Journal of Law and Policy

Volume 13
Issue 1
SCIENCE FOR JUDGES III:
Maintaining the Integrity of Scientific Research and
Forensic Evidence in Criminal Proceedings

Article 8

2005

Fingerprint Evidence

Sandy L. Zabell

FINGERPRINT EVIDENCE

*Sandy L. Zabell, Ph.D.**

I. FINGERPRINTS REVISITED?

Recent years have seen an increasing number of challenges to fingerprint evidence.[1] Given its long standing as the "gold standard" of human identification, this may seem surprising, but there are, in fact, several natural reasons for this unexpected development. The first and most important of these is undoubtedly the spectacular rise to prominence of DNA technologies in the forensic arena. DNA identification has not only transformed and revolutionized forensic science, it has also created a new set of standards that have raised expectations for forensic science in general.[2] Traditional areas of criminalistics that have large

* Professor of Mathematics and Statistics at Northwestern University. I am grateful to Simon Cole for generously sharing his extensive knowledge of this subject, and for providing many helpful references and comments. Thanks also to Margaret Berger, David Kaye, Gregory O'Reilly for helpful comments on an initial draft of the manuscript, and Matthew Tulchin for editorial assistance.

[1] *See* U.S. v. Navarro-Fletes, No. 01-30247, 2002 WL 31420123, at *1-2 (9th Cir. Oct. 24, 2002) (ruling that the district court did not err in admitting fingerprint testimony); U.S. v. Havvard, 260 F.3d 597 (7th Cir. 2001) (affirming that latent fingerprint identification satisfies *Daubert*); U.S. v. Salim, 189 F. Supp. 2d 93, 101 (S.D.N.Y. 2002) (finding that the methodology undertaken by the government's fingerprint witness met the *Daubert* standard); U.S. v. Llera Plaza, 188 F. Supp. 2d 549 (E.D. Pa. 2002) (discussing whether the FBI's fingerprint identification technique satisfied *Daubert*).

[2] *See* National Institute of Justice Investigative and Forensic Sciences, *at* http://www.ojp.usdoj.gov/nij/sciencetech/ifs.htm (last visited Nov. 26, 2004). The Institute states: "The success achieved by DNA evidence in the criminal justice system has raised the bar for all forensic disciplines. The criminal justice community, as well as the general public, now carries the same high

subjective and judgmental components, such as bite mark analysis, are now subject to much greater skepticism and more searching scrutiny.[3]

Even given these new expectations, however, how can fingerprint analysis, so long the paradigm for human identification and so apparently simple in concept, be subject to serious question? The answer to this apparent paradox lies in recognizing the distinction between a latent print (one taken from a crime scene), and a rolled or inked print (a print taken under controlled conditions, such as at a police station). Latent prints may exhibit only a small portion of the surface of the finger and may be smudged, distorted, or both, depending on how they were deposited. For these reasons, latent prints are an "inevitable source of error in making comparisons," as they generally "contain less clarity, less content, and less undistorted information than a fingerprint taken under controlled conditions, and much, much less detail compared to the actual patterns of ridges and grooves of a finger."[4] (See Figure 1).

expectations for all forensic evidence." *Id.*

[3] *See* Flynn McRoberts & Steve Mills, *From the Start, A Faulty Science*, CHI. TRIB., Oct. 19, 2004, at A1. David Faigman, a Professor at the University of California Hastings College of the Law and co-editor of *Modern Scientific Evidence*, remarks, "I think bite marks probably ought to be the poster child for bad forensic science." *Id.*

[4] Lyn Haber & Ralph N. Haber, *Error Rates for Human Latent Fingerprint Examiners*, *in* AUTOMATIC FINGERPRINT RECOGNITION SYSTEMS 341 (Nalini Ratha & Ruud Bolle eds., 2004).

Figure 1: Fingerprint Comparisons[5]

Alex Garcia; Chicago Tribune
Fig. 1(a) – Inked print of Richard Jackson

Alex Garcia; Chicago Tribune
Fig. 1(b) – Latent print said to match the print in Fig. 1(a)

Another important reason for the increased scrutiny of fingerprint evidence is the increasing number of documented misidentifications based on fingerprint analysis.[6] Such misidentifications are of interest for several reasons: they illustrate the subjective nature of fingerprint evidence; they directly contradict a number of claims advanced by the fingerprint profession; and they provide concrete illustrations of just what can go wrong.

[5] Images appear courtesy of the *Chicago Tribune*. The photographs were taken by Alex Garcia and originally appeared in Flynn McRoberts et al., *Forensics Under the Microscope: Unproven Techniques Sway Courts, Erode Justice*, CHI. TRIB., Oct. 17, 2004, at A1.

[6] *See infra* Part II; McRoberts et al., *supra* note 5; Simon Cole, *The Myth of Fingerprints*, N.Y. TIMES, May 13, 2001, § 6 (Magazine) at 13.

II. FINGERPRINT MISIDENTIFICATIONS: TWO CASE STUDIES

A. Commonwealth v. Cowans[7]

On May 30, 1997, an African-American male shot and wounded Officer Gregory Gallagher of the Boston Police Department while that officer was on duty.[8] The assailant's baseball hat fell off during the initial struggle between the two men.[9] Shortly after the shooting, an African-American male holding a gun gained entry into the nearby residence of Ms. Bonnie Lacy.[10] The individual removed his sweatshirt, wiped his gun off, asked for and received a glass of water, and then left.[11]

Officer Gallagher later identified Mr. Stephan Cowans as his assailant in a photographic lineup that included the pictures of eight individuals.[12] The officer also subsequently identified Cowans in a standard lineup that included the suspect.[13] A witness who saw the presumed assailant shortly after the shooting confirmed the identification, although Ms. Lacy did not.[14] In addition to the eyewitness evidence, investigators located a fingerprint on the glass used by the individual who had gained entry to Ms. Lacy's house.[15] The print was matched to that of Mr. Cowans by two fingerprint examiners working for the Boston Police Department.[16] A fingerprint examiner retained by the

[7] *Commonwealth v. Cowans*, 756 N.E.2d 622 (Mass. App. Ct. 2001)

[8] *Id.* at 624.

[9] Jonathan Saltzman, *Judge Sets Bail in DNA Case*, BOSTON GLOBE, Jan. 22, 2004, at B4.

[10] *Cowans,* 756 N.E.2d at 625.

[11] *Id.*

[12] *Id.*

[13] *Id.*

[14] *Id.*

[15] Jonathan Saltzman, *Judge Sets Bail in DNA Case*, BOSTON GLOBE, Jan. 22, 2004, at B4. *See also* Jonathan Saltzman & Mac Daniel, *Man Freed in 1997 Shooting of Officer Judge Gives Ruling After Fingerprint Revelation*, BOSTON GLOBE, Jan. 24, 2004, at A1.

[16] Jack Thomas, *2 Police Officers Are Put On Leave*, BOSTON GLOBE, Apr.

defense later confirmed the fingerprint match.[17] On the basis of this evidence, Mr. Cowans was convicted of shooting a police officer and sentenced to thirty to forty-five years in state prison.[18]

In the pre-DNA world, Mr. Cowans would no doubt have spent much of his adult life behind bars. However, in May 2003 (six years after Cowans's conviction), at the defendant's request, Orchid Cellmark Laboratories performed DNA testing on both the glass and the baseball hat found at the crime scene.[19] The DNA profile found on the glass did not match that of Mr. Cowans, but it did match that of the primary contributor to the DNA on the baseball cap.[20]

In January 2004, at the request of the Commonwealth of Massachusetts, further testing was performed on the sweatshirt. The resulting DNA profile matched the common profile found on the glass and the baseball hat.[21] Initially, Suffolk Assistant District Attorney David E. Meier stated that, given the "compelling" evidence of the fingerprint on the glass, his office would retry Cowans if the conviction were overturned.[22] Two days later, after the fingerprint had been re-examined, however, Meier changed his mind.[23] In addressing Superior Court Judge Peter Lauriat, Meier explained that the fingerprint evidence presented at trial did not match that of Cowans: "I can conclusively and unequivocally state, your honor, that that purported match was a mistake."[24] Mr. Cowans was then released, having spent six years in jail for a

---

24, 2004, at B1 (stating that two fingerprint technicians "testified that a fingerprint lifted from a glass used by the assailant matched Cowans's print").

[17] Interview with James Dilday, Partner, Grayer & Dilday (Apr. 19, 2004).

[18] *See* Saltzman & Daniel, *supra* note 15.

[19] *Id.*

[20] The cap contained a mixture of two or more sources of DNA, but none of these could have originated with Mr. Cowans. *See* Saltzman, *supra* note 9.

[21] Jack Thomas, *'I Was Not the Man Who Did This' Cleared of Shooting Charge, Stephan Cowans Looks Back at Six Years Lost in Prison*, BOSTON GLOBE, Apr. 28, 2004, at F1. *See also* The Innocence Project, Stephen Cowans Case Profile, *available at* http://www.innocenceproject.org/case/display-_profile.php?id=141 (last visited Feb. 23, 2005).

[22] Saltzman, *supra* note 9.

[23] *Id.*

[24] Saltzman & Daniel, *supra* note 15.

crime he did not commit.[25]

B. The Mayfield Affair

On March 11, 2004, a terrorist bomb attack on a Madrid train station resulted in 191 deaths and some 2,000 people injured.[26] The Spanish authorities found a bag of detonators near the site of the explosion with a fingerprint on it that did not match any in their databank.[27] The authorities forwarded the print to several investigative organizations, including the Federal Bureau of Investigation (FBI). After searching its fingerprint database, the FBI located a possible match in the prints of Mr. Brandon Mayfield, an attorney in Portland, Oregon.[28]

From the start, there were troubling aspects about the match. Mr. Mayfield had ties to Muslim individuals and organizations thought to make him suspect, but there was no evidence that he had been out of the country for many years.[29] Nevertheless, the FBI examiners concluded that the print was a "100 percent positive identification," and so informed the Spanish authorities on April 2, 2004.[30]

The Spanish disagreed. On April 13, 2004, the Spanish authorities reported in a memorandum to the FBI that the match was "conclusively negative."[31] Where the FBI found fifteen points of agreement for the fingerprint, the Spanish found only seven.[32]

---

[25] *Id.*

[26] *See* Elaine Sciolino, *Ten Bombs Shatter Trains in Madrid, Killing 192*, N.Y. TIMES, Mar. 12, 2004, at A1.

[27] Susan Schmidt & Blaine Harden, *Lawyer's Fingerprint Linked to Bombing Bag, Detonators Found in Stolen Van in Spain*, WASH. POST, May 8, 2004, at A3.

[28] *Id.*

[29] *Id.*

[30] Robert B. Stacey, *A Report on the Erroneous Fingerprint Individualization in the Madrid Train Bombing Case*, 54 J. FORENSIC IDENTIFICATION 706, 710.

[31] Sarah Kershaw, *Spain and U.S. at Odds on Mistaken Terror Arrest*, N.Y. TIMES, June 5, 2004, at A1.

[32] *Id.*

Nevertheless, the FBI continued to maintain that the latent print on the bag matched that of Mr. Mayfield and arranged a meeting with Spanish officials in Madrid on April 21, 2004 to present their analysis. The meeting did nothing to change the opinion of the FBI and, subsequently, Mr. Mayfield was arrested on May 6, 2004 on a material witness warrant.[33]

Fortunately for Mr. Mayfield, the Spanish authorities persisted with their investigation and, shortly after Mayfield's arrest, announced that they had matched the latent print to an Algerian named Ouhnane Daoud.[34] The final blow came when the Spanish authorities "found traces of Daoud's DNA in a rural cottage outside Madrid where investigators believe the terrorist cell held planning sessions and assembled the backpack bombs used in the attack."[35] Mr. Mayfield was finally released after spending two weeks in jail.[36]

What went wrong? FBI officials initially gave conflicting accounts. In June 2004, *The New York Times* reported on the agency's changing positions:

> F.B.I. officials told Congress members in the briefings last week that they had come up with the match after working off a 'second-generation' digital print—meaning a copy of a copy. But they gave a somewhat different explanation in interviews this week, saying they were now uncertain what generation the digital print represented. But the F.B.I. official who spoke to *The New York Times* on condition of anonymity added that the real issue was the quality of the latent print that the Spaniards originally took from the blue bag.
>
> The determination by an F.B.I. examiner that the print was useable was hasty and erroneous, F.B.I. officials said, and

---

[33] *Id.*

[34] Susan Schmidt, *Oregon Lawyer's Status Remains Murky*, WASH. POST, May 22, 2004, at A2.

[35] Tomas Alex Tizon et al., *Critics Galvanized by Oregon Lawyer's Case*, L.A. TIMES, May 22, 2004, at A13.

[36] David Heath & Hal Bernton, *Portland Lawyer Released in Probe of Spain Bombings*, SEATTLE TIMES, May 21, 2004, at A1.

> set the agency off in the wrong direction and corrupted the rest of the process. [37]

Neither of these conflicting explanations is satisfactory. Latent prints of any type, digitized or not, almost always involve elements of distortion and a loss of information; part of an examiner's claimed expertise is the ability to determine precisely when a latent is useable, and the FBI, in particular, has extensive experience, going back a decade, in the use of digitized prints. [38] The basic rule is supposed to be that if there is any doubt regarding the purported match, an identification is not reported. Either that did not happen in Mayfield, or erroneous identifications are possible, even when examiners have (or express) no doubt about the identification.

Nor is the "quality of the latent print that the Spanish originally took" an adequate explanation for the error, given that Spanish authorities had no difficulty in determining from the outset that the latent did not match Mr. Mayfield's print. Instead, the explanation seems to lie elsewhere. David Ashbaugh, in his highly regarded book *Quantitative-Qualitative Friction Ridge Analysis*, notes:

> Experienced identification specialists have learned through training and practice the limits of how much distortion or difference is still considered within the parameters of agreement. For the benefit of those who do not have much experience, if each area of friction ridge detail being compared requires justification for why the formation appears slightly different or why it is not spatially correct, *be cautious, one may be talking oneself into agreement that is not really there*.[39]

It would appear that a situation similar to the one Ashbaugh describes occurred in the Mayfield case.

In responding to the investigative methods used by the FBI in the Madrid bombing case, Pedro Luis Melida Lledo, head of the

---

[37] Kershaw, *supra* note 31, at A1.

[38] *See* Christopher Brislawn, *Fingerprints Go Digital,* 42 NOTICES AM. MATHEMATICAL SOC'Y 1278 (1995).

[39] DAVID R. ASHBAUGH, QUANTITATIVE-QUALITATIVE FRICTION RIDGE ANALYSIS: AN INTRODUCTION TO BASIC AND ADVANCED RIDGEOLOGY 146 (1999) (emphasis added).

Spanish National Police's fingerprint unit, noted, "They had a justification for everything . . . [b]ut I just couldn't see it."[40] According to Melida, Mayfield's print differed from the latent found on the blue bag both in that the arcs on the lower portion of his print pointed downward rather than upward, as they did in the latent, and in the number of concentric rings (or crests) that were visible to investigators.[41] Melida offered the following analogy for what might have happened to the FBI investigators:

> You're trying to match a woman's face to a picture . . . [b]ut you see that woman has a mole, and the face in the picture doesn't. Well, maybe it's covered up with make-up, you say. O.K., but the woman has straight hair and it's curly in the picture. Maybe the woman in the picture had a permanent?[42]

Other examples of fingerprint misidentifications are known. (The inked and rolled prints in Figure 1, for example, furnish another instance.)[43] As the Mayfield case illustrates, the process of comparing latent and inked prints is inherently subjective and subject to error, and an awareness of the perils implicit in such an approach is a requisite first step in considering possible remedies. Understanding how some of the FBI's most senior examiners could have persuaded themselves that the two prints matched is perhaps best achieved, however, in a more general context, to which we now turn.

---

[40] Kershaw, *supra* note 31, at A1.

[41] *Id.*

[42] *Id.*

[43] *See* Mary Anne Janco, *Murder Case is Formally Dropped: Richard Jackson's Fingerprints Did Not Match Those Found at the Scene*, PHILA. INQUIRER, Mar. 8, 2000, at B1; Anne Barnard, *Convicted in Slaying, Man Wins Freedom: An FBI Investigation Found that Fingerprints at Murder Scene Were Not Those of Richard Jackson*, PHILA. INQUIRER, Dec. 24, 1999, at B1; Mary Anne Janco, *Release of Convicted Killer is Sought*, PHILA. INQUIRER, Nov. 24, 1999, at B1; Rachel Scheier, *New Trial Sought in U. Darby Slaying: A Prosecution Fingerprint Expert's Testimony is in Question*, PHILA. INQUIRER, Aug. 16, 1999, at B1.

III. THE NEED FOR SCIENCE

Because the requirements for scientific rigor are so demanding, let us pause briefly to discuss why they exist. Science is not just one of several competing, equally valid forms of knowledge. Scientific procedures have evolved, and science is accorded great respect precisely because it is recognized that one can only obtain truly reliable knowledge by using its protocols and practices.

Suppose that a scientist is testing a particular procedure having a claimed effect. For example, consider a case in which a new medication or a new surgical procedure claims to be superior to an old medication or surgical procedure. There are two important elements in the design of a scientifically valid experiment to test this. The experiment should employ controls, and the individuals or units being tested should be randomly allocated to the treatment and control groups. The allocation should be random in order to eliminate any possible bias in assigning individuals to the two groups. Ideally, the evaluation should also be *blind*: both the evaluator and the subject should not be told whether the subject is in the treatment or control group.[44] There are many examples in the scientific literature illustrating the importance of properly designed experiments; the following case is particularly instructive.

A. The Portacaval Shunt[45]

The portacaval shunt was a surgical procedure thought for several years to be effective in treating cirrhosis of the liver by redirecting blood flow. After many studies in support of the procedure had been published in medical journals, Grace, Muench, and Chalmers published a review of the studies.[46] Their analysis examined the structure of fifty-one portacaval shunt studies and the conclusions reported in each paper regarding the shunt's

---

[44] The nature of the evaluation may make blinding impossible, however.

[45] *See* DAVID FREEDMAN ET AL., STATISTICS 8-9 (3d ed. 1997). This outstanding textbook contains a number of other instructive examples of this nature.

[46] N. D. Grace et al., *The Present Status of Shunts for Portal Hypertension in Cirrhosis*, 50 GASTROENTEROLOGY 684 (1966).

effectiveness. The results of the fifty-one studies are summarized below.

Design of Experiment	**Degree of Enthusiasm for Procedure**		
	Marked	Moderate	None
No Controls	24	7	1
Nonrandomized controls	10	3	2
Randomized controls	0	1	3

The results are instructive when enthusiasm for the shunt procedure is compared to whether or not the design of the experiment employed controls and, if so, which type. It is apparent from the table that the procedure has little, if any, value: when controls were employed and allocated in a random fashion to treatment and control groups, three of the four studies with controls found no evidence of the procedure's effectiveness. Most of the studies, however, were poorly designed: thirty-two of the fifty-one did not employ controls at all, and fifteen of the nineteen that did use controls failed to use randomization to allocate subjects to the treatment and control groups. As a result, most of the studies got it wrong: 75% of the studies without controls expressed marked enthusiasm for the procedure, as did 66% of those that employed nonrandomized controls.

What is going on here? In brief, it is a well-understood phenomenon that *the human mind often sees what it expects to see, hopes to see, or wants to see*.[47] The reason science requires rigorous standards of testing is precisely because strongly held beliefs and expert opinions do not always withstand scrutiny or comparisons to serious objective competitors. This is not simply a case of science versus non-science. In the past, spectacular

[47] It should be stressed that this is a separate phenomenon from deliberate fraud. In the phenomenon we are discussing, there is deception at work, but it is self-deception.

examples of nonexistent phenomena, such as N-rays, polywater, and cold fusion, have been reported and maintained by many investigators until serious peer review revealed the claims to be erroneous.[48] The failure to follow proper procedures, even or especially by scientists, can have serious consequences.

B. Clinical v. Statistical Prediction

A related phenomenon is that of clinical versus statistical prediction. A classic study performed by Paul Meehl found that the predictions of experts based on clinical interviews were often far less accurate than predictions based on statistical or actuarial instruments.[49] Although experts are often more comfortable relying on their instincts, this reliance does not always translate into superior predictive ability.[50]

A related point is the distinction between the scientific uses of the terms "reliability" and "validity." A measurement procedure is said to be *reliable* if, when repeated, it gives rise to the same result. It is, in other words, consistent. In contrast, a procedure is said to be *valid* if it measures what it claims to measure. For example, in the popular Analysis, Comparison, Evaluation, and Verification (ACE-V) paradigm for fingerprint identification,[51] the verification stage, in which a second examiner confirms the assessment of the original examiner, may increase the consistency of the assessments. But while the verification stage has implications for the reliability of latent print comparisons, it does not assure their validity.

---

[48] *See, e.g.,* Irving Klotz, *The N-ray Affair,* SCIENTIFIC AMERICAN, May 1980; FELIX FRANKS, POLYWATER (1981); GARY TAUBES, BAD SCIENCE: THE SHORT LIFE AND WEIRD TIMES OF COLD FUSION (1983).

[49] PAUL MEEHL, CLINICAL VERSUS STATISTICAL PREDICTION: A THEORETICAL ANALYSIS AND REVIEW OF THE EVIDENCE (Jason Aronson ed., Univ. of Minn. Press 1996) (1954).

[50] Anne M. Heinz et al., *Sentencing By Parole Board: An Evaluation,* 67 J. CRIM. L. & CRIMINOLOGY 1, 1-31 (1976) (discussing the phenomenon in the context of parole decisions).

[51] *See,* ASHBAUGH, *supra* note 39, at 196.

IV. FINGERPRINT OPINIONS

Suppose that a latent print and a rolled print are judged to be a match by an examiner. Such fingerprint identifications can be thought of as implicitly making several assertions based on the knowledge and experience of the examiner. First, through an identification an examiner suggests that the latent print and the rolled prints could have a common source; that is, although no pair of latent and rolled prints is ever identical, the examiner concludes on the basis of his knowledge and experience that the differences between the latent and rolled prints fall within the expected or possible range of variation. Second, an identification indicates that the latent and rolled prints exhibit a degree of similarity substantially greater than one ordinarily sees in pairs of prints coming from different individuals. Such a judgment in principle also reflects the accumulated knowledge and experience of the examiner, even though, in practice, an examiner might not have actually seen very many "near misses." Lastly, the examiner asserts that only one person in the world could be the source of the latent print.

This last assertion is problematic because, unlike the first two, it is difficult to determine what precise knowledge or experience on the part of an examiner could serve as the basis for such a conclusion. Ashbaugh describes the process:

> An opinion of individualization is unambiguous. The details in both prints are in agreement and, in the opinion of the identification specialist, there is sufficient uniqueness [sic] present in the friction ridge detail to eliminate all other possible donors. This opinion is subjective and it is based on the knowledge and experience of the examiner.[52]

There are obvious questions about such an approach. Suppose that the fingerprint examiner is wrong and that, rather than there being a unique "match," instead 1 in 1,000,000 persons have fingerprints bearing as great a degree of similarity to a particular latent print as the particular fingerprint in question. How could an examiner distinguish between these two scenarios based on his knowledge

[52] *Id.* at 146.

and experience? The world would look basically the same; people do not usually see events that happen only 1 in 1,000,000 times. How does an investigator, based on personal experience, discriminate between 1 in 1,000,000 and zero? The examiners in Mayfield thought they knew, but obviously they did not.[53]

"How much correspondence between two fingerprints is sufficient to conclude that they were both made by the same finger?"[54] David Stoney, a distinguished expert on fingerprints, tells us:

> An adequate answer to [this question] is not currently available. The best answer at present to the question 'How much is enough?' is that this is up to the individual expert fingerprint examiner to determine, based on that examiner's training, skill, and experience. Thus we have an ill-defined, flexible, and explicitly subjective criterion for establishing fingerprint identification . . . .
>
> Any unbiased, intelligent assessment of fingerprint identification practices today reveals that there are, in reality, no standards. That is, the amount of correspondence in fraction ridge detail that is necessary for a conclusion of identity has not been established.[55]

A declaration that a fingerprint match has occurred is essentially a statistical statement. As the detail available in a latent fingerprint increases, fewer and fewer people will have consistent fingerprints: first, 1 in 10, then 1 in 1000, and, finally, 1 in 1,000,000. Based on intuition, the examiner thinks he knows where to draw the line, but the question remains: what is the justification for such a judgment?

---

[53] The advent in recent years of computer-searchable databanks of digitized prints may well have played some role here; a match criterion unlikely to generate false positives when used by an examiner fifty years ago while searching a databank numbering in the tens of thousands might well generate false positives when searching a modern database numbering in the tens of millions.

[54] David A. Stoney, *Measurement of Fingerprint Individuality, in* ADVANCES IN FINGERPRINT TECHNOLOGY 327, 329 (Henry C. Lee & Robert E. Gaensslen eds., 2d ed. 2001).

[55] *Id.* at 329-30.

A. Comparing Identification Techniques: Fingerprints v. DNA Analysis

The scientific process for analyzing DNA stands in contrast to the intuitive process for identifying fingerprint matches. A comparison of the two techniques is instructive. Consider the following steps in the presentation and analysis of both types of evidence:

DNA	**Fingerprints**
DNA is unique	Fingerprints are unique
Only a portion of genome examined	Latent only a partial print
Thirteen loci examined	Friction ridge detail identified
Statistical calculation	Subjective judgment

Because the totality of each person's DNA is unique, examining the human genome can potentially enable us to identify a person. But practical considerations of time and expense limit us to examining only a small portion of the genome. Current forensic practice usually entails looking at a total of thirteen loci (locations on the twenty-three pairs of chromosome found in human cells), and the degree of rarity of the resulting DNA profile is scientifically determined by means of a statistical population genetic calculation.

There is an obvious parallel between DNA and fingerprint analysis—up to a point. It is claimed that fingerprints, the friction ridge detail on the surface of the skin, have the theoretical ability to uniquely identify a person.[56] Unfortunately, however, fingerprint examiners are often faced with the Aristotelian reality of a latent

---

[56] *See, e.g.,* FRANCIS GALTON, FINGERPRINTS (1892); STEPHEN M. STIGLER, *Galton and Identification by Fingerprints, in* STATISTICS ON THE TABLE 131-40 (1999).

print rather than the Platonic ideal of a pristine inked print. The analysis of latent prints rests on an examination of Galton points of comparison (or, more generally, "friction ridge detail")—the fingerprint analogues of the short tandem repeat (STR) loci used in modern forensic DNA analysis. It is here that the apparent parallel between fingerprint analysis and DNA analysis intersects reality. In contrast to the scientifically-based statistical calculations performed by a forensic scientist in analyzing DNA profile frequencies, each fingerprint examiner renders an opinion as to the similarity of friction ridge detail based on his subjective judgment.[57]

In order to better understand this fundamental difference in the two identification procedures, it is important to consider in detail the steps involved in each. The next section discusses the statistical calculation of DNA profile frequencies and compares the steps in that process to the corresponding steps in fingerprint analysis.

V. COMPUTING DNA PROFILE FREQUENCIES

Current forensic DNA technology in the United States is commonly based on the determination of a thirteen-locus DNA profile. A typical example of such a DNA profile is given on the next page:

---

[57] The statistical calculation of a DNA profile involves the multiple steps of collecting a database, estimating allele frequencies using the database, testing the database for Hardy-Weinberg and linkage equilibrium, using the so-called "product rule" to compute a combined profile frequency, using population genetic adjustments to correct for possible population substructure, and sometimes even attaching confidence limits to the resulting estimate.

Table 1: A Typical Thirteen Locus DNA Profile

Locus	Profile
D3S1358	16, 18
VWA	15, 20
FGA	24, 26
D8S1179	14, 15
D21S11	28, 30
D18S51	10, 16
D5S818	10, 13
D7S820	11, 11
D13S317	12, 15
D16S539	9, 14
THO1	6, 8
TPOX	6, 8
CSF1PO	7, 8

Forensic scientists determine DNA profiles and profile frequencies using a number of steps.[58] First, we know through research that at various points on the chromosomes, short sequences of nucleotides are repeated in tandem, one after another. These sequences are known as short tandem repeats (STRs). The exact number of repeats varies from one individual to another, but is constant for a given person. For example, consider the locus designated D3S1358. In this example, we are concerned with the DNA on chromosome 3 (hence the "D3"). The DNA profile for an individual at this location (or "locus") might be 16, 18, indicating that at one of the many possible locations where short tandem repeats occur on chromosome 3, here, at location number 1358, one of the chromosomes of the pair contains the repeated sequence 16 times, and the other, 18 times. The 16 and the 18 are determined objectively using computer software that generates an electropherogram, a printout that looks somewhat like an EKG and exhibits a sequence of peaks. The presence of a peak indicates the presence of a repeat of a given size.

After the DNA profile of interest has been established, a previously collected database of DNA profiles is consulted and the frequencies of occurrence of the different repeat sizes (or "alleles") in the DNA profile of interest are determined. (The term "allele" refers to a possible state of the DNA at a given locus. In the example above, the 16 and 18 are two different alleles.) For example, in the case of locus D3S1358, an FBI database indicates that among 203 Caucasian individuals having a total of 406 alleles, the 16 allele occurs 94 times and the 18 allele occurs 66 times. Thus, the frequency of the 16 allele among Caucasians in the United States is estimated to be 94/406, or 23%, and the frequency of the 18 allele is estimated to be 66/406, or 16%.

The frequency of the joint occurrence of the 16, 18 "genotype" is then estimated using a formula that states that the frequency is twice the product of the individual allele frequencies (2 x 23% x

---

[58] *See* JOHN M. BUTLER, FORENSIC DNA TYPING: BIOLOGY AND TECHNOLOGY BEHIND STR MARKERS 25 (2001); IAN W. EVETT & BRUCE S. WEIR, INTERPRETING DNA EVIDENCE: STATISTICAL GENETICS FOR FORENSIC SCIENTISTS (1998).

16%, or approximately 7%).[59] The use of this formula is justified by a finding that the population is in at least approximate "Hardy-Weinberg" equilibrium.[60] The specifics of "Hardy-Weinberg" equilibrium will not be discussed here, but it is important to note that the presence or absence of equilibrium, at least approximately, is a fact amenable to scientific determination by the use of appropriate statistical tests. Because there was some initial controversy regarding the forensic use of DNA, the National Academy of Sciences, through its operating arm the National Research Council (NRC), convened two panels that issued reports (one in 1992 and one in 1996) on the subject. One of the key findings of the second NRC panel report on forensic DNA is that the use of the above product formula (subject to certain caveats and population genetic adjustments, depending on the circumstances) is scientifically justified.[61]

Similarly, in the case of the vWA locus, the 15 allele occurs among Caucasians approximately 44/392, or 11% of the time, and the 20 allele occurs 4/392, or 1% of the time. Thus, the 15, 20 profile occurs 2 x 11% x 1% = 0.22% of the time, or approximately 1 in every 455 times.

Finally, the examiner must determine how unusual it is for a person to have the overall profile at all thirteen loci. In order to compute this, an examiner uses a second product rule that requires the multiplication of profile frequencies at each separate locus. For example, if an examiner knew only the profiles for D3S1358 and vWA, he would multiply the frequencies for both the 16, 18 at D3S1358 and the 15, 20 at vWA, that is, 0.22% x 7% = 0.0154%, or about 1 in 6500 times. The scientific justification for this formula lies in a determination that the population is in "linkage," or gametic, equilibrium. Once again, this fact is amenable to scientific determination and the second NRC report concludes that,

---

[59] *Twice* because there are two possible ways the profile could arise, depending on which alleles the mother and the father contribute. Thus, the mother could contribute the 16 and the father the 18, or vice versa.

[60] BRUCE S. WEIR, GENETIC DATA ANALYSIS II: METHODS FOR DISCRETE POPULATION GENETIC DATA 92-103 (1996).

[61] NATIONAL RESEARCH COUNCIL, THE EVALUATION OF FORENSIC DNA EVIDENCE (1996).

subject to certain caveats alluded to earlier, the formula's use is scientifically justified.[62]

Certain elements of approximation occur in the above calculations; for example, profile frequencies are based on estimates derived from samples, the populations in question may not be in complete equilibrium, and relatives might be possible contributors. Adjustments and corrections are available to account for all of these possibilities.

In contrast, the fingerprint examination process enjoys none of the statistical qualities of DNA analysis. "Galton points," or elements of friction ridge detail, are not objectively determined quantities. For example, as noted above, in the Mayfield case, the FBI claimed fifteen matching points, while the Spanish found only seven.[63] Indeed, as forensic scientists Christophe Champod and Ian Evett emphasize, point requirements only provide an illusory form of transparency to the identification process.[64] This lack of objectivity in the determination of points formed an important part of the argument in the influential 1996 paper of Evett and Ray Williams, a paper that played a key role in leading the United Kingdom to eventually abandon its sixteen-point standard.[65]

Similarly, there are no analogues in the realm of fingerprint evidence to the collection of databases for the purpose of computing allele and DNA profile frequencies. Although examiners have a general sense of which fingerprint characteristics are common and which are rare, the judgments inherent in this work are not based on published statistical studies and methods do not exist for estimating the rarity of latent prints. Instead, as noted by Stoney, the identification process appeals primarily to the experience, training, and expertise of the examiner.

Consider, for example, the following candid admission by Robert D. Olson, Sr.:

---

[62] *Id.*

[63] Kershaw, *supra* note 31, at A1.

[64] Christophe Champod & Ian W. Evett, *A Probabilistic Approach to Fingerprint Evidence,* 51 J. FORENSIC IDENTIFICATION 101, 101-22 (2001).

[65] *See* Ian W. Evett & Ray L. Williams, *A Review of the Sixteen Point Fingerprint Standard in England and Wales,* 46 J. FORENSIC IDENTIFICATION 49, 49-73 (1996).

> All training programs for latent print examiners place great emphasis on practical experience during the phases of instruction regarding the evaluation and comparison of latent prints. This emphasis is well founded and has considerable merit, but it has been stressed so heavily that little written information exists regarding the methodologies and procedures for making a comparison of two prints. This lack of information has resulted in . . . the failure of many persons to recognize the scientific nature of latent print identification.[66]

The last sentence of the quotation is a contradiction in terms. Olson describes a master-apprentice form of mentoring and, whatever else such an arrangement might be, it is not science. The failure to publish specific "methodologies and procedures" is totally antithetical to good scientific practice and renders impossible any form of critical external review.

VI. UNIQUENESS

It is often asserted that fingerprints are unique. This is really asking the wrong question. A comparison to a similar issue in the interpretation of DNA evidence is instructive. Before DNA evidence is introduced in court, an expert sometimes states that every person's DNA (except that of identical twins) is unique to that person. Such a statement is true, but also misleading. The current systems used to type DNA, as noted earlier, examine only a very small portion of the human genome. It is not the unrealized potential of the entire genome, but the statistical calculations regarding the small fraction examined that informs us about the strength of the evidence. Similarly, the issue in fingerprint identification is not whether the surface of the human finger theoretically contains enough information to permit unique identification, but whether the latent print being used in a specific

[66] Robert D. Olson, Sr. & Henry C. Lee, *Identification of Latent Prints, in* ADVANCES OF FINGERPRINT TECHNOLOGY 41-61 (Henry C. Lee & Robert E. Gaensslen eds., 2d ed. 2001).

case is sufficient to arrive at such a conclusion.[67]

Although there is a substantial literature on the uniqueness of fingerprints,[68] it is surprising how little true scientific support for the proposition exists. "From a statistical viewpoint, the scientific foundation for fingerprint individuality is incredibly weak."[69] Several studies have attempted to propose statistical models for fingerprint identification. The most recent of these studies is an FBI-sponsored study performed by Meagher, Budowle, and Ziesig (MBZ).[70] In the MBZ study, the digital images of 50,000 fingerprints were selected and all possible pairwise comparisons were made (including the comparison of an image with itself), using an unspecified quantitative measure of similarity. In each case it was found, not surprisingly, that the image of a fingerprint was far more similar to itself than any of the others and the probability of a fortuitous match occurring was estimated based on the similarity score used by MBZ. The MBZ study concluded: The probability of any non-mate fingerprint being identical to any

---

[67] Of course, the answer to the first question is not irrelevant because it tells us whether an affirmative answer to the second question is even possible.

[68] *See, e.g.,* Galton, *supra* note 56; Stoney, *supra* note 54; David H. Kaye, *The Non-Science of Fingerprinting: United States v. Llera-Plaza,* 21 QUINNIPIAC L. REV. 1073 (2003).

[69] Stoney, *supra* note 54, at 383. Stoney's paper provides an outstanding and comprehensive review of the existing literature on the uniqueness of fingerprints. However, Stoney's conclusions about the nature of these past studies are depressing. Stoney writes:

> Beginning with Galton and extending through Meagher, Budowle, and Ziesig, there have been a dozen or so statistical models proposed. These vary considerably in their complexity, but in general there has been much speculation and little data. Champod's work is perhaps the exception, bringing forth the first realistic means to predict frequencies of occurrence of specific combinations of ridge minutiae.

Id.

[70] For detailed discussions of the MBZ study, see Stoney, *supra* note 54, at 378-83; David H. Kaye, *Questioning a Courtroom Proof of the Uniqueness of Fingerprints*, 71 INT'L STAT. REV. 521 (2003); Simon A. Cole, *Grandfathering Evidence: Fingerprint Admissibility Rulings from Jennings to Llera Plaza and Back Again*, 41 AM. CRIM. L. REV. 1226–1231 (2004) [hereinafter Cole, *Grandfathering Evidence*].

particular fingerprint is $1/10^{97}$ (1 followed by 97 zeroes).[71]

This FBI-sponsored study has been justly attacked on several grounds. First and foremost, even if one were to compare rolled prints rather than latent prints, the appropriate comparison would be the degree of similarity between two different prints from the *same* individual versus two prints from different, randomly selected individuals. As Professor David Kaye notes:

> [T]he study merely demonstrates the trivial fact that the same two-dimensional representation of the surface of a finger is far more similar to itself than to such a representation of the surface of a finger from any other person in the data set. As such it ignores the fact that two prints rolled successively from the same individual are not identical . . . .[72]

Dr. Stoney also criticized the FBI study on similar grounds, noting that "[t]here would never be an occasion to compare a single fingerprint to itself."[73]

Obviously the more relevant study would have been to compare instead *distinct* pairs of prints coming from 50,000 different individuals, each individual contributing two prints. This could be thought of as modeling a search of the Automated Fingerprint Identification System (AFIS) database of inked prints using the inked print of a suspect whose prints might be in AFIS under a different name due to a prior offense. However, the still far more informative study would have been to compare latent prints to target rolled prints.

---

71 Stoney, *supra* note 54, at 379.

72 Kaye, *supra* note 68, at 527. David Stoney agrees with Kaye:

> The Meagher, Budowle, and Ziesig experiments do not include comparing two different prints of the same finger with each other. All actual fingerprint comparisons have this aspect, and every fingerprint examiner knows that no two impressions from the same finger are exactly alike . . . .
>
> The model, therefore, makes an extremely elementary, fundamental mistake.

Stoney, *supra* note 54, at 380-81.

73 Stoney, *supra* note 54, at 380-81.

In fact, the FBI study did attempt to address the latent versus inked print issue by using cropped versions of the inked prints as surrogates for latent prints. Once again, this aspect of the FBI study has been harshly criticized; the use of a portion of an inked print to represent a latent print is totally misleading because latent prints in general suffer not only from a loss of information, but also distortions in the deposited image.[74] In his critique of the FBI study, Dr. Stoney explained that it "employed the thoroughly discredited practice of using subsets of inked prints to simulate latent prints."[75] For this and other reasons, Dr. Stoney concluded that the study was deeply flawed and "highly misleading" and concluded that "it [was] remarkable that a study with such fundamental flaws was presented in court."[76]

Both Professor Kaye and Dr. Stoney are far from alone in their harsh criticism of the FBI study. Champod and Evett, for example, declared "entirely unsupportable" the study's conclusion that the probability of a random print being identical to a particular fingerprint is 1 in 10^{97}, commenting, "The figure of 10^{-97} [computed by MBZ] so transcends reality that we are amazed that it was admitted into evidence."[77]

Curiously, as Professor Kaye notes, an initial oversight on the part of the authors of the MBZ study inadvertently permitted one to compare a small number of rolled prints from the same individual.[78] This was because it turned out that the 50,000 prints did not, in fact, all come from 50,000 different individuals; in a

---

[74] This loss of information is due to the fact that only a portion of the finger may leave a print.

[75] Stoney, *supra* note 54, at 382-83.

[76] *Id.* Stoney elaborates on his criticism: "It was specifically designed to "prove the uniqueness" of fingerprints in a *Daubert* hearing, and incorporates a profound ignorance of both forensic science and statistics. Perhaps the most remarkable aspect of these experiments is that they continue to be introduced in such hearings." *Id.*

[77] *See* Champod & Evett, *supra* note 64. For statistical aficionados, the MBZ calculation makes the absurd assumption that one can accurately calculate probabilities more than twenty standard deviations in the tail of an approximately normal distribution.

[78] Kaye, *supra* note 68, at 1079.

small number of cases, some of the prints were, in fact, duplicate prints taken from the same individual. These duplicate pairs (although of course still fairly similar) were sufficiently dissimilar to suggest that one might well see comparable pairs of prints exhibiting a comparable level of similarity coming from *different* individuals, provided only that a large enough group of prints were examined.[79]

VII. PROFICIENCY TESTS

Despite the absence of objective standards, scientific validation, and adequate statistical studies, a natural question to ask is how well fingerprint examiners actually perform. Proficiency tests do not validate a procedure per se, but they can provide some insight into error rates. In 1995, the Collaborative Testing Service (CTS) administered a proficiency test that, for the first time, was "designed, assembled, and reviewed" by the International Association for Identification (IAI).[80] The results were disappointing. Four suspect cards with prints of all ten fingers were provided together with seven latents.[81] Of 156 people taking the test, only 68 (44%) correctly classified all seven latents.[82] Overall, the tests contained a total of 48 incorrect identifications.[83] David Grieve, the editor of the *Journal of Forensic Identification*, describes the reaction of the forensic community to the results of the CTS test as ranging from "shock to disbelief,"[84] and added:

> Errors of this magnitude within a discipline singularly admired and respected for its touted absolute certainty as an identification process have produced chilling and mind-numbing realities. Thirty-four participants, an incredible 22% of those involved, substituted presumed but false

---

[79] Of course, this depends on the particular type of AFIS software being used.

[80] David Grieve, *Possession of Truth,* 46 J. FORENSIC IDENTIFICATION 521, 523 (1996).

[81] *Id.*

[82] *Id.*

[83] *Id.*

[84] *Id.* at 525.

> certainty for truth. By any measure, this represents a profile of practice that is unacceptable and thus demands positive action by the entire community.[85]

What is striking about these comments is that they do not come from a critic of the fingerprint community, but from the editor of one of its premier publications.[86]

It is important to recognize the limits of the information provided by proficiency tests. Properly designed proficiency tests may give some sense of the possible magnitude of errors in casework on the part of examiners, but they do not provide scientific validation for assertions that a pattern is unique in the human population. Such tests are most useful when they are external and blind, that is, the examiners do not have an interest in the outcome of the test and the individuals being examined do not know that they are being tested. For precisely the same reasons, if a fingerprint match is being verified by a second examiner, the second examiner should not know the conclusion of the first examiner or other facts about the case that might affect his judgment. For example, in the Cowans case, the knowledge that a fellow officer had been shot, had an unobstructed view of the assailant, and had positively identified the suspect may have influenced the judgment of subsequent examiners looking at the latent print on the glass.

VIII. VALIDATION

The fingerprint community appears to believe that past performance provides a form of validation. The community claims that fingerprints have been used successfully for a nearly a century in the United States and, despite the millions of prints on file with

---

[85] *Id.*

[86] The results of subsequent tests have not been quite as poor in terms of erroneous identifications, but still have an error rate in the neighborhood of 10%. The improvement may reflect additional conservatism on the part of the test takers rather than changes in actual practice. For further discussion of subsequent tests, see Lyn Haber & Ralph N. Haber, *Error Rates for Human Latent Fingerprint Examiners, in* AUTOMATIC FINGERPRINT RECOGNITION SYSTEMS 339-60 (Nalini Ratha & Ruud Bolle eds., 2004).

the FBI, no matching pair coming from a distinct source is known to exist. However, these arguments do not inspire confidence. First, it need hardly be said that mere courtroom use does not constitute validation. In a case in which there are twenty-five witnesses to a shooting and the assailant is pinned to the ground, finding a matching print tells us little. The real focus should be on errors in cases in which fingerprint evidence played a key or even decisive role. A conviction in such a case does not validate the identification because the conviction itself was presumably a direct result of the fingerprint evidence.

In the pre-DNA era, finding new evidence to establish that a conviction was erroneous was a formidable undertaking. At present, however, the use of DNA for purposes of human identification has radically altered the playing field. As of June 2002, DNA testing has resulted in the exoneration of at least 108 Americans.[87] The realization that there is potentially a substantial error rate in death penalty convictions, many of these discovered because of post-conviction DNA testing, has led to a statewide moratorium on the death penalty in Illinois.[88] Thus, far from being surprising, it was perhaps inevitable that the solid science of forensic DNA identification would begin to play a role in identifying erroneous fingerprint matches. This was certainly true for the Cowans case in Boston and may also have played a role in the Mayfield case.

These two cases may merely be the harbinger of things to come. In forensic DNA testing, fingerprint examiners face, for the first time, a competing form of evidence that both the public and their own colleagues in law enforcement will certainly favor as being both more scientifically based and more credible. That being said, we still do not know the actual rate of error for fingerprint identification in criminal cases. As Donald Kennedy, the editor-in-

---

[87] *Confidence in Criminal Justice Act of 2002: Hearing on S. 2446 Before the Senate Comm. On the Judiciary*, 107th Cong. (2002) (statement of Barry Scheck, Co-founder, The Innocence Project, Benjamin N. Cardozo School of Law), *available at* 2002 WL 20318239.

[88] Press Release, State of Illinois Governor's Office, Governor Ryan Declares Moratorium on Executions, Will Appoint Commission to Review Capital Punishment System (Jan. 31, 2000) (on file with author).

chief of *Science*, notes, "It's not that fingerprint analysis is unreliable [but] . . . that its reliability is unverified by either statistical models of fingerprint variation or by consistent data on error rates."[89]

The argument that no latent print has ever been found to match the rolled print of a different person is similarly misleading because no systematic search for such pairs on the entire databank of millions of fingerprints has ever been performed. Moreover, the Mayfield case arguably provides a counterexample to the claim that no two inked prints can match the same latent; the FBI only retreated after Spanish authorities identified a superior candidate.[90] The deeply disturbing question is whether Mr. Mayfield would have ever been exonerated had the terrorist act taken place in the United States, rather than Spain, and had Mr. Daoud's print not been discovered.[91]

Unfortunately, current practice can make it difficult to identify potentially competing candidates for matches. In one California jurisdiction, a match to a latent was found by an AFIS search. The software generated ten candidate fingerprints, one of which was identified as a match to the latent. However, the information regarding the other nine prints was discarded and, it was claimed, could not be recovered. Such a procedure prevents independent experts from determining whether another print might match as well as or better than that of the candidate originally identified.

Several problems complicate true validation. First and foremost, there is the absence of any objective standard. Many examiners believe that such standards cannot be achieved and that no objective set of criteria can capture the full detail available in a fingerprint. Even if true, and the clinical-statistical prediction debate seems relevant here, it remains unclear how far one could get with objective measures. One possibility is to advocate explicit probabilistic measures of rarity. These measures would have the

---

[89] Donald Kennedy, Editorial, *Neuroscience and Neuroethics*, 302 SCIENCE 1625, 1625 (2003).

[90] *See supra* II.B.

[91] It is important to keep in mind that three senior examiners of the FBI and an independent outside examiner all signed on to the match.

merits of objectivity, transparency, and verifiability.[92] Moreover, it may be unnecessary to design an automated procedure to extract all of the information related to a single fingerprint, given that DNA analysis functions successfully with only a very small portion of the genetic information available. For example, pattern recognition experts Pankanti, Prabhakar, and Jain report:

> Our results show that (1) contrary to the popular belief, fingerprint matching is not infallible and leads to some false associations, (2) while there is an overwhelming amount of discriminatory information present in the fingerprints, the strength of the evidence degrades drastically with noise in the sensed fingerprint images, (3) the performance of the state-of-the-art automatic fingerprint matchers is not even close to the theoretical limit, and (4) because automatic fingerprint verification systems based on minutia use only a part of the discriminatory information present in the fingerprints, it may be desirable to explore additional complementary representations of fingerprints for automatic matching.[93]

Modern science has transformed many disciplines; there is no reason why the same should not be true for fingerprint identification.

IX. MAYFIELD REVISITED

The Mayfield case affords important insights into the current standards of fingerprint identification in the United States. The misidentification that took place cannot be explained away as the error of a single incompetent individual, nor can it be rationalized as the regrettable product of a substandard laboratory. Nor is it surprising that the safeguards employed by the system failed. The practice of having additional individuals review a match (the "V" in ACE-V) has little value if such a review is not blind. Internal

---

92 *See* Champod & Evett, *supra* note 64.

93 Sharath Pankanti et al., *On the Individuality of Fingerprints*, 24 IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE 1010 (Aug. 2002).

review can be superficial and subject to pressure; the use of external experts to review a case suffers from the reality that such individuals, often retired police examiners themselves, may be overawed by the FBI or state and local agencies, and, human nature being what it is, in any case predisposed to believe the original finding.

Mayfield also has things to tell us about the scientific status of fingerprint examination. The fingerprint profession invariably attributes errors in identifications to individuals rather than the underlying methodology itself.[94] But, given its unavoidable subjective component, in latent print examination people *are* the process.

This can be seen in Mayfield. In mid-June 2004, the FBI convened an outside committee of experts to review the FBI's handling of the case; the findings of that committee were summarized in a report written by Robert Stacey of the FBI.[95] The Stacey report states: "The error was a human error and not a methodological or technology failure."[96] Such a position is hardly credible. This was not, at a minimum, "a human error" (note the singular tense), but several errors, on the part of senior FBI fingerprint examiner Terry Green (who made the initial identification); supervisory FBI fingerprint specialist Michael Wieners (the unit chief) and fingerprint examiner John T. Massey (a retired FBI examiner who had thirty years of experience), both of whom verified the identification; and Kenneth Moses, director of Forensic Identification Services of San Francisco, the court-appointed independent expert.[97]

Initially, the FBI sought to lay the blame on the quality of the

---

[94] *See* SIMON COLE, SUSPECT IDENTITIES: A HISTORY OF FINGERPRINTING AND CRIMINAL IDENTIFICATION (2001).

[95] Stacey, *supra* note 30, at 706-18.

[96] *Id.* at 714.

[97] *Id.* at 710; *See also* Ben Jacklet & Todd Murphy, *Now Free, Mayfield Turns Furious*, PORTLAND TRIB., May 25, 2004, *available at* http://www.portlandtribune.com/archview.cgi?id=24535. For the fingerprint profession's general proclivity to distinguish between human and methodological errors and attribute all fingerprint misidentifications to the former, *see generally* Cole, *Grandfathering Evidence*, *supra* note 70.

latent; "Robert Jordan the FBI agent in charge of Oregon . . . said the FBI's initial determination about Mayfield's fingerprint was 'based on an image of substandard quality.'"[98] The June committee summarily dismissed this: "All of the committee members agree that the quality of the images that were used to make the erroneous identification was not a factor."[99]

But if it was not the quality of the latent, then what caused the misidentification? Here the committee pointed to a phenomenon discussed earlier, "confirmation bias," or "context effect," that is, "the mind-set in which the expectations with which people approach a task of observation affects their perceptions and interpretations of what they observe."[100]

Unfortunately, it is apparent that the fingerprint community only partially recognizes the nature of the problem. In discussing possible remedies, the Stacey report states that:

> Procedures that require descriptive documentation . . . and blind verification (i.e., previous results unknown to the verifier) should be implemented *on designated cases* . . . . The original examiner's document should be sealed or withheld from the verifier. The verifier would then conduct his or her examination independently . . . .[101]

"Designated cases" are later defined as being "high-profile cases or cases with latents of poor quality."[102] It is a positive sign that the report recognizes the importance of blind verification, but puzzling why it felt this step to be necessary only in "designated cases." The reason appears to be that the report believes the problem in Mayfield arose in part because of the "extremely high-profile" nature of the case, a circumstance the report alludes to several

---

98 *Spain Bomb Case Tossed*, PITT. POST-GAZETTE, May 25, 2004, at A7.

99 Stacey, *supra* note 30, at 718.

100 *Id.* at 713. "Context effects" have been discussed in the forensic literature before. See, e.g., M. J. Saks et al., *Context Effects in Forensic Science: A Review and Application of the Science of Science to Crime Laboratory Practice in the United States*, 43 SCIENCE & JUSTICE 119 (2003), for an excellent discussion.

101 Stacey, *supra* note 30, at 715 (emphasis added).

102 *Id.* at 717.

times.[103] In fact, there is no reason why *all* verifications should not be blind, and—for the reasons discussed earlier in this paper—every reason why they should. The potential for bias exists in all cases, not just "high-profile" ones. (The high-profile cases are, of course, the ones in which misidentifications are most likely to be detected.)

The report asserts, perhaps with this issue in mind, that "[l]atent print examiners routinely conduct verifications in which they know the previous examiners' results and yet those results do not influence the examiner's conclusions," but no evidence for this lapidary statement is furnished, and there is no reason to credit it.[104]

The Stacey report also seems naïve in its prescription for dealing with disagreements. It states:

> The verifiers must be willing to oppose any examiner if the verifiers do not see the details needed to effect the identification decision. The quality assurance program should make examiners feel that they can disagree about any identification. The examiners should be encouraged to step forward, without fear of reprisal, if they disagree. This part of the scientific method must be institutionalized.[105]

The verifiers should not "feel that they can disagree" because there should be nothing for them to either agree or disagree about. Blindness is an absolute prerequisite for independent evaluation. And if disagreement arises after blind verification, then the matter should be referred to a committee of third parties. Expecting, for example, a junior verifier to stand up to a senior member of a unit, especially in an organization like the FBI, is totally unrealistic. It is, in any case, profoundly disturbing that such issues were only beginning to be seriously addressed by the FBI in the year 2004.[106]

---

[103] *Id.* at 713, 716-17.

[104] *Id.*

[105] *Id* at 715.

[106] It is interesting to speculate on what might have happened in the Mayfield case if, instead of the Spanish authorities, the FBI had faced a local or state group. Indeed, in *U.S. v. Mitchell*, the FBI asked fifty-three law enforcement agencies whether Mr. Mitchell's fingerprints matched either of two latents. U.S. v. Mitchell, 365 F.3d 215, 223-25 (3d Cir. 2004). Of the thirty-nine

There are, finally, two aspects of the Mayfield identification that the Stacey report unfortunately does not address. First, there is the issue of the discrepancy in the number of matching "points"; fifteen in the case of the FBI and seven in the case of the Spanish. If there is a zero methodological error rate for analysis and comparison, this should not have happened. Second, there is the issue of the FBI examiners ignoring the dissenting view of the Spanish. The Stacey report remarks dryly only that "this was interesting, considering that the identification is filled with dissimilarities that were easily observed when a detailed analysis of the latent print was conducted."[107] This is a most unsatisfactory point at which to leave the matter. No matter how high-profile the case, how could four experienced examiners overlook such indicia?

There is at least one plausible explanation. Latent prints can suffer from both a loss of detail and distortions in the detail present. The process of identification necessarily involves a weighing of the amount of agreement in comparing two prints against the extent of dissimilarities that may just reflect distortions arising from the way in which the latent was deposited. There was a considerable amount of agreement between Mr. Mayfield's print and the latent; the Stacey report refers twice to the "power of the IAFIS [Integrated Automated Fingerprint Identification System] correlation."[108] The "power of the IAFIS correlation" simply means that, among the candidate matching prints generated by IAFIS, one of them was so similar to (that is, highly "correlated with") Mayfield's that the three senior FBI examiners involved in the examination concluded that Mayfield had to be the source of the latent. This example demonstrates that, when searching tens of millions of inked prints, the fingerprint community has no real idea of just how close a near miss can be.

The persistence of both the FBI examiners and the independent court-appointed expert in their assessments can be viewed not as

---

agencies that responded, nine believed that Mr. Mitchell's prints did not match either latent, but most of these agencies retreated from this position when the FBI continued to press the matter. *Id.*

[107] Stacey, *supra* note 30, at 714.

[108] *Id*. at 713.

the result of an error, in the sense of a departure from some unspecified protocol, but as merely reflecting their subjective judgment that the strength of the correlation and their ability to rationalize the apparent differences present indicated that the two prints had a common source. The Spanish obviously had different standards. The ability of the review committee to rationalize away the differences in findings as due to the high-profile nature of the case, rather than being the result of a methodological error, may itself reflect a form of "confirmation bias" (recall this is "the mind-set in which the expectations with which people approach a task of observation affects their perceptions and interpretations of what they observe").

X. DISCUSSION

> *"We just did our job and made a mistake . . . . That's how I like to think of it—an honest mistake . . . . I'll preach fingerprints till I die. They're infallible. I still consider myself one of the best in the world."*[109]

The forensic use of fingerprints is now more than a century old.[110] Introduced when forensic science was in its infancy, fingerprints were subject to only limited scrutiny before being generally accepted by the courts.[111]

Central to this acceptance appears to be what Simon Cole has termed "the fingerprint examiner's fallacy"—the argument that fingerprint identification is valid because "fingerprints are unique."[112] This is akin, as Cole notes, to arguing that eyewitness identification is dependable on the grounds that the human face is unique.[113] Identification of any kind involves the extraction and

---

[109] John Massey, referring to the Mayfield misidentification. McRoberts et al., *supra* note 5. The quotation is remarkable: an error was made, the examiner considers himself one of the best in the world, but fingerprints are infallible.

[110] COLE, *supra* note 94.

[111] *Id.*

[112] Cole, *Grandfather Evidence*, *supra* note 97, at 1197–1203.

[113] *Id.* at 1202.

analysis of features; the fundamental issue is not the "uniqueness" of the object under scrutiny (be it the human face, the friction ridge patterns of the human finger, or the sequence of bases in human DNA), but the accuracy of the process used to extract features and analyze them.

One reason for the widespread acceptance of this logical fallacy may be the failure to recognize the crucial difference between an inked and a latent print. If we think of Figure 1(a), an inked print, then the dependability of the process may seem self-evident; however, the true issue is our ability to extract information from the all-too frequent reality of Figure 1(b), the latent print.

From the start, the fingerprint community has claimed the ability, given sufficient detail, to identify the source of a latent to the exclusion of all others. For example, examiners have made assertions that fingerprints are "infallible," that an identification is "100 percent positive," and that the "methodological error rate" for parts of the process is zero.[114] Such claims have no scientific basis, a fact now recognized by some members of the profession itself.[115] When fingerprint identification errors have been discovered in the past, the fingerprint community has almost invariably attributed them to incompetent individuals rather problems or limitations in the methodology itself. But there is no "methodology," apart from the individual, in the sense of a universally accepted and objective set of protocols that can be applied to a set of prints to establish identity of source. Fingerprint proponents, however, often ascribe fingerprint identification the status of a science, referring to the ACE-V "methodology" as being part of the scientific method.

---

[114] *Id.* at 1231–32.

[115] Grieve, *supra* note 80, at 528. Grieve writes:

> [T]his categorical requirement of absolute certainty has no particular scientific principle but has evolved from a practice shaped more from allegiance to dogma than a foundation in science. Once begun, the assumption of absolute certainty as the only possible conclusion has been maintained by a system of societal indoctrination, not reason, and has achieved such a ritualistic sanctity that even mild suggestions that its premise should be re-examined are instantly regarded as acts of blasphemy. Whatever this may be, it is not science.

Id.

ACE-V is an acronym, not a methodology. It is merely the common sense description of what anyone would do if they were examining a latent and a candidate source print. All of this is not to say that fingerprint examination may not usually get things right. The problem is that we have no true idea of the underlying error rate.

Society and the courts accord science great deference and respect. When a scientist testifies in court, he is often viewed as an impartial and objective expert reporting the indisputable facts of science to us. If a forensic scientist tells us that it is 100% certain that an individual was present at a crime scene and has the blood of the victim on his clothes, this may be decisive in an otherwise tenuous case. This places great responsibility, but also great power, in the hands of a witness.

Today we accord forensic DNA evidence such a role, but it was hard earned and is well deserved. If a forensic scientist claims that a person is, to a high degree of likelihood, the source of the evidentiary DNA found at the crime scene, there is usually a solid scientific foundation for such an assertion, ordinarily documented in great detail in a case file that can be reviewed by outside experts, including individuals other than crime laboratory technicians. Each new major advance in the forensic use of DNA is accompanied by validation studies and tested in the fire of admissibility hearings. "Trust but verify" is the watchword.

The same cannot be said for fingerprint evidence. Latent print examination necessarily contains a large subjective component, something that automatically rules out certainty. The ability of the human mind to see what it hopes or expects is truly remarkable, and this ability flourishes in the absence of stringent safeguards.[116]

In the past, the fingerprint community has defended its lack of scientific grounding, in part, by appealing to its track record in the courts. The importance of *Cowans* and Mayfield, among other things, is that they underscore the shakiness of such an argument.

---

[116] *See* Michel D. Risinger et al., *The Daubert/Kumho Implications of Observer Effects in Forensic Science: Hidden Problems of Expectation and Suggestion*, 90 CAL. L. REV. 1 (2002) (providing an excellent review of the ability of the human mind to see what it hopes or expects to see). *See also* EVON Z. VOGT & RAY HYMAN, WATER WITCHING, U.S.A. (2d ed. 1979) (offering a classic discussion of this phenomenon).

Obtaining a conviction does not validate the identification.

Despite the present unsatisfactory state of affairs, there are some obvious remedies. A rigorous system of mandatory, frequent, external blind proficiency testing needs to be implemented. Second, a mechanism for routine, random, blind audits of latent identifications should be established. Third, the government needs to fund research into the validity and reliability of fingerprint identification, the development of pattern recognition software, and the quantification of the uncertainty inherent in latent print identifications.

Finally, the courts have a role to play as well. Limits should be placed on the testimony of fingerprint examiners ("100 percent positive identification"), so that their testimony reflects the true limits of their expertise. "Whereof one cannot speak, thereof one must remain silent."[117]

[117] "Wovon Mann nicht sprechen kann, darüber muss Mann schweigen." (The concluding sentence of LUDWIG WITTGENSTEIN, TRACTATUS LOGICO-PHILOSOPHICUS (1921)).