Introduction

Imagine that you have been selected to sit on a jury in a criminal trial. An expert takes the stand. He emphasizes his credentials as a senior fingerprint examiner at the Federal Bureau of Investigation (FBI), in Quantico, Virginia, the preeminent crime lab in the country. A series of explosions had killed over 190 people in a terrorist attack. The expert proceeds to confidently describe a forensic hit, comparing a fingerprint of the defendant's to a print found on a bag of detonators in a white van parked near the scene. The expert explains that he studied high-resolution images of the prints on a computer screen, identified fifteen points they shared, and reached a firm conclusion: a "100 percent identification." Next, he asked two experienced colleagues to review the prints: the chief of his unit, and a retired FBI examiner with thirty-five years of experience. Each of the three experts agreed 100 percent with his conclusion.

The judge instructs you, as a juror, to carefully observe all of the evidence in the case. The judge tells you that to convict you must be certain beyond a reasonable doubt that the defendant was the culprit. Would you convict?

"That's not my fingerprint, your honor," says the defendant, in response to this evidence.

"What . . ." the judge responds. "It wasn't your fingerprint?"

"If it is, I don't know how it got there," he insists. "It is not my fingerprint."

In the real case, Brandon Mayfield, a Portland, Oregon lawyer, pleaded for his freedom in an Oregon federal courtroom. Federal agents testified that they identified his fingerprint on a plastic bag with detonators found near the bombing of four commuter trains in Madrid, Spain, that killed 193 people and injured about two thousand more. The judge sided with the FBI and ordered Mayfield detained as a material witness to terrorism. Mayfield knew that he was innocent. He had never set foot in Spain. He had converted to Islam years earlier, and the FBI theorized that perhaps he had formed an allegiance to militant Islamic groups and traveled under a fake name. His case would come to reshape the course of forensics, but only after he faced the prospect of indefinite detention and the death penalty.

Would you convict a person if the only evidence in the case was a fingerprint comparison? What if it was a bite mark, a drug test, or a DNA test? Before making the momentous decision to convict a person, you should ask how reliable the evidence is. Forensic evidence refers broadly to evidence in legal matters that uses scientific methods; my focus in this book is on the wide array of forensic evidence used in criminal cases. What is most surprising is that many forensic examiners do not use methods that are based on solid scientific research. Indeed, with the exception of DNA testing, the experts who link evidence to particular defendants at criminal trials cannot give you a straight answer to the question, "How reliable is your evidence?" Techniques like fingerprinting have been used for over a hundred years. Surely, someone must know how reliable they are. Yet not only is reliability untested and unknown, but the experts do not candidly admit to the judge and the jury that deep uncertainty lies at the foundations of their work. They do not admit that no one has carefully tested the reliability of the methods they use or the work they do on routine cases. Nor do they admit that the crime lab where they work lacks a rigorous testing program. Instead, forensic analysts testify in court just like the actors on popular forensics shows: they claim to find a perfect match. Take, for example, an episode of the popular show *CSI: Miami,* where the investigators, like in Mayfield's case, had just a single fingerprint. Crack

investigator Eric "Delko" Delektorsky looks at the image: "Got a tented arch." Delko runs the print through a computer, which displays a supposed 99.32 percent hit to a person they didn't expect: the victim's fiancé. Moments later, you see arresting officers escorting the fiancé out of the house. Case closed.[1]

The FBI analysts in Brandon Mayfield's case were even more certain: they were "100 percent" certain. They were so certain that when Spanish authorities issued a report with a negative conclusion, contrary to the FBI's fingerprint identification, the FBI fingerprint analysts forcefully disagreed and flew to Madrid, Spain, to present their findings, with blown-up photos illustrating their work. The FBI placed Mayfield under twenty-four-hour surveillance, and then they arrested him. Mayfield's lawyer counseled him that he could be detained indefinitely and might face the death penalty. Then, on May 20, 2004, the prosecutor stood up in court and told the judge something unexpected: that morning the government "received some information from Spain" which "casts some doubt on the identification." Spanish authorities "determined completely" that the print belonged to a known Algerian terrorist. The FBI agreed to release Mayfield, dropped all charges a few days later, apologized to Mayfield, and a federal investigation followed.[2]

Our crime labs need an autopsy. The episode profoundly harmed Mayfield and his family. The failure of these FBI agents brought home how little we know about the reliability of forensic evidence. We need to know why these errors occur. After all, fingerprints have been used in court for over a hundred years. Fingerprint examiners insisted for decades that they had an error rate of zero. If three experienced fingerprints experts could get it so badly wrong, in a high-profile case, then how reliable is fingerprinting? How about all of the other forensics? The problem cried out for a serious scientific inquiry. A mini-autopsy did occur in response to the Mayfield case itself, when the U. S. Department of Justice wrote a lengthy report, identifying specific problems with the work done in Mayfield's case. However, the investigators did not try to answer the most fundamental question that you would want answered if you were sitting on a jury: How reliable is a fingerprint comparison? No scientific studies had been done on the question. The investigators briefly noted in their report that, according to critics, the basic premises of fingerprint

work remain scientifically unproven. They further noted that defense lawyers and academics had questioned whether one can reach a conclusion using fingerprints with absolute certainty, but then left it at that.[3]

A few lone voices, mostly in academia, had raised reliability concerns for years, but without any success in the courts. Particularly when DNA testing became more common in the 1990s, leading scientists, law professors, social scientists, and a few defense lawyers began to ask what research supported the traditional non-DNA forensics. They were ignored. The U. S. Supreme Court issued its landmark *Daubert* decision in 1993, holding that federal judges must act as gatekeepers to ensure that experts use reliable methods. Lawyers expected that judges would finally scrutinize forensic science in court, particularly after many states adopted this federal rule. A few judges hesitantly raised reliability concerns about fingerprinting, but the FBI shut them down with confident assertions that such techniques were foolproof.

Slowly, however, revelations from cases like Mayfield's began to erode the wall of silence that law enforcement, forensic analysts, and prosecutors had built around forensic evidence. Part I of this book describes how lawyers, scientists, and investigators uncovered the full scope of this crisis. It was not only Mayfield who deserved answers. Not long after Mayfield was cleared, Keith Harward, a prisoner in Virginia, began to write letters seeking DNA testing. At trial, dentists claimed his teeth matched bite marks on the victim. He was innocent, but no one seemed to be listening until his letter reached the Innocence Project in New York, founded by lawyers Barry Scheck and Peter Neufeld. Meanwhile, a public defender in Washington, DC, Sandra Levick, had located a string of cold cases in which FBI agents gave highly overstated testimony about forensic hair comparisons. She too began to pursue more modern DNA testing to try to prove innocence in old cases.

Researchers began to ask new questions. Leading statisticians began to ask what probabilities exist for forensics, since there is no such thing as a 100 percent match; every conclusion has some degree of uncertainty. Itiel Dror, a psychologist, began to study the role that cognitive biases play in forensics. Peter Neufeld and I began to examine the trial testimony in hundreds of cases of innocent people freed by DNA testing. We were surprised to find that in over half of the cases, forensic errors contributed to

the original convictions.[4] A few dissenting forensic scientists began to cautiously ask questions from within their professional communities.

Entire crime labs now came under scrutiny. The same year that Mayfield was arrested, in 2004, the entire Houston crime lab was closed due to rampant errors. A well-known former prosecutor, Michael Bromwich, audited this mass disaster. Levick would soon learn that Bromwich had previously audited problematic testimony in FBI hair cases, including in her client's cases. Levick's persistence would trigger a new audit of thousands of old FBI cases. In 95 percent of the cases, FBI experts testified erroneously and misrepresented the reliability of the technique, including in death penalty cases. Other crises were brewing during this time, but without anyone detecting the problem. Sonja Farak and Annie Dookham falsified their work at labs in Massachusetts, until eventually they were caught and forty-thousand-plus cases were overturned. Labs in large and small cities, from Chicago, Illinois, to Cleveland, Ohio, to Amherst, Massachusetts, to entire state crime labs in West Virginia and Montana, all had audits, reviews, and cases reopened. Journalists began to pay attention to stories of botched forensic analysis. Some began to suspect that the authorities might have executed innocent people based on flawed forensics.

Responding to a growing national problem, the U.S. Congress called on the preeminent scientific organization in the country, the National Academy of Sciences, to investigate and report. Federal appellate judge Harry Edwards, who co-chaired the committee of leading scientists, crime lab directors, judges, and lawyers, had, like many others, always assumed that forensics were foolproof evidence. Hearing about hundreds of cases like Mayfield's shocked the lawyers and the scientists in the group. It was as if everything stopped at the national meeting of the American Academy of Forensic Sciences (AAFS) on February 16, 2009, the day the report came out. The three-hundred-page tome could not have been clearer. A single sentence summed it up: "With the exception of nuclear DNA analysis, however, no forensic method has been rigorously shown to have the capacity to consistently, and with a high degree of certainty, demonstrate a connection between evidence and a specific individual or source." What did that mean? Only DNA testing, the scientists said, could reliably connect evidence to individuals. No other forensics were reliable enough to make a

definitive hit. As crucial as forensics can be in criminal cases, much of forensic evidence used is "without any meaningful scientific validation."[5]

As that important report showed, many types of forensic comparisons lack reliability, but the problem becomes even more troubling when one looks at the entire process, from the crime scene, to the lab, and then to the courtroom. Part II of this book explores each of the ways that forensics can go wrong. Although forensic methods have error rates, including false hits and misses, the reliability of experts has rarely been carefully tested, even for long-standing and widely used techniques like fingerprint and firearms comparisons. When researchers do uncover error rates, forensic analysts often do not disclose them in court. We do not know how reliable particular forensics professionals are; most labs do not seriously test them. When judges deem them experts, they take the stand and use overstated language and proclaim infallibility to the jury. Bias affects forensic examiners, who typically work as an arm of law enforcement.

When, despite the 2009 report, little had changed in our labs and courtrooms, a second group of top scientists, led by mathematician and geneticist Eric Lander, came together in response. In 2016 the President's Council of Advisors on Science and Technology (PCAST) issued a report that emphasized a simple message: If we do not know how reliable a forensic technique is, we should not use it until this fundamental question is answered. Lander and his fellow scientists said that some techniques, such as firearms and bite mark comparisons, lack validation. Other techniques, like fingerprint evidence, they found to be valid, but with error rates far higher than many people assume. Again, forensic experts, prosecutors, and judges largely ignored this scientific report. Although few new studies have been done to try to measure error rates, analysts continue to reach unsupported conclusions, prosecutors continue to rely on them, and judges have only grudgingly raised questions about evidence like firearms and bite mark comparisons.

Today, it is more important than ever that we get forensic evidence right. Part III turns from forensic experts to our crime labs. During this time of increasing scrutiny of forensics, crime labs did not shrink but grew in size. Demand for lab services and backlogs in testing have dramatically increased. Crime labs now process millions of requests every year, as more criminal cases depend on forensic testing. The FBI led efforts to expand