they happen at all.

One of the problems with the ACE-V procedure lies in sloppy execution. For example, the protocol calls for the analysis and comparison steps to be separated, with a detailed description of the mark being made before an examiner ever sees an exemplar. This is to prevent circular reasoning, in which the presence of the exemplar inspires the 'discovery' of previously unnoticed features in the mark. But this separation doesn't always happen, says forensic-science consultant Lyn Haber, who together with her husband, psychologist Ralph Haber, co-authored the 2009 book *Challenges to Fingerprints*. To save time, she says, many examiners do the analysis and comparison simultaneously. The FBI highlighted this as a factor contributing to the Mayfield error.

Misprints

Another problem is that the ACE-V protocol itself is sloppy, at least by academic standards. For example, it calls for the final verification step to be independent of the initial analysis, but does not lay down strict guidelines for what that means. So in practice, the verifier often works in the same department as the first examiner and knows whose work he or she is checking — not a form of 'independence' with which many scientists would be comfortable.

Nor is ACE-V especially strict about what examiners can and cannot know about the case on which they are working. This is especially worrying in light of a study[3] published in 2006 in which Dror and his colleagues showed that both experienced and novice fingerprint examiners can be swayed by contextual information. In one experiment, the researchers presented six examiners with marks that, unbeknown to them, they had analysed before. This time, the examiners were furnished with certain details about the case — that the suspect had confessed to the crime, for example, or that the suspect was in police custody at the time the crime was committed. In 17% of their examinations, they changed their decision in the direction suggested by the information. This point is emphasized by the conclusion in last year's NAS report that "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 many critics this is the central issue: fingerprint analysis is fundamentally subjective. Examiners often have to work with incomplete or distorted prints — where a finger slid across a surface, for example — and they have to select the relevant features from what is available. What is judged relevant therefore changes from case to case and examiner to examiner.

Several research groups are now looking at this problem, with a view to understanding and improving the way that experts make a judgement. The FBI has an ongoing study looking at the quantity and quality of information needed to make a correct decision. Dror's group is doing a controlled study of the errors made by examiners in which they are given marks, told they have been taken from a crime scene — they were actually made by Dror — and asked to identify them.

Expert testimony

Other critics have wondered whether any examiner truly qualifies as an expert. As the Habers point out in their book, examiners rarely find out whether their decision was correct, because the truth about a crime is often not known. As a result, they write, "even years of experience may not improve [an examiner's] accuracy".

Some fingerprint examiners have simply rejected these criticisms. In 2007, for example, the chairman of Britain's Fingerprint Society, Martin Leadbetter, wrote in the society's magazine[4] that examiners who allow themselves to be swayed by outside information are either incompetent or so immature they "should seek employment at Disneyland".

But others have taken the criticisms to heart. After hearing about Dror's research on bias, Kevin Kershaw, head of forensic identification services at Greater Manchester Police, one of Britain's largest police forces, decided to buffer his examiners from potentially biasing information by preventing investigating officers from coming on-site to wait for results, and potentially talking to the examiners about the case. This is made easier by the fact that in Manchester, as in many British police forces, the forensic division is separated from the others. In the United States, by contrast, most of the fingerprint work is done inside police departments — a situation that the NAS report recommended be changed.

Kershaw also invited Dror to come and teach his examiners about the dangers of bias, and he changed his service so that the verifier no longer knows whose work he or she is checking. Finally, as Dror's research indicated that the decisions that are most susceptible to bias are those in which the mark is unclear or hard to interpret, Kershaw introduced blind arbitration in cases in which examiners disagree.

D. RYAN/AP PHOTO

Brandon Mayfield was falsely accused of terrorism on the basis of a fingerprint found at the scene.

Safeguards against bias are relatively easy to put in place, but another potential source of error might be harder to eliminate. It has to do with how faithfully the pattern on a finger is reproduced when it is inked or scanned to create an exemplar. No reproduction is perfect, notes Christophe Champod, an expert in forensic identification at the University of Lausanne in Switzerland. So a mark recovered from a crime scene could match exemplars from more than one person, or vice versa, he says. Exacerbating the problem is the continued growth in the exemplar databases that AFIS have to search.

> **"This is all about adding a culture of science to the forensic-science community."**

Champod thinks that the language of certainty that examiners are forced to use hides this uncertainty from the court. He proposes that fingerprint evidence be interpreted in probabilistic terms — bringing it in line with other forensic domains — and that examiners should be free to talk about probable or possible matches. In a criminal case, this would mean that an examiner could testify that there was, say, a 95% chance of a match if the defendant left the mark, but a one in a billion chance of a match if someone else left it.

To be able to quote such odds, however, examiners would need to refer to surveys showing how fingerprint patterns vary across populations and how often various components or combinations of components crop up. For example, is a particular configuration of bifurcations, ridge endings and the like found in 40% of a given population or in 0.4%? Some research has been done on this issue, but not on a sufficiently large or systematic scale. Nevertheless, Champod is optimistic that a probabilistic system is within reach. Unlike with DNA, he says, which has strong subpopulation effects, fingerprint patterns vary little between populations, simplifying the task.

A probabilistic approach would not replace the examiner or address bias, but it would render the decision-making process less opaque. "Once certainty is quantified, it becomes transparent," says Champod. Ultimately, however, it is for the courts to decide how much weight they accord to fingerprint evidence. The fact that courts still routinely treat it as infallible — which means a single incriminating fingerprint can still send someone to jail — strikes Mnookin as "distressing jurisprudence". Champod, too, would like to see its importance downgraded. "Fingerprint evidence should be expressed by fingerprint examiners only as corroborative evidence," he says. If other strands of evidence limit the pool of suspects, then a fingerprint is much less likely to be misattributed.

To date, judges haven't shown much inclination to alter the status quo. But to be fair, says Barry Scheck, co-director of the Innocence Project — a group in New York that campaigns to overturn wrongful convictions — they haven't been given a viable alternative. The probabilistic approach is not yet ready for court. But that may be about to change if researchers can come up with ways to help fingerprinting profession re-establish itself on a more scientific footing.

A cultural change will also be needed, within both the fingerprint community and the legal system. "This is all about adding a culture of science to the forensic-science community," says Harry Edwards, a senior judge on the District of Columbia circuit and co-chair of the NAS committee that produced last year's report. "From what I have seen, we still have a long way to go." ■

Laura Spinney is a freelance writer based in Lausanne, Switzerland.

1. *A Review of the FBI's Handling of the Brandon Mayfield Case* (Office of the Inspector General Oversight and Review Division, 2006).
2. *Strengthening Forensic Science in the United States: A Path Forward* (National Academies, 2009).
3. Dror, I. E. & Charlton, D. J. *Forensic Identification* **56,** 600–616 (2006).
4. Leadbetter, M. *Fingerprint Whorld* **33,** 231 (2007).

See Editorial, page 325; Opinion, page 351; and online at www.nature.com/scienceincourt.