

The Myth of Fingerprints  
A forensic science stands trial  
By Simon Cole

A COURTROOM LOOKS DIFFERENT WHEN YOU ARE SITTING ON THE WITNESS STAND. The imposing trappings of a medieval institution; the small crowd of silent onlookers; the elation that comes of debating issues one has long considered in isolation, tempered by the claustrophobia of the interrogation and the naked deployment of authority – it all seemed so reminiscent of...well, of my doctoral defense.

When I first chose to write my dissertation on how fingerprint examiners achieve credibility as expert witnesses, I never thought that I would end up an expert witness myself. But there I sat in the witness box, fidgeting nervously, swearing to tell the truth. Here was a role reversal worthy of these postmodern times, one that brought new meaning to the term "participant observation."

But for Byron Mitchell, a quiet, husky, bespectacled bear of a man in his mid-fifties, the proceedings were anything but academic. Mitchell was on trial for allegedly driving the getaway car in the 1991 robbery of an armored vehicle. No one had been hurt, but Mitchell, facing a conviction for armed robbery, was looking at spending most of the rest of his life in prison. The forensic evidence that linked him to the getaway car was the toughest sort to beat: two fingerprints lifted from the steering wheel and the gearshift. Mitchell's only hope was to undermine the credibility of the world's most widely trusted form of forensic proof. It was this endeavor that brought him to a federal courtroom in Philadelphia on a sweltering July morning, where he sat surrounded by four earnest public defenders, a determined prosecutor, some of the country's leading fingerprint examiners, three iconoclastic academics, and a patient but skeptical judge.

Most laypersons regard fingerprint identification as the epitome of science in the service of the law – indeed, as the model to which all other forensic sciences aspire. And yet this isn't quite so. Fingerprint identification puts the legal definition of science, recently articulated in the landmark 1993 Supreme Court decision *Daubert v. Merrell Dow*, to one of its most difficult tests.

"Fingerprint evidence," Arizona State University law professor Michael Saks asserted in 1998, "may present courts applying Daubert with their most extreme dilemma.... A vote to admit fingerprints is a rejection of conventional science as the criterion for admission. A vote for science is a vote to exclude fingerprint expert opinions."

The authority of fingerprint evidence rests on two contested assumptions. Although conventional wisdom since the nineteenth century has accepted the doctrine that no two fingerprints are alike, no one has really proven the proposition's validity. But if the question of the uniqueness of fingerprints seems pedantic, consider a more practical concern: How reliable is fingerprint evidence anyway? Can forensic technicians really match a fragmentary or smudged print taken from a crime scene to one and only one human fingertip, to the exclusion of all others in the world? At a pretrial hearing in the Mitchell case, this important question would receive its first airing in a U.S. courtroom.

FOR HUNDREDS OF YEARS, the uniqueness of fingerprint patterns has been so widely assumed that it never seemed necessary to prove it. As far back as the fourteenth century, the Persian historian Rashid ad-Din, reporting the use of fingerprints as signatures in China, declared that "experience shows that no two individuals have fingers precisely alike." At the beginning of the modern era of fingerprinting, John S. Billings, a U.S. Army physician, declared, "Just as each individual is in some respects peculiar and unique, so... even the minute ridges and furrows at the end of his forefingers differ from that of all other forefingers and is [*sic*] sufficient to identify."

The first person to fashion a statistical foundation for this assumption was the British gentleman scientist Sir Francis Galton. He calculated the probability that any two fingerprints would resemble each other in all particulars as one in sixty-four billion. Galton also noticed that the papillary ridges on the fingertip often end abruptly or split and rejoin one another. In his 1892 book *Finger Prints*, Galton labeled these points at which ridges end or split "minutiae"; they would later be known as "ridge characteristics," "points of similarity," and "Galton details." There are about thirty-five to fifty such points on a typical finger. By correlating the minutiae with one another, an expert could match two fingerprints.

Inspector Charles Collins of Scotland Yard put Galton's observations to work in the famous 1905 Deptford murder trial, in which the brothers Albert and Alfred Stratton stood accused of murdering a shopkeeper and his wife in a London suburb. Collins compared Alfred Stratton's inked fingerprint with a bloody fingerprint found on a cash box and concluded that there were eleven matching "characteristics." The inspector's testimony convinced the jury but not the judge, who, bound by the jury's verdict, "almost apologetically" sentenced the Stratton brothers to hang. How secure was Collins's contention that he had matched the print to Alfred Stratton's finger, to the exclusion of all others? And how many matching minutiae did Collins require in order to reach his conclusion? No one really knew.

Oddly, the leading critic of Collins's methods was a Scottish physician named Henry Faulds – the very man who had first suggested to Galton (via Galton's cousin, Charles Darwin) that "greasy finger marks" might be used to link criminals to their crimes. In a pamphlet published just after the trial, Faulds complained that "the popular fiction, that no two fingers can be alike," was being treated as "a sober fact of the highest scientific certainty." And yet, he pointed out, "the only proof of it is seemingly the same 'fact' repeated in other words – that Scotland Yard by its system of classification has never been able to find two fingers alike." The absence of disproof, in other words, was being treated as proof. Moreover, based on his own experience studying fingerprints, Faulds claimed that "'repeat patterns' in single fingers are often found which come so near, the one to the other, that the least smudginess in the printing of them might easily veil important divergences with appalling results."

Why had Faulds gone from early champion of forensic fingerprinting to outspoken naysayer? To be sure, Scotland Yard's bigwigs had minimized the physician's contribution to the controversial new technique, and Faulds bridled at the slight. But he was also concerned that fingerprinting was slipping out of the hands of men like himself whom he considered "scientists" and into the hands of men like Collins whom he regarded as poorly educated clerks. "Dactyloscopy," as he called it, had once been a subject fit only for professionals like Galton and himself. But now, Faulds complained, it was becoming common "to treat the matter as one which can be dealt with

by subordinate officials untrained in scientific observation."

Faulds was not entirely mistaken: During the first two decades of the twentieth century, control of fingerprinting shifted from scientists like Faulds and Galton to police identification clerks and detectives. In court, these clerks and detectives gave themselves the title FPE, but whether the E stood for "expert" or "examiner" was a subject of intense debate. And yet even as fingerprint identification fell less to scientists and more to cops, courts in the United States and around the world certified the practice as a "science."

Exactly what made fingerprint identification a science was not entirely clear. In 1928, in response to a defendant's demand for evidence "as to the state of this so-called science," the Vermont Supreme Court replied, "No such evidence was required. The subject is one of the things that does not have to be proved." Two years later, the Oklahoma Criminal Court of Appeals announced, "We take judicial knowledge that there are no two sets of finger prints exactly alike." In support of this opinion, the court cited a wide range of published literature – including legal treatises, court cases, criminalistics manuals, the *Encyclopaedia Britannica*, and Mark Twain's novella *Pudd'nhead Wilson* – all of which backed the assertion that there were no two prints alike. In 1941, the Texas Court of Criminal Appeals received a case in which a fingerprint was the only evidence against the defendant. The court undertook a thorough examination of the history of fingerprint identification and decided the time had come to reverse the burden of proof. "Instead of the state being called upon...to offer proof that no two finger prints are alike," the court wrote, "it may now be considered in order for those taking the opposite view to assume the burden of proving their position."

IT MAY BE TRUE THAT ALL FINGERPRINTS ARE UNIQUE – in the sense that, as Gottfried Wilhelm Leibniz argued, all natural objects can be differentiated if examined in enough detail. But forensic fingerprint identification almost never deals in whole fingerprints. Rather, technicians use "latent" fingerprints – invisible impressions that they "develop" using a powder or a chemical developing agent. Latent prints are usually fragmentary, blurred, overlapping, and otherwise distorted. The challenge is to match the latent print to a pristine inked (or, these days, optically scanned) print taken under ideal conditions at the police station.

Today just about everybody, including attorneys, judges, and jurors, seems to accept that FPEs can reliably match latent prints to one and only one source finger. And yet some forensic scientists trained in fields like chemistry and toxicology remain skeptical. They charge that the FPEs avoid statistics, rely on mere hunches, and then proceed to couch their conclusions in terms of absolute certainty. The strongly held belief among FPEs that latent fingerprints can be matched to one person alone, wrote David Stoney in a 1997 legal practice manual, is "the product of probabilistic intuitions widely shared among fingerprint examiners, not of scientific research. There is no justification based on conventional science, no theoretical model, statistics, or an empirical validation process." Detailing the scrutiny visited on DNA typing, Stoney, director of the McCrone Research Institute, an Illinois nonprofit specializing in applied microscopy, declared, "Woe to fingerprint practice were such criteria applied!"

At the time, Stoney's words caught the eye of Robert Epstein, a public defender in Philadelphia. Epstein was helping Byron Mitchell's attorney prepare for Mitchell's second armed robbery trial. Epstein had already succeeded in getting Mitchell's first conviction overturned by the U.S. Court

of Appeals for the Third Circuit in 1998, when a crucial piece of evidence – an anonymous note that had been left on the windshield of the getaway car – had been deemed hearsay. Now the government was preparing to try Mitchell again. With the anonymous note excluded, the crucial evidence would be the two latent fingerprints lifted from the car.

After reading Stoney's article, Epstein realized that an opportunity had arrived: He would challenge the very admissibility of fingerprint identification as scientific evidence. He immediately contacted James Starrs, a professor of law and forensic science at George Washington University. Best known for his efforts to exhume historical figures like Jesse James and Meriwether Lewis to determine whether they had been murdered, Starrs is an enthusiastic proponent and practitioner of forensic science. At the same time, he is a critic of forensic techniques that he considers "garbage," like voice-stress analysis and handwriting identification. Fingerprint matching, he recently concluded, belongs in that category.

When Epstein put Starrs on his witness list, the prosecuting lawyers requested a pretrial hearing to examine Starrs's qualifications; they hoped to dismiss the testimony by casting doubt on the witness's scientific credibility. The defense countered by requesting another pretrial hearing – this one to address directly the question of whether fingerprints constituted reliable scientific evidence. It was the first time such a hearing had ever been held in an American court.

JUST WHAT DOES THE LAW CONSIDER SCIENCE? It's a significant question, because the label the courts accord expert witness testimony can sway jurors from skepticism to sympathy. Until recently, the criterion for scientific evidence was "general acceptance" within a particular field. But this standard, critics pointed out, might apply as easily to astrology as to astronomy.

The Supreme Court reviewed the general-acceptance standard for the first time in the 1993 tort case, *Daubert v. Merrell Dow*, which examined allegations of birth defects caused by the morning-sickness drug Bendectin. Numerous parties, including the American Association for the Advancement of Science (AAAS) and the National Academy of Science, submitted friend-of-the-court briefs in order to help the Court reach a more sophisticated legal definition of scientific evidence. Justice Harry Blackmun's opinion acknowledged the difficulty, well known to philosophers of science, of devising hard-and-fast rules for distinguishing science from pseudoscience. Nonetheless, the Court agreed on four criteria: the testability, or "falsifiability," of the hypothesis; "submissions to the scrutiny of the scientific community" through mechanisms like peer review; the existence of an error rate and standards; and, again, general acceptance. Blackmun had borrowed the crucial notion of falsifiability from the philosopher of science Sir Karl Popper. Science, Popper argued, advances not by proving theories right but by proving them wrong, through a process of elimination. Scientific theories, according to Popper, can be "falsified," or proven wrong, by experiment. Unscientific theories – Marxist dialectical history and Freudian psychology were Popper's favorites – are formed in such a way that they cannot be falsified by data.

In adopting its new criteria, the Court cautioned that they were not strict requirements. They were to be taken only as guidelines for an assessment that, ultimately, was up to the judge's discretion. This flexible decision pleased nearly everyone. Plaintiff and defense lawyers alike read the decision as favorable to their interests, and the AAAS commended the Court for adopting the definition of science preferred by scientists.

But would latent fingerprint identification meet the new standards? If fingerprinting failed the Daubert test, would courts throw out fingerprint evidence, or would they throw out Daubert?

FACED WITH A DAUBERT HEARING, fingerprinting advocates couldn't exactly go out and compare prints for every finger in the world in order to prove the uniqueness hypothesis. Instead, the prosecution argued that the millions of fingerprint comparisons performed by thousands of law enforcement agencies worldwide in the course of case work constituted a kind of testing in itself. Furthermore, assistant U.S. attorney Paul Sarmousakis pointed to Galton's statistical model in his brief, contending that the sheer improbability of fingerprint duplication played to the prosecution's side. Sarmousakis further drew on a twins-based argument advanced by André Moenssens, a professor of law at the University of Missouri at Kansas City. General fingerprint patterns are thought to be genetically influenced, but the details of fingerprints are epigenetic – that is, determined by environmental forces, such as temperature and pressure in the womb. Therefore, although the fingerprints of identical twins are "more similar" than those of fraternal twins or unrelated individuals, they are still demonstrably different. If identical twins don't have identical fingerprints, Moenssens argued, we might assume that no one does.

In its brief, the defense attempted to dispense with all these arguments. The statistical models were flawed, the defense maintained, noting that none had been empirically tested. And Moenssens's twins argument did not address the problem of forensic identification – namely, whether fragments of two different fingerprints might accidentally match. As for the notion that case work itself constitutes empirical testing, the defense countered that such research fails to meet a basic requirement of falsifiability: There is no mechanism, short of extraordinary occurrences like postconviction exoneration, for determining that a latent fingerprint has been erroneously identified.

Perhaps anticipating such challenges, the Federal Bureau of Investigation made an eleventh-hour dash for empirical validation. In February 1999, Stephen Meagher, a supervisory fingerprint specialist at the FBI, sent a letter to all fifty state crime laboratories that began, "The FBI needs your immediate help!" Enclosed were copies of the two latent prints from the getaway car and an inked fingerprint card containing a full set of Mitchell's ten fingerprints. The respondents were to examine the latents in order to determine whether they matched any of the prints on the card. The idea, presumably, was to show that latent print identification was objective by demonstrating that different practitioners would agree on whether or not the prints matched.

But rather than clarifying matters, Meagher's survey added yet another wrinkle to the case. Seven laboratories failed to match one of the latent prints with Mitchell's inked prints, and five failed to match the second latent print. Suddenly, Epstein had a number of qualified FPEs who said that the latents from the getaway car could not be positively identified as Mitchell's prints. The prosecution had a problem.

Meagher quickly sent a follow-up package to all the laboratories that failed to identify one or both of the prints. The package contained better photographs and enlargements of both the original lifted prints and the prints from Mitchell's fingerprint card that were supposed to match them. The enlargements arrived in plastic sleeves on which Meagher had marked corresponding ridge characteristics with red dots. Meagher's letter asked the state examiners to "test your prior

conclusions against these enlarged photographs with the marked characteristics" and to return the results as soon as possible. Almost all the respondents changed their minds and declared that both prints matched.

From the defense's point of view, the FBI's follow-up "survey" was a blatant attempt to intimidate state examiners into changing their opinions. According to ideal protocols, FPEs should not use enlargements to examine latent prints. Moreover, marking up a print for another examiner is considered highly improper. At the hearing, however, Meagher compared the follow-up to a training exercise: "If an individual fails to make an identification that we believe they should have been able to, we would take that information back to that individual, show them the characteristics which they should take into consideration, ask them to reassess their position and, you know, use the information that's now presented to them to try to come up with the same conclusion. That is, that the two prints were identical." The study was not a scientific experiment, Meagher insisted; it was "just a survey."

In any event, guarding itself against the possibility that the survey would fail to sway the court, the FBI used a computer model to clinch its case. This idea had been knocking around law enforcement agencies ever since the mid-1980s, when Automated Fingerprint Identification Systems (AFIS) began to emerge. With millions of fingerprints optically scanned into fingerprint databases, why not have the computer search for matches between them? Such an endeavor might answer Faulds's old criticism – that police agencies had never looked for two very similar fingerprint fragments from different people.

In conjunction with Lockheed Martin, the aerospace company that built its AFIS, the FBI performed what it called a "50K vs. 50K Fingerprint Comparison Test." The test involved running each of fifty thousand single fingerprints through the AFIS against the whole set, and generating a similarity score (a "Z score") for each comparison. The result was that no two different fingerprints generated a Z score anywhere near the Z score for a fingerprint compared with itself.

Could the FBI rest its case? Not so fast. As Epstein revealed in his cross-examination of Donald Zizek, a Lockheed engineer, the test had generated the benchmark Z scores for matches by comparing each fingerprint with itself. But forensic fingerprint identification is supposed to compare two different impressions from the same finger. As Stoney noted, "No two things, no two representations of this person's signature, no two representations of their fingerprint will be exactly alike." The test, therefore, ought to have used two different impressions from the same finger to establish a baseline score for a match. "It is really shocking," Stoney testified, "to see something presented...that doesn't have that basic element of forensic examination in part of the study."

THE GOVERNMENT'S STAR WITNESS at the Daubert hearing was David Ashbaugh, a tall, crisp, and genial Canadian Mountie. An indefatigable writer and researcher, Ashbaugh has more or less single-handedly tried to articulate a scientific foundation for latent fingerprint identification that meets modern standards of "science." In doing so, his body of work has also called some cherished principles of fingerprint identification into question.

For most of the twentieth century, FPEs used the minutiae method Galton devised in the 1890s.

What methodological debates there were centered on how many minutiae, or "points of similarity," ensured a positive identification. Some police departments thought twelve, some thought ten, some were comfortable with eight. By the 1970s, however, many FPEs, aware of the growing sophistication of other forensic sciences, had realized that the various numerical standards were not based on empirical studies but were simply cautionary guidelines meant to preclude misidentifications. In 1973, the professional organization of FPEs, the International Association for Identification (IAI), declared that there was "no scientific basis" for adhering to any minimum standard. Instead, the IAI argued that the criteria for a match should be the examiner's expert opinion.

Since the 1980s, Ashbaugh has expanded upon the IAI's position, sternly opposing the use of point matching in fingerprint identification. Ashbaugh advocates a holistic "quantitative-qualitative evaluation process" that encompasses all available information in latent prints, including points. But what constitutes a match in Ashbaugh's system is not clear. "That final decision," Ashbaugh testified at the Daubert hearing, "is a subjective decision. It's based on your knowledge and experience and your ability."

Not only Ashbaugh but all the government witnesses claimed that point counting was an antiquated method, rendered obsolete by Ashbaugh's work. Indeed, some refused even to use the term "points." At Mitchell's original trial, however, FBI special agent Duane Johnson had shown the jury charts with matching points. And at Mitchell's second trial (held in February 2000, after the hearing), the defense called numerous state FPEs, who emphasized point counts and professed no knowledge of either Ashbaugh or his method. But according to Ashbaugh, even FPEs who think they are counting points are unconsciously doing quantitative-qualitative evaluation. Indeed, the FBI's Meagher insisted that courtroom charts displaying matching points are "just a simplistic way of explaining the identification process to the jury." The point minimums, meanwhile, were "strictly" for "quality assurance,...not for individualization."

In championing Ash baugh's holistic evaluative technique, the prosecution claimed that fingerprint identification was essentially an error-free science. True, erroneous fingerprint identifications had been exposed in the past, and proficiency tests reveal high rates of error by FPEs. But at the hearing, Bruce Budowle, a senior scientist at the FBI laboratory in Quantico, Virginia, who holds a doctorate in genetics and is one of the country's leading authorities on DNA typing, maintained that because FPEs improve and refine their methods in response to exposed gaffes, using these errors "to calculate an error rate is meaningless and misrepresents the state of the art."

Meagher took these claims a step further, contending that the true error rate for fingerprint identification was zero. He did this by distinguishing "methodological error" from "practitioner error," concluding that even if practitioners occasionally made mistakes or operated under less-than-optimal conditions, the methodology of fingerprint identification was beyond reproach. "If the scientific method is followed," Meagher testified, "the error in the analysis and comparative process will be zero."

AFTER THE PROSECUTION RESTED, the defense countered with Starrs. Although recovering from recent surgery, Starrs, leaning on a wooden cane, clearly relished taking the witness stand. He called himself an "eccentric" and an "iconoclast"; he made frequent reference

to his eight children and to his wife, who he said possessed "a fiery Irish temper"; he engaged in repartee with both the judge and the attorneys; and when he suspected the prosecutor of trying to corner him, he said so, declaring, "I will not be trapped." Starrs attacked fingerprinting advocates for "the claim of absolute certainty...the failure to carry out controlled empirical data searching experimentation, a failure to recognize the value of considerations of the error rate." Error, after all, is a necessary accompaniment to measurement and experiment: No academic scientist would claim, or even want to claim, to have a zero error rate. Nor would he wish to calculate an error rate based on ideal conditions not present in actual practice. Starrs further deplored "the lack of objectivity and uniformity and systemization with respect to standards, if any, of the fingerprint analysis." He concluded: "It is a miasma. It just has absolutely no basis at all in solid science."

Starrs was followed by David Stoney and myself. Stoney argued that case work did not constitute testing. "Remember," he testified, "to be science it has to be capable of proving that I'm wrong." But case work provides no mechanism for exposing errors. Bringing to bear my work on the history of fingerprint identification as a forensic tool, I tried to explain how and why the practice's premises had never been tested. But called to the stand to challenge a discipline's claims to expertise, I, predictably, ended up having my own expertise challenged. In the end, Judge J. Curtis Joyner ruled that only fingerprint examiners were qualified to critique fingerprint identification and that neither Starrs nor I would be permitted to testify in front of the jury. Stoney might be allowed, but only as a fingerprint examiner with an opinion on the case prints themselves.

By persuading the judge to bar expert testimony against the scientific standing of fingerprinting, the prosecution effectively won the Daubert hearing. The barring of such testimony, of course, is reminiscent of an argument that has plagued science studies since its inception. Practitioners of the disciplines we study object to our work on the grounds that we don't have enough training in their disciplines to criticize their knowledge claims. But can vigorous skepticism, which courts and philosophers agree should characterize a science, be achieved by a discipline's practitioners alone?

Last summer's Daubert hearing failed to strike down fingerprint evidence, but it raised important issues that continue to spark fierce debate. In the wake of the Mitchell case, somewhere between fifteen and twenty Daubert challenges to fingerprint evidence have been filed around the country. At the same time, conflict continues to divide the fingerprinting community, where Ashbaugh's opponents still defend point standards.

As for Byron Mitchell, at his second trial, the defense subpoenaed all the state examiners who failed to match the latent prints during Meagher's first survey. But the defense's attempt to insinuate that the FBI had intimidated these examiners into changing their opinions apparently did not sway the jury: Mitchell was convicted on all counts. The jury deliberated a little more than two hours. Epstein, however, plans to appeal his case yet again. And the deliberation over the scientific basis of fingerprint evidence, no doubt, has just begun.

Simon Cole is the author of *Suspect Identities: A History of Fingerprinting and Criminal Identification*, forthcoming this spring from [Harvard University Press](#).