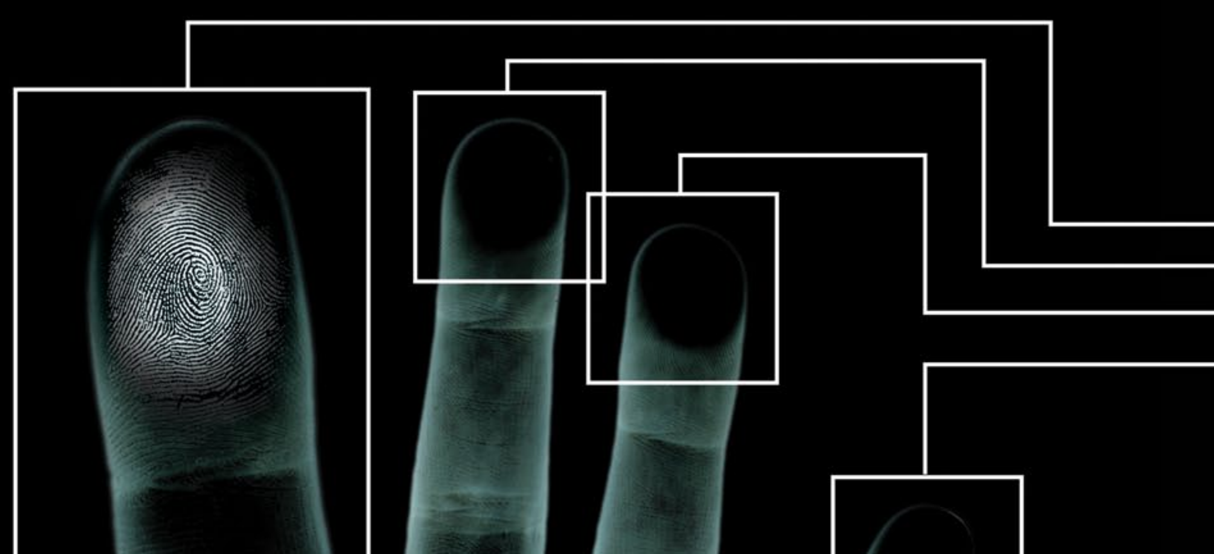
CHAPTER 14

SCIENTIFIC RESEARCH SUPPORTING THE FOUNDATIONS OF FRICTION RIDGE EXAMINATIONS

GLENN LANGENBURG

CONTENTS



CHAPTER 14

SCIENTIFIC RESEARCH SUPPORTING THE FOUNDATIONS OF FRICTION RIDGE EXAMINATIONS

GLENN LANGENBURG

14.1 Introduction

When some people think of research, what comes to mind are images of individuals in white lab coats, looking up intermittently to take data measurements and jot down notes. This is a very limited and narrow view of research. Investigative reporters, attorneys, police detectives, engineers, authors, actors, and, of course, scientists, all perform research. The scientist, however, performs scientific research. Simply defined, research is an inquiry into any subject or phenomenon. Scientific research, then, can be defined as a scientific inquiry into a subject or phenomenon.

What makes an inquiry "scientific"? What is science? What is scientific method? What are the rules for a scientific inquiry? The answers to these questions are not simple, and are the subject of an entire realm of philosophy of science. This chapter will review some of these topics, relating the issue to friction ridge skin science. The reader, however, is encouraged to read more regarding the philosophy of science to better understand the complexity of science and scientific inquiry.

14.2 The Nature of Scientific Inquiry

14.2.1 Science and Falsifiability

The word *science* is derived from the Latin *scientia* (meaning knowledge), which is itself derived from the Latin verb *scire* (to know). Science can be defined as a body of knowledge obtained by systematic observation or experimentation. This definition is very broad, and, under such a permissive definition, many fields of study may be defined as science. Scientific creationism, theological science, Freudian psychoanalysis, and homeopathic medicine could arguably be classified as sciences.

Sir Karl Popper (1902–1994) recognized the difficulty of defining science. Popper, perhaps one of the most respected and widely known philosophers of science, separated science from nonscience with one simple principle:

falsifiability. Separation, or demarcation, could be done if a theory or law could possibly be falsified or proven wrong (Popper, 1959, 1972). A theory or law would fail this litmus test if there was no test or experiment that could be performed to prove the theory or law incorrect. Popper believed that a theory or law can never be proven conclusively, no matter the extent of testing, data, or experimentation. However, testing that provides results which contradict a theory or law can conclusively refute the theory or law, or in some instances, give cause to alter the theory or law. Thus, a scientific law or theory is conclusively falsifiable although it is not conclusively verifiable (Carroll, 2003).

Although the Popperian view of science is a widely held view amongst scientists, it is important to note that the U.S. Supreme Court has also taken this view of science (*Daubert,* 1993, p 593). Justice Blackmun, writing for the majority, cited Popper, specifically noting that a scientific explanation or theory must be capable of empirical testing. The issue of falsification was also raised during the *Daubert* hearing for the admissibility of latent print evidence during *U.S. v Mitchell* (July 13, 1999). (For an explanation of *Daubert* hearings, see Chapter 13.)

14.2.2 Scientific Laws and Theories

There is a grand misconception, even within the scientific community, that scientists first make observations; then they postulate a hypothesis; after rigorous testing, the hypothesis is accepted, thus becoming a theory; then the theory, after enjoying many years of success, without any instances of being refuted, is accepted as a scientific law. This hierarchical structure is a myth (McComas, 1996). Schoolhouse Rock (Frishberg and Yohe, 1975) described such a hierarchy for bills on their journey to becoming laws. Such is not the case in science.

Scientific laws and theories, though related, represent different knowledge within science. McComas stated, "Laws are generalizations, principles or patterns in nature and theories are the explanations of those generalizations".

Scientific laws describe general principles, patterns, and phenomena in the universe. Scientific theories explain why these general principles, patterns, and phenomena occur. The verbs associated with laws and theories speak to the nature of these concepts: scientific laws are discovered; scientific theories are invented (McComas, 1996).

Exactly what defines a law and exactly what defines a theory is contested within the philosophy of science. In fact, some philosophers of science (Van Fraassen, 1989, pp 180–181) believe that no laws exist at all. However, the majority of modern philosophers of science believe that laws exist and there are two popular competing definitions: *systems* and *universals* (Thornton, 2005).

The systems definition of a law defines a law within a deductive system. Axioms are stated that allow deductive conclusions. The strength of the law is within the truth of the generalized statement and its simplicity. As an example, if "all human friction ridge skin is unique", and I am a human, then one can deduce from the law (if true) that my friction ridge skin is unique. Instances of nonunique friction ridge skin would obviously show the law to be false.

The universals definition of a law defines the law as a relationship or "contingent necessitation" between universals (universals being just about anything). The wording of such a law would be similar to:

- Humans exist.
- Unique friction ridge skin exists.
- The law is the relationship of these two entities: Humans possess unique friction ridge skin.

In either case, laws can be described by the following features (Hempel and Oppenheim, 1948; Zynda, 1994):

- Laws are universal.
- Laws have unlimited scope.
- Laws contain no designation of individual, particular objects.
- Laws contain only "purely qualitative" predicates.

Theories, on the other hand, are explanations for laws. For example, Sir Isaac Newton discovered the "Law of Gravity". This law is universal, unlimited, not just applicable to a unique object, and is descriptive and predictive. However, this law does not explain how and why gravity works. Scientists of Newton's era proposed waves of gravity emitted from objects, attracting each other, operating similarly to magnetism. The attractive forces of gravity comprised the Theory of Gravity. Later, Albert Einstein found instances where the theory did not hold up (e.g., light bending toward massive objects in space). According to the accepted theory of the time, Einstein's observations were not possible. Einstein proposed a new and revolutionary theory of gravity to explain this phenomenon. Einstein's new theory was called the "General Theory of Relativity" and described curvatures in the space–time continuum. These curvatures were due to massive objects exerting their force of gravity on the space–time continuum, very similar to bowling balls placed on an outstretched blanket. Einstein's proposed theory was not initially accepted, but after years of tests and experiments, his theory gained acceptance.

This is the true nature of science. Laws are discovered. Theories are invented to explain them. The laws and theories are tested by experiments, observations, and hypothesis testing. Hypotheses are woven together into the theories as the theories are modified. Theories are never proven, only continually tested and updated. Theories can be accepted for hundreds of years, but with the advent of newer technology, theories are subjected to new tests and rigors, and eventually outdated or incomplete theories give way, absorbed into new, mature theories. The science of friction ridge skin has experienced exactly such trials.

14.2.3 Laws and Theories in Friction Ridge Examination

If we accept the definition that a scientific law is a generalized description of patterns and phenomena in nature and a scientific theory is the explanation for that law, then what theories and laws exist within the discipline of friction ridge science?

The two most basic laws are:

1) Human friction ridge skin is unique.

> Each individual possesses a unique arrangement of friction ridge skin. Specifically, the ridge arrangements, the robust arrangements of the minutiae within the ridge patterns, and the shapes and structures of the ridges all combine to form a unique arrangement of friction ridge skin in the hands and feet of each individual.

2) Human friction ridge skin is persistent (permanent) throughout the individual's lifetime.

> Specifically, what is meant by persistence is that the sequence of the ridges and the arrangement of the robust minutiae do not change throughout a person's lifetime. This is not to say that the friction ridge skin does not change over time. It does. Friction ridge skin expands as people grow from childhood to adulthood. Skin cells constantly slough off. The substructure of the skin changes over time and ridge heights decrease (Chacko and Vaidya, 1968). The number of visible incipient ridges increases as we age (Stücker et al., 2001). Hairline creases and wrinkles proliferate as we age. All these factors describe a dynamic and changing friction ridge skin. Yet the arrangement of the minutiae and the ridge sequences is very robust and reproducible. There is evidence to support that third-level details (e.g., ridge shapes and pore locations) are persistent; this is explored later in the chapter (see section 14.3.2.2).

The next question of interest is, Are these scientific laws? According to Popper, to satisfy the criteria for scientific laws, these laws must be falsifiable. Clearly, both laws are easily falsifiable. One must simply find instances where different individuals have indistinguishable friction ridge skin or instances where the arrangement of the ridges in friction ridge skin is observed to naturally change over time (excluding injury or trauma, of course). However, in the history of this discipline, no such instances have been demonstrated.

Suppose one individual, in the entire world, actually did have a fingerprint that matched someone else's fingerprint. Obviously, the forensic community would be shocked, and the verity of the law would be questioned. But in a purely Popperian view (Thornton, 2005):

> No observation is free from the possibility of error—consequently we may question whether our experimental result was what it appeared to be. Thus, while advocating falsifiability as the criterion of demarcation for science, Popper explicitly allows for the fact that in practice a single conflicting or counter-instance is never sufficient methodologically to falsify a theory [or law], and that scientific theories [or laws] are often retained even though much of the available evidence conflicts with them, or is anomalous with respect to them.

Thus, Popper advocated constant testing to refute a theory or law. A single instance of falsifiability should spawn additional testing.

Fundamental theories exist that explain the two laws of uniqueness and persistency. Uniqueness is explained by biological variations (genetic influences and random localized stresses) within the developing fetus. Persistence is maintained by the substructural formations of the developing skin (hemidesmosomes, papillae, and basal layer).

These are theories that explain the laws. These theories have empirical evidence and testing that support, but do not conclusively prove, them. Additional information may be learned that will cause these theories to be adjusted and incorporate the new data. Thus, science is evolving and dynamic.

14.2.4 Hypothesis Testing

Theories and laws are commonly challenged through hypothesis testing. The results of testing a hypothesis can support or refute a theory or law. In some instances, the results will call for modifications to be made to a law or theory, which in turn leads to further hypotheses to test under the new or modified law.

Although there are no rigorous formulas or recipes for testing hypotheses and designing experiments (nor should there be), a generic model for hypothesis testing can be described. The steps of this model are often referred to as "scientific method". Huber and Headrick (1999) noted that the term scientific method is a misnomer. They stated that scientific method is derived from epistemology (the study of knowledge and justified belief, according to the *Stanford Encyclopedia of Philosophy*). Francis Bacon defined a basic approach to scientific method encapsulated in four steps: (1) observe, (2) measure, (3) explain, and (4) verify (Huber and Headrick, 1999). This description in modern times has been modified into a hypothesis testing model. The basic steps of the hypothesis testing model have been described as:[1]

- Observation.
- Hypothesis formulation.
- Experimentation.
- Data analysis and conclusion.
- Reproducibility.
- Communication of results.

The researcher must first make a specific observation or note a general problem or query. Then a hypothesis is formulated (often referred to as the "null hypothesis"). The hypothesis is testable and falsifiable. A counter-hypothesis is also formulated. A suitable experiment is designed to test the specific hypothesis. Data from the experiment are collected. These data may be qualitative or quantitative. The data are evaluated, often statistically (though that is not a requirement), and conclusions are drawn whether to accept the hypothesis or reject the hypothesis and accept the null hypothesis. The results of the experiment should be reproducible by another scientist following the methodology. Finally, the results should be communicated to others. This is important not only for sharing the knowledge but also for peer review and critical analysis.

14.2.5 Comparison Methodology and Theory

As an extension of the law that friction ridge skin is unique, if during the deposition of a latent print, the details of the friction ridge skin are sufficiently recorded on a surface via residues on the friction ridge skin, then theoretically *the latent print image can be individualized to the source friction ridge skin.*

This is what Hempel and Oppenheim (1948) refer to as a derived theory (as opposed to a fundamental theory). The derived theory allows application of the principle to specific objects or individuals that would be prohibited by the universality and generality requirements of a law or fundamental theory. However, the theory that latent prints can be attributed to a unique source of friction ridge skin raises some questions that are difficult to answer.

Even if the friction ridge skin is unique down to the cells and ridge units, this issue is secondary to whether a latent print (which will not contain all of the information in the source skin) can be correctly attributed to its source. How much information must be transferred for the examiner to reliably individualize the latent print? What happens to the reliability of the details when subjected to distortions? What tolerances are acceptable regarding distortions and the flexibility of skin?

Ultimately, the latent print will be compared to a source (via known standard reproductions) by an expert. The comparison methodology generally accepted in the United States is the ACE-V methodology. This is an acronym for analysis, comparison, evaluation, and verification. The stages of ACE-V methodology are defined as: *Analysis*—Assessment of the quantity and quality of ridge detail present in an impression; *Comparison*—A side-by-side comparison of the two

[1] This basic model can be found in most elementary collegiate science texts in various forms.

impressions; *Evaluation*—The decision process to declare an individualization, exclusion, or inconclusive opinion; *Verification*—Verification of the result by another competent examiner. The ACE process was initially described by Huber as a logical, methodological process for the comparison of handwriting evidence (Huber, 1959). (For more about ACE-V, see Chapter 9.)

It has been argued elsewhere that ACE-V "methodology" is not in any real sense a methodology and is more akin to a "protocol" (Champod et al., 2004). A methodology would typically encompass very explicit steps, instructions, criteria, and a transparent decision model. This has not been accomplished. The ACE-V protocol, however, serves as an appropriate model and descriptor for performing any sort of forensic comparative examination, whereby evidence from an unknown source is compared against appropriate known exemplars to reach an opinion regarding the source of the evidence. As such a protocol, it offers good suggestions for general forensic examinations such as (1) analysis of the unknown should be done separately, prior to comparison to the known exemplar, and (2) there must be verification of the conclusion and peer review of the reasoning used to reach the proffered conclusion.

Wertheim has suggested that ACE-V is analogous to the scientific method (Wertheim, 2000, pp 1–8). Huber and Headrick made a similar analogy for the ACE process with respect to handwriting comparisons (Huber and Headrick, 1999, pp 351–355). The analysis is the assessment (observation) that a latent print has detail sufficient for a comparison. A hypothesis is formed: the latent print originated from Individual A; a null hypothesis is formed: the latent print did not originate from Individual A. The images are compared and agreement is found or not found (experimentation). Based on the degree of agreement (data), one concludes that there is sufficient evidence during the evaluation stage to individualize or exclude (support or reject the hypothesis as a conclusion). The process is then verified by another expert during verification (reproducibility).

As Hughes (1998, pp 611–615) has noted, the practice of friction ridge examination is an applied science. The discipline borrows from other sciences to support and justify the practice of comparing friction ridge images by a specific comparison methodology.

14.3 Scientific Research Related to Friction Ridge Examination

14.3.1 Friction Ridge Skin Is Unique

In order to prove the axiom of unique friction ridge skin to be true, every area of friction ridge skin on the planet (and all the skin of past and future generations) would need to be examined. Obviously, this will never be possible. Therefore, to support this premise, the discipline looks to three areas of support:

- Empirical observations and evidence.
- The theory of the formation of friction ridge skin (i.e., the biological formation).
- Fingerprint individuality models based on probability and statistics.

14.3.1.1 Observations. The empirical evidence, for many years, was generally viewed by the discipline as the *pièce de résistance* of evidence for the claim that friction ridge skin is unique. An expert would anticipate under vigorous cross-examination during trials to be asked, "Well, how do you *know* that no two fingerprints are alike?" The typical answer of course was, "Because in all the history of fingerprints, all the billions of comparisons worldwide, no two fingerprints have ever been found to be identical, from different sources, and this includes identical twins."

Although this fact is important and should not be dismissed, it does not satisfy the argument and does not prove that one person's particular print does not have a matching mate somewhere out there on the planet. All that can be inferred from this fact is that, presently, no two people have been found to have matching fingerprints. Taking it a step further, it does not satisfy that one particular latent print, with just enough distortion and low clarity, might not be mistaken to be from a different source, given that the false source was very similar in appearance to the true source of skin. Latent print examiners should be cautious about resting merely on empirical evidence to support the uniqueness of friction ridge skin. Furthermore, the number of actual comparisons that have been performed, when compared to the total number of possible comparisons available (i.e., every human's friction ridge skin against every other human's friction ridge skin), is only the smallest fraction (cf. by inference, "The Snowflake Paradigm" by Thornton, 1986). Therefore, given what would undoubtedly be an exceptionally small probability (i.e., matching fingerprints between two different people), an impossibly large number of comparisons would need to be done to even have a realistic chance of finding such a match in the population. So even if matching fingerprints were to exist in the population, the chance of discovering them is simply too remote.

The literature lacks research that was specifically conducted to prove that no two areas of friction ridge skin are alike. The absence of such a study stems from (1) as discussed previously, its impossibility, and (2) the profession's consistent reliance on its collective experience and case studies to demonstrate the point. Additionally, it could be argued that, until *U.S.* v *Mitchell* (1999), the premises and validity of friction ridge skin examinations had not been seriously challenged or scrutinized; therefore, the impetus to scientifically test the law, under the rigors of present-day science, has not existed.

Still, although there is not (and cannot be) any definitive way to prove that all friction ridge skin is unique, there exists empirical evidence that supports the premise. Evidence from "look-alikes" (i.e., close nonmatches—friction ridge skin impressions from two different sources that are very similar in appearance) (IEEGFI-II, 2004, p 13) has been helpful. Evidence from look-alikes can be found in monozygotic twin research and two Automated Fingerprint Identification System (AFIS) studies.

Studies of Monozygotic Twins. If one wanted to find areas of matching friction ridge skin from two different individuals, it would seem that the population of monozygotic twins would be a good place to start the search. Galton (2005, pp 185–187, originally published in 1892) first explored this avenue. He found similarities in patterns, but the minutiae were different. Similarly, other researchers, exploring the hereditary aspects of fingerprints, have examined the prints of monozygotic twins. The works of Wilder, Grüneberg, Bonnevie, and Newman are summarized by Cummins and Midlo (1943, pp 210–245). These researchers all investigated the similarities of fingerprints between monozygotic twins. Their findings mirrored the conclusions of Galton.

Okajima (1967, pp 660–673) found a higher correlation for the number of minutiae present between the fingerprints of identical twins than the number of minutiae present between the fingerprints of fraternal twins. Lin and colleagues (1982, pp 290–304) further investigated this relationship. They examined the correlations for fingerprint pattern, ridge count, and minutiae positioning for 196 pairs of twins (including both identical and fraternal twins). They found that the correlations followed the trend (in decreasing order of correlation): identical twins, fraternal twins, related siblings, and lastly, unrelated individuals. Their work echoed that of previous researchers noted by Cummins and Midlo (1943, pp 235–245). Lin and colleagues (1982) concluded that "although fingerprints [of identical twins] may have a high degree of similarity . . . variations in minutiae distribution still permit their differentiation" (Lin et al., 1982, p 304).

In more recent times, German *(U.S.* v *Mitchell,* July 8, 1999, pp 2–56), in preparation for a *Daubert* hearing, performed similar analyses as Lin and colleagues (1982) with a database of fingerprints of 500 pairs of twins (including both identical and fraternal twins). Again, similarities in patterns, ridge count, and minutiae locations were noted between identical twins, but the prints were still differentiable. German further noted that even in the smallest areas of agreement (clusters of two to three minutiae located in similar positions), he could differentiate the prints based on third-level detail (i.e., the shapes of the ridges and pore locations). However, it should be noted that the work of German was not published. Therefore, it was not peer reviewed and can only be found in the testimony during the *Daubert* hearing in the *Mitchell* case. Moreover, unlike Lin and colleagues (1982), the German study was not conducted with well-defined hypotheses to be tested, the methods to test the hypotheses were not clear prior to the commencement of the work, and it is not clear what metrics were used to determine the strength of the similarities and dissimilarities when comparing mated monozygotic twin prints.

Srihari and colleagues also conducted a large study of twins' fingerprints (Srihari et al., 2008). They used 298 sets of twins and 3 sets of triplets. The researchers used a minutiae-based automatic fingerprint identification algorithm to compute comparison scores. The researchers compared each identical twin to his or her mated identical twin. They also compared scores between twins' fingerprints and unrelated twins' fingerprints, fraternal twins, and non-twins. Comparing the distributions of scores produced, the researchers found that twin pairs have more similarities in level 1 detail and level 2 detail than the general population, but are still discriminable.

All of the previous studies with twins dealt exclusively with known exemplars of their friction ridge skin. What is lacking from the literature is whether an examiner can correctly attribute a latent impression to the correct friction ridge skin source when identical twins have deposited latent prints. The only data of this nature can be found in the 1995 Collaborative Testing Services (CTS) latent print examiner proficiency test (CTS, 1995; Grieve, 1996, pp 521–528). This particular CTS proficiency test included a bloody impression from an individual whose fingerprint exemplars were not provided for the proficiency test. Instead, the fingerprint exemplars from the donor's identical twin, who did not create the bloody impression, were provided. Approximately one in five participants in this proficiency test erroneously individualized the impression to the incorrect source.

Empirical Data. It is unknown exactly which individual or culture first recognized the individuality of fingerprints. From the ancient Middle East to the ancient Chinese, there is evidence in these cultures of an awareness of the uniqueness of fingerprints. (For a timeline of fingerprint science, see Chapter 1.) It was not until 1788 that Dr. J. C. A. Mayer recorded:

> Although the arrangement of skin ridges is never duplicated in two persons, nevertheless the similarities are closer among some individuals. In others the differences are marked, yet in spite of their peculiarities of arrangement all have a certain likeness. (Cummins and Midlo, 1943, p 13)

Mayer is considered the first individual to record the assertion that friction ridge skin is unique.

Many more early pioneers investigating this phenomenon followed, including Sir William Herschel and Dr. Henry Faulds. However, neither Herschel nor Faulds published hard data in support of their theories. In his 1880 letter to *Nature* (Faulds, 1880, p 605), Faulds reported several conclusions, including "absolute identity" of criminals from crime scene latent impressions. However, Faulds never provided the data for his basis, stating only that he examined a "large number of nature-prints" taken from individuals in Japan. His later writings (Faulds, 1911) refer to his examination of "many thousands of living fingers".

In 1970, the International Association for Identification (IAI) organized a committee known as the Standardization Committee. The primary task of the committee was "to determine the minimum number of minutiae of friction ridge characteristics which must be present in two impressions in order to establish positive identification" (McCann, 1971, p 10). For 3 years, the committee addressed this issue and in 1973, the Standardization Committee reached a consensus: "No valid basis exists at this time for requiring that a pre-determined minimum number of friction ridge characteristics must be present in two impressions in order to establish positive identification. The foregoing reference to friction ridge characteristics applies equally to fingerprints, palm prints, toe prints and sole prints of the human body" (McCann, 1973, p 13). This conclusion was arrived at through interviews with professionals in the field, a review of the literature, surveys sent to various international identification bureaus, and the generally accepted view of the profession. It is important to note that during the interviews and surveys, no agency reported any knowledge of an instance where two individuals were found to have matching fingerprints or any other matching areas of friction ridge skin (Moenssens, 2006).

As for concrete empirical studies, two are notable. Fingerprint expert Stephen Meagher (*U.S.* v *Mitchell,* July 8, 1999, pp 56–229; July 9, 1999, pp 2–31), in preparation for a *Daubert* hearing, conducted a survey. He sent images of two latent prints (the images that had been identified to the defendant in this case) to all 50 state laboratories.[2] All agencies were asked to search the two latent prints in their local AFIS databases. Only one agency reported identifications: Pennsylvania, the state in which the defendant had been arrested. Eaton (2005, 2006) reported similar findings in an unpublished pilot study. A single common loop latent print with 12 minutiae, and a second image of the same print, cropped to show 8 minutiae, were sent to 50 agencies (in 9 countries). These agencies searched the images in their AFIS databases. The only agency to report an individualization was the Western Identification Network, which was the only agency that maintained a copy of the civilian tenprint card for the donor of the latent print in this experiment.

Although neither of these results offer substantial proof that all friction ridge skin is unique, it is important to note that, after comparing these latent prints to hundreds of millions of fingerprints combined in the AFIS databases, no

---

[2] It should be noted however (and this concern was raised during the *Mitchell* testimony) that the surveys did not always reach the intended participants. In some states, the surveys were sent to the Criminal Justice Information Services Division instead of the latent print unit. Therefore, the distribution of the surveys may not have been properly controlled.

agency reported a match to anyone other than the correct known source. In effect, Meagher and Eaton were not able to falsify the individuality of fingerprints in these noteworthy, albeit limited, instances.

14.3.1.2 Biological Basis. On the basis of a holistic and qualitative understanding of the morphogenetic processes of friction ridge skin formation, latent print examiners have predominantly supported the statement: Nature never repeats itself (McRoberts, 1996; Thornton, 1986). This position has been further supported by the views of numerous biologists, zoologists, and anatomists who have explored the proffered model for friction ridge skin formation (Wilder and Wentworth, 1918, 1932; Cummins and Midlo, 1943; Hale, 1952; Okajima, 1967; Misumi and Akiyoshi, 1984; Montagna and Parakkal, 1974, Montagna et al., 1992; Babler, 1978, 1990, 1991).

Early authors generally referred to the variability of minutiae alone, and thus a probabilistic approach to fingerprint individuality, as evidence for the uniqueness of friction ridge skin (Galton, 2005, pp 100–113; Wilder and Wentworth, 1932, pp 309–328). Cummins (2000, pp 79–90) and Hale (1952, pp 147–173) recognized that the variability in minutiae formations and appearance were attributable to random mechanical stresses during friction ridge formation. The patterns of friction ridge skin and the arrangement of the minutiae, in conjunction with variability in the edge formations (Chatterjee, 1962), pore locations (Locard, 1912; Faulds, 1912, pp 29–39), and ridge widths and heights (Cummins et al., 1941; Ashbaugh, 1999, pp 61–65), provide a seemingly infinite palette of variation, even in the smallest regions. Montagna and colleagues have generally noted that skin (friction ridge and nonfriction ridge skin) differs from individual to individual and is not repeated elsewhere in regions on each individual (Montagna and Parakkal, 1974; Montagna et al., 1992). Montagna and colleagues noted in their observations and study of friction ridge skin and nonfriction ridge skin:

> The palmar and plantar surfaces are filigreed by continuous and discontinuous alternating ridges and sulci [furrows]; the details of these markings and their configurations are collectively known as dermatoglyphics. Each area has unique regional and individual structural variations not matched elsewhere in the same or in any other individual. (Montagna et al., 1992, p 8)

The biological model for the morphogenesis of friction ridge skin supports the perspective for the uniqueness of friction ridge skin. Although not necessarily providing concrete evidence to test the uniqueness of friction ridge skin, the theory does explain why the law holds true.

The biological basis for flexion crease formation has been studied by several researchers (Kimura and Kitigawa, 1986, 1988; Popich and Smith, 1970). With respect to the study of palmar features, empirical frequencies have been reported by Tietze and Witthuhn (2001). They reported frequencies of creases, ridge flow, patterns, and other distinct formations from 35,000 pairs of palmprints. Although these observations do not show "uniqueness" of palmar features, these data are helpful for assessing the rarity of these features.

14.3.1.3 Probability Models for Fingerprint Individuality. Though many early pioneers recorded their empirical observations, it was Sir Francis Galton who developed the first probability model for individuality, resulting from his systematic analysis and study of fingerprints. From Galton's model in 1892 to the present, there have been approximately two dozen or so models, each improving or refining aspects of previous models.

This section will summarize the significant research and models available. The summaries given are very basic and brief. Excellent summaries, discussions, and critiques of these models, including the assumptions, limitations, and strengths of each, have been provided elsewhere (see Stoney and Thornton, 1986a, pp 1187–1213; Stoney, 2001, pp 327–387; Pankanti et al., 2001, pp 805–812).

The Galton Model (1892) (Galton, 2005, pp 100–113). Although Galton devised the first probability model for fingerprint individuality, it was very crude. Using enlargements of fingerprints, Galton dropped square pieces of paper of varying size randomly over the enlargements. He then attempted to predict whether the pieces of paper covered minutiae. Galton built his model on his ability to predict the occurrence of minutiae, dependent on the configuration of the surrounding ridges. He did not base his model on the actual frequencies and distributions of minutiae. Furthermore, he used unrealistic factors to estimate probability of differing pattern types and the number of ridges in a particular region of the print. From these calculations, he arrived at the probability of finding any given arrangement of minutiae in a fingerprint to be 1.45×10^{-11} (i.e., 1 in 68 billion).

The Henry Model (1900) (Henry, 1900, pp 54–58). The second model, proposed by Sir Edward Henry, was a drastic deviation from Galton's approach. Henry proposed that each minutia was an independent, identically distributed event (each occurrence of minutia has the same probability and is not dependent or influenced by any other minutiae). The probability of a minutiae event was 1/4 (.25). The probability of finding 12 matching minutiae was then $(1/4)^{12} = 6 \times 10^{-8}$ (i.e., approximately 1 in 17 million). To account for pattern type, according to Henry's model, pattern type was deemed equivalent to two more minutiae (multiplying the previous results for minutiae by 1/16). Thus, if given a whorl print with 12 minutiae, the probability of finding a whorl print with 12 matching minutiae is $(1/4)^{14}$ or 4×10^{-9} (i.e., approximately 1 in 270 million).

The Balthazard Model (1911) (Balthazard, 1911, pp 1862–1864). Using Henry's approach, Dr. Victor Balthazard (a French medical examiner) also used the probability of a minutia event equal to 1/4, but while Henry's was arbitrary, Balthazard based his use of 1/4 on whether a bifurcation or ridge ending pointed to the left or to the right. He proposed that each of these four possibilities (bifurcation left or right, ridge ending left or right) is equally likely to occur, and thus he arrived at a probability of 1/4 for a minutia event. His model did not include a factor for pattern type. He then reasoned that, in order for his model to satisfy the expectation of only one person on the planet to have a matching configuration to the print, 17 minutiae in agreement would need to be found. By his model, finding 17 matching minutiae had a probability of $(1/4)^{17} = 6 \times 10^{-11}$ (i.e., 1 in 17 billion). He also conceded that if one was certain the donor was restricted to a certain geographical region, then a positive identity could be established with a lower number of minutiae (e.g., 10 to 12 minutiae). In effect, Balthazard proposed the first "minimum point" threshold.

The Locard Model (1914) (Locard, 1914, pp 526–548; Champod, 1995, pp 136–163). The Locard model is not a statistical model, but rather a pragmatic opinion derived from the statistical models of Dr. Edmond Locard's era. Locard established his tripartite rule:

1) If more than 12 concurring minutiae are present and the fingerprint is very clear, then the certainty of identity is beyond debate.

2) If 8 to 12 concurring minutiae are found, then identification is marginal and certainty of identity is dependent on:

   a. the quality (clarity) of the fingerprint,

   b. the rarity of the minutiae type,

   c. the presence of a core and delta in a clear area of the print,

   d. the presence of pores, and

   e. the perfect agreement of the width of the ridges and furrows, the direction of the ridge flow, and the angular value of the bifurcation.

3) If a limited number of characteristic features are present, the fingerprint cannot provide certainty for an identification, but only a presumption proportional to the number of points available and their clarity.

In instances of parts 1 and 2 of the rule, positive identification can be established following discussion of the case by at least two competent and experienced examiners. Locard arrived at these conclusions based on his own experience and observations and the works of Galton, Balthazard, and Ramos.[3] Part 3 of the rule, as noted by Champod (1995, pp 136–150), is highly suggestive of a probabilistic approach to fingerprint evidence and conclusions.

The Bose Model (1917) (Roxburgh, 1933, pp 189–214). Rai Sahib Hem Chandra Bose used the Henry model and also used a probability of 1/4 for a minutia event; however, he clearly did so on a poor assumption. He chose 1/4 as a probability on the basis of his contention that there are four types of minutiae events, all equally likely to occur: a dot, bifurcation, ending ridge, or continuous ridge. Clearly, there are many more continuous ridge events than minutiae and certainly more ridge endings and bifurcations than dots distributed in a typical fingerprint.

The Wilder and Wentworth Model (1918) (Wilder and Wentworth, 1918, pp 319–322). Dr. Harris Wilder and Bert Wentworth used the Henry model as well, but instead of an assumed probability of minutia occurrence of 1/4, they used 1/50. They gave only this reason as justification:

> We have no definite data for knowing the percentage of occurrence of [minutiae in a specific

[3] Galdino Ramos. *De Identificacao,* Rio de Janeiro, 1906. Locard (1914) referenced Ramos' work, stating that Ramos calculated that it would take 4,660,337 centuries before two people were born with the same fingerprints. Locard, however, sharply disagreed with Ramos' calculations, stating that they were in error because Ramos used an incorrect number of minutiae in the fingerprint as his basis for calculations. Locard did not state how Ramos computed his values, and thus it cannot be known whether Ramos overestimated or underestimated in his calculations.

> pattern]. . . . As a matter of fact it is absurd to use anywhere near as small a ratio as 4 to 1, for the percentage of occurrence of any one of these details; it would be rather 1 in 50, or 1 in 100 . . . (Wilder and Wentworth, 1918, p 321).

The Galton model only recognized and used approximately 35 minutiae on the "bulb" of the finger (i.e., in the central portion of the tip of the finger) (Galton, 2005, pp 97–98). Wilder and Wentworth (as did Balthazard) recognized that there are "60 to 100 separate details" in a full fingerprint (Wilder and Wentworth, 1932, p 319).

The Pearson Model (1930) (Pearson, 1930, p 182). Karl Pearson, an eminent mathematician and statistician of the late 19th century (famous for his many contributions to the field of statistics, including the well-known chi-square test), did not create a fingerprint model per se. Rather, in writing the biography of his good friend and colleague Sir Francis Galton, Pearson critiqued Galton's model. Pearson suggested that a more appropriate estimate of the probability of a minutiae event was 1/36, rather than 1/2 as Galton had used.

The Roxburgh Model (1933) (Roxburgh, 1933, pp 189–214). T. J. Y. Roxburgh's model incorporated several innovative concepts. First, it included a factor for the number of intervening ridges from a minutia to the origin, using a polar coordinate system. All previous (and subsequent) models used rectangular areas or Cartesian coordinate systems. Second, Roxburgh included a clarity factor, recognizing that clarity can be low due to smearing or smudging and sometimes the type of minutiae present in a print may be ambiguous. The factor, termed "Q" for quality, allowed for the adjustment of probabilities based on the quality of a minutia. The Roxburgh model also incorporated factors for pattern type and minutiae type (the latter similar to the Balthazard model).

Roxburgh also provided a table of probabilities for matching crime scene latent prints as a measure of the probability of finding that arrangement of minutiae. The table listed probabilities for 1 through 35 matching minutiae for 4 classes of clarity: "ideal", "good", "poor", and "worst". On the basis of these calculations, he provided a second table (Table 14–1) for the minimum number of minutiae needed to declare a positive identification between a crime scene latent print[4] and a known exemplar. Roxburgh included a factor for error, with upper and lower limits of margin of error of 1/500,000 (if the finger designation is unknown) and 1/50,000, respectively. Roxburgh wrote:

> Taking the value of 1/50,000 as the margin of safety, we see then that with a good average print, 8 to 9 points are sufficient for safety; for a poor average print, 9 to 10 points are required; and for a poor print 11 points; and for a very poor print, not showing the form and centre, 15 or 16 points. For a very good print (approaching an ideal print), 7 to 8 points would suffice. (Roxburgh, 1933, p 212)

Roxburgh essentially calculated minimum thresholds based on a quantitative–qualitative examination.

The Cummins and Midlo Model (1943) (Cummins and Midlo, 1943, pp 147–155). The model used by Dr. Harold Cummins and Dr. Charles Midlo is identical to the Wilder and Wentworth model, with the exception of a factor for pattern type. They reasoned that the probability of obtaining the most common fingerprint pattern (an ulnar loop) with similar ridge counts (based on 11 ridges) was 1/31. Thus, as an upper bound, this factor is multiplied with the probability of a minutiae arrangement.

The Amy Model (1946–1948) (Amy, 1946a, pp 80–87; 1946b, 188–195; 1948, pp 96–101). Lucien Amy developed a model that incorporated two essential factors of individuality: the number and position of minutiae and the type of minutiae. Amy first derived data for the type of minutiae from observing frequencies of occurrence in 100 fingerprints. All previous models either arbitrarily assigned frequencies or assumed equal frequencies. Amy used the Balthazard criteria of bifurcation to the left or right and ridge ending to the left or right, but found that these minutiae types were not uniformly distributed. From these distributions, Amy calculated a factor for minutiae type (including orientation).

Amy then calculated the total number of possible minutiae arrangements, given a number of minutiae. He did so using a binomial distribution. This sort of probability distribution and modeling would be akin to calculating how many different ways you can arrange a certain number of cars in a parking lot with a fixed number of spaces, where each car would be parked in a space, but not all spaces filled, and finally, the lot itself having a fixed, given size.

[4] Technically, the table was useful for any two images based on the quality of the images, i.e., comparing an "ideal" inked print against a smudged "worst case" inked print.

Table 14–1

Roxburgh's calculation for the minimum number of minutiae needed to declare a positive match, with a margin of error of 1 in 50,000.[1]

Population or Number in Class[2]	Character of Print				
	(i) Ideal	(ii) Good Average	(iii) Poor Average	(iv) Poor	(v) Worst Case
10^1	2	3	3	3	8
10^2	3	3	4	4	9
10^3	4	4	5	6	10
10^4	4/5	5	6	7	11
10^5	5	6	6/7	8	12
10^6	6	7	7/8	9	13
10^7	7	8	8	10	14
10^8	7	8	9	11	15/16
10^9	8	9	10	12	16/17
1.6×10^9 (world)	8	9	10/11	12/13	17
1.6×10^{10} (finger unknown)	9	10	11	13	18

Notes

(1) Table 14–1 shows the number of points that are required for safety for five types of prints. The first four columns are based on decreasing levels of quality; the fifth column was obtained by using the lowest quality print and taking a margin of error of 1/50,000.

(2) The figures are given in each case for the designation of the finger being known. If unknown, the class is multiplied by 10, and the number of points required is as for the next class below in the table.

(Adapted from Roxburgh, 1933.)

To calculate the probability of duplicating a given minutiae arrangement, Amy multiplied these two factors (minutiae type and minutiae arrangements) together and added a correction factor for clusters of minutiae.

Amy also calculated, based on his model, the chance of a false match. Amy showed that as the number of comparisons for a particular arrangement increased, so did the probability of finding a match and so did the chance of a false match. The chance of finding similar configurations in a billion people is much higher than when comparing against one or two individuals. Amy's observations follow directly from the concept that even the rarest of events have expectations of occurrence when the number of trials is very large. This is a critical concept, especially when the potential effects of large AFIS databases are considered, and the possible correlation to recent events (e.g., the Brandon Mayfield incident—section 14.3.3.4) must be considered (Stacey, 2004, pp 706–718).

Amy suggested that if a minutiae configuration is compared against one or two suspects and a match is declared, this is stronger evidence than if a minutiae configuration is compared against one billion individuals. Thus the strength of the match is decreased for a large number of comparisons and the likelihood of a false match is increased, or the criteria for a match must become more stringent when comparing against a large population to achieve the same level of reliability. However, Amy's position is that the truth of the conclusion depends both on the strength of the evidence (the match) and the size of the relevant population.

With respect to a similar debate regarding DNA evidence and DNA database searches, Donnelly and Friedman treated the strength of the evidence (the rarity of a profile) and the strength of the identification decision (the chance the profile originated from the defendant) separately (Donnelly and Friedman, 1999, pp 1–9). According to them, a DNA match either comes from a single suspect provided by police investigation (what they referred to as a "confirmation" case) or the match comes from a large database search (what they referred to as a "trawl" case). In either case, the rarity of the profile does not decrease.[5] However, the chance the profile originated from the defendant (and thus the strength of the prosecutor's case) would depend on whether the suspect was selected from a trawl case or confirmation case. From a statistical approach, the prior probabilities for the prosecutor's hypothesis (guilt) are drastically different in a confirmation case versus the trawl case. In the confirmation case, the police presumably had prior information through investigation to arrive at a particular suspect. The DNA match now adds significant weight to the case. In the "trawl" case, absent any other evidence to tie the suspect to the scene, the prosecutor's case is much weaker given only the DNA match produced from a large database, where there is a greater potential for a false match. The parallels to friction ridge examinations and AFIS databases are important to note, especially as the profession explores a probabilistic approach to friction ridge examinations.

The Trauring Model (1963) (Trauring, 1963, pp 938–940). The model by Mitchell Trauring was not a model for calculating fingerprint individuality per se, but rather for estimating the probability of a false match to an *individual* if searched in a proposed theoretical automated fingerprint identification system. The Trauring model is very similar in assumptions and calculations to the Balthazard model and was derived from the Galton model. However, instead of using the probability of 1/2 (0.50) for a minutia event, Trauring calculated the probability of a minutia event to be 0.1944. This value was based on his observations of minutiae density and his estimate of finding "test" minutiae in a quadrilateral region bounded by a set of "reference" minutiae.

The Kingston Model (1964) (Kingston, 1964; Stoney and Thornton, 1986a, pp 1204–1209). The model by Charles R. Kingston is similar in approach and complexity to the Amy model. Kingston calculated three critical probabilities for assessing fingerprint individuality: (1) observed number of minutiae for a region of a given size, (2) observed arrangements for the minutiae, and (3) observed minutiae type.

Kingston's first factor, probability of observed number of minutiae, was calculated from observations of minutiae density from 100 fingerprints. Kingston found this distribution followed a statistical model known as a Poisson distribution. (Amy had used a binomial distribution, but under these conditions, the binomial distribution is approximately a Poisson distribution.) Thus for a fingerprint area of a specific size, Kingston could calculate the probability of finding *x* number of minutiae in this space.

Also similar to Amy and to the previous analogy of cars in a parking lot, Kingston calculated the number of positions and arrangements for a given number of minutiae. The analogy of the parking lot is even more apropos to Kingston's model, as Kingston's model was based on the assignment of the first minutia into a position, then the second minutia would occupy another position, and so forth. This is similar to cars queued up to park where, after the first car has parked, the second car must find another spot, and so forth.

Kingston's final factor, the minutia type, was based on observed frequencies for almost 2,500 minutiae. Unlike the previous models, which assumed and estimated various distributions, or relied solely upon simple bifurcations and ridge endings, Kingston calculated relative frequencies for ridge endings, bifurcation, dots, enclosures, bridges, triradii, and "other" minutiae.

The Gupta Model (1968) (Gupta, 1968, pp 130–134; Stoney and Thornton, 1986a, p 1191). The model by S. R. Gupta is the last of the simple models based on the Henry model. Gupta made observations of minutiae position frequencies from 1,000 fingerprints. Unlike his predecessors, he was not examining the frequency (rarity) of a particular type of minutiae; rather, he examined how often a particular type of minutiae appeared in a specific position. Referring back to the parking lot analogy, it is akin to observing how often a Ford parks in a particular parking spot (versus a Chrysler, General Motors, or Toyota vehicle). He estimated that bifurcations and ridge endings generally appeared in a particular position with a frequency of 1/10, and less common features (e.g., dots, spurs) with a frequency of 1/100.

[5] Some sources believe the rarity of the profile would not change at all under these two scenarios. Donnelly and Friedman (1999) argued that the rarity of the profile actually would change and have more weight after a database search, because a large portion of the population has been effectively excluded as a potential donor, thus empirically demonstrating the rarity of the DNA profile. Literally, the denominator to calculate the rarity of a profile would change after a large database search, because it would be known how many individuals did not have the profile. Significant debate surrounds this issue. The debate illustrates a classic difference between the frequentist and Bayesian approaches.

Gupta also included a factor for pattern type and ridge count for the pattern.

The Osterburg Model (1977–1980) (Osterburg et al., 1977, pp 772–778; Sclove, 1979, pp 588–595; 1980, 675–695). The Osterburg model was proposed by Osterburg, Parthasarathy, Raghavan, and Sclove in 1977. The model was modified by additional work by Sclove in 1979 and 1980. The basic Osterburg method was to divide a fingerprint into square cells, with each cell possessing an area of 1 sq mm. Osterburg observed the relative frequencies of 13 different ridge events in all of these cells. These events included no event (an empty cell), ending ridge, bifurcation, island, dot, and so forth. He calculated the rarity of these events. Notably, he only used 39 fingerprints to do so.

He then reasoned that the rarity of a fingerprint arrangement would be the product of all the individual minutiae frequencies and empty cells. Given a partial 72 sq mm fingerprint, if one has 12 ridge endings (each occupying 1 cell) and 60 empty cells, the probability of this event is $(0.766)^{60} (0.0832)^{12} = 1.25 \times 10^{-20}$, where 0.766 and 0.0832 are Osterburg's observed frequencies of an empty cell and a ridge ending, respectively.

Finally, Osterburg corrected for the number of possible positions this grouping of minutiae can take. This factor was dependent on the size of this partial fingerprint physically fitting into all the fully rolled fingerprint blocks on a tenprint card. Again referring back to the parking lot analogy, it is similar to taking a row of cars and empty spaces from a lot and seeing how many ways you can physically fit that chunk into the entire parking lot. This approach is somewhat similar to Amy's.

One of the largest problems with the Osterburg model is the assumption that each cell event is independent. For example, if a cell contains a minutia, it is unlikely that the surrounding eight cells will also contain minutiae. Minutiae generally do not all group together. Sclove recognized that the presence or absence of minutiae in a group of cells will influence the presence or absence of minutiae in neighboring cells. Sclove modified Osterburg's event frequencies to reflect this dependency.

The Stoney and Thornton Model (1985–1989). Chronologically to this point, knowledge of fingerprint individuality models in the fingerprint community was scarce. Stoney and Thornton, in part to satisfy a portion of Stoney's thesis requirement, critically reviewed all the previously mentioned models, noting each model's flaws and strengths (Stoney and Thornton, 1986a, pp 1187–1216). On the basis of their review, Stoney and Thornton then proposed a set of criteria that the ideal model would possess for calculating the individuality of a print, as well as determining the probabilistic strength of a match. Stoney and Thornton identified that the ideal model must include the following features:

1) *Ridge structure and description of minutiae locations*
Ridge counts must be considered for measuring distances between features. For features on the same ridge, linear distances should be used, provided there are acceptable tolerances for distortion. (Though this author would suggest, when clarity is sufficiently high, one could count the intervening ridge units, which would not be subject to linear distance distortion.)

2) *Description of minutia distribution*
Minutiae are not uniformly distributed across a fingerprint and can vary in density (as noted by Kingston) and conditional relationship (as noted by Sclove). An accurate distribution of minutiae for a specific region must be a property of the ideal model.

3) *Orientation of minutiae*
With the exception of the dot or very short ridge, minutiae possess an orientation along the ridge flow that must be considered.

4) *Variation in minutiae types*
Relative frequencies for minutiae must be considered and the ideal model should have consideration for the absence of minutiae (similar to the Osterburg/Sclove model).

5) *Variation among prints from the same source*
The ideal model should account for the flexibility of skin where some features (e.g., ridge flow and linear distances) would not be as robust as other features (e.g., minutiae location on a ridge and ridge counts between minutiae). Poor clarity, distortion, and variability within the source must all be considered.

6) *Number of orientations and comparisons*
The number of ways to orient a fingerprint fragment can vary. For example, a delta could logically be oriented in three different ways. Also, on an individual with a loop pattern on each finger and toe, and several deltas in the palms and on the soles of the feet, a single delta formation could be compared nearly 60 different ways to one individual alone. The more orientations a print can assume will result in more comparisons that are possible. As Amy

observed, the more comparisons that are performed, the more opportunities that occur for a false match.

The model proposed by Stoney and Thornton was a study of minutiae pairs, within the ridge structure of the print. They performed statistical analyses on 2645 minutiae pairs from 412 fingerprints (all male distal tips of thumbs) (Stoney and Thornton, 1987, pp 1182–1203) and attempted to meet all of the ideal conditions that they had proposed. They were able to meet most of their conditions and developed a model for describing minutiae (Stoney and Thornton, 1986b, pp 1217–1234).

In the Stoney and Thornton model, each pair of minutiae is described by the minutiae events (i.e., type of minutiae, orientation, intervening ridge count, and linear distance) and spatial position of the pair within the entire fingerprint pattern. The combination of all the minutiae pairs is a measure of individuality for that print. Thus Stoney and Thornton described a model that incorporated many of the essential components for determining the individuality of friction ridge arrangements.

Champod and Margot Model (1995–1996) (Champod and Margot, 1996a, 1996b; Stoney, 2001, pp 373–378). Until this point, all previous calculations and minutiae observations had been done by hand and involved small databases of fingerprints (Stoney and Thornton's model thus far used the largest database of 412 prints, albeit thumbtips). The Champod and Margot model was the first to utilize a computerized algorithm to process the fingerprint images. They used a database of 977 fingerprints composed of ulnar loops from the middle and index fingers and whorls from the middle finger.

Champod and Margot, similar to Stoney and Thornton, first performed a systematic statistical description of the minutiae in the fingerprints. They calculated the minutiae density and distribution of minutiae for various regions in the print, the frequencies of the minutiae types, the orientation of the minutiae, and lengths of compound minutiae (e.g., short ridges, enclosures).

Using their data, they then calculated probabilities for specific minutiae configurations and combinations. These probabilities indicate the probability of reoccurrence for a specific minutiae configuration and thus can be expressed as a measure of the strength of the match.

The Meagher, Budowle, and Ziesig Model (1999) (*U.S. v Mitchell,* July 8, 1999, pp 157–198; July 9, 1999, pp 29–139). This model, often referred to as the "50K versus 50K study", was an experiment conducted by the FBI in conjunction with Lockheed Martin, Inc., in response to the first *Daubert* challenge in *U.S. v Byron Mitchell*. This study has not been published, but descriptions of the study and data are found within the documents and testimony provided by Stephen Meagher, Bruce Budowle, and Donald Ziesig in *Mitchell*.

The primary experiment conducted by Meagher and colleagues utilized AFIS computer algorithms to compare each of 50,000 fingerprint images (all left loops from white males) against itself[6] and then the remaining 49,999 images in the database. The result of each comparison produced a score proportional to the degree of correlation between the two images. It is critical to note that all previous models possess calculations of individuality based on predicted minutiae arrangements; however, the scores in this model are a function of the AFIS algorithms and matcher logic.

Presumably, the highest score would result when an image is compared against itself. All of the other 49,999 comparison scores were then normalized (to fit a standard normal curve) to the highest score. The top 500 scores for each print were then examined. From these data, Meagher et al. concluded that, on the basis of the highest normalized score (averaged from all 50,000 trials), the probability of two identical, fully rolled fingerprints is less than 1×10^{-97}.

Meagher and colleagues conducted a second experiment, identical to the first, with the exception that in these trials, "simulated" latent prints were used. These simulated latent prints were cropped images of the original, showing only the central 21.7% area of the original image. The value of 21.7% was used because it constituted the average area of a latent print from a survey, conducted by this group, of 300 actual latent prints.

Each simulated latent print was searched against its parent image and the other 49,999 other images. The scores were calculated, ordered, and the top 500 scores examined.

[6] It is important to note that the image is compared against itself. Therefore, the model does not account for intraclass variability, that is, multiple representations of the *same fingerprint* showing variations in minutiae positioning due to distortion and stretching of the skin. This is not to say, for example, two inked prints from the same finger; rather, the image is literally compared against itself. One would obviously expect that the highest match score produced will be from the comparison of the image to itself. This was the case in all 50,000 trials. This important distinction is also a key point of criticism and is considered a fundamental flaw in the model by some reviewers (Stoney, 2001, pp 380–383; Wayman, 2000).

The scores were stratified for minutiae counts in the simulated latent prints; the counts of minutiae in these simulated prints ranged from 4 to 18 minutiae. Meagher and colleagues calculated probabilities of a false match in this second experiment ranged from 1×10^{-27} (for 4 minutiae) to 1×10^{-97} (for 18 minutiae).

The Pankanti, Prabhakar, and Jain Model (2001) (Pankanti et al., 2001, pp 805–812). The model proposed by Pankanti, Prabhakar, and Jain is more of an assessment for probabilities of false match rates in an AFIS model than an assessment for the individuality of a fingerprint. The model essentially calculates the number of possible arrangements of ridge endings and bifurcations, as seen from the view of an AFIS. However, an important new inclusion is the introduction of intraclass variation for a specific print (i.e., how much variance can be observed for a single fingerprint when several standards are taken from the same fingerprint).

Pankanti and colleagues determined the tolerance for minutiae from a database of 450 mated pairs. These images were pairs of the same fingerprint taken at least one week apart. For each minutia, the corresponding minutia was located in the mate. The spatial differences were calculated for all the corresponding minutiae in the pairs and, on the basis of the best fit of their data, they calculated the theoretical tolerance for locating minutiae. It is important to note that their calculated metric for tolerance is a spatial one (with linear [x,y] and angular [θ] components), not a ridge-based one (as previously noted by Stoney as a critical component). Thus in this model, the computer would accept "matching" minutiae if they possessed a similar location in space (x,y, θ) even if the ridge counts differed significantly from a fixed point.

Using an electronic capture device, Pankanti and colleagues collected a total of 4 images from each of 4 fingers from 167 individuals, for a total of 668 fingerprint images, each in quadruplicate. They repeated this process for a second capture device. They created two databases, one for each of the two capture devices. Given that each fingerprint in the database had four images of the same finger, captured separately, Pankanti and colleagues measured the differences in the minutiae locations for each image to determine the acceptable tolerance based on natural variations for that finger.

On the basis of these calculations, Pankanti and colleagues derived an expression to calculate the probability of a matching fingerprint pattern, given the specific size of a print and the number of minutiae available to match. They calculated that to match 36 minutiae out of an arrangement of 36 minutiae (similar to Galton's proposed 35 minutiae in an average print and including only ridge endings and bifurcations) the probability was 5.47×10^{-59}. To match any 12 of these minutiae, given the same parameters, the probability was 6.10×10^{-8}. (This, of course, implies that 24 of these minutiae do not match, and this would be unacceptable as a model for comparative analysis.) The group calculated the probability for matching all 12 minutiae, given only a 12 minutiae arrangement. This probability was 1.22×10^{-20}.

The group also calculated, using similar parameters and some basic assumptions, a table that was based on many of the previous models for the probability of matching 36 minutiae (considered by this model a full fingerprint) and 12 minutiae (12 on the basis of the "12-point rule", which some have attributed to Locard's tripartite rule). Amy's, Kingston's, and Champod's models were not included because these models were more complex than the other models and included variables not considered by this group (e.g., Kingston's inclusion of minutiae type).

The author of this chapter chose to perform calculations for eight minutiae, given his personal experiences. The author has witnessed examiners in the United States effecting individualizations with eight minutiae and little to no third-level detail. In effect, individualizations have been declared solely on an arrangement of eight minutiae, with minimal, if any, consideration for the frequency of the minutiae type, locale in the print (i.e., delta versus periphery), or complexity of the arrangement. The author calculated as a lower bound, on the basis of the equations provided by Pankanti and colleagues, probabilities for matching eight common minutiae from these models. Pankanti and colleagues' calculations, the author's additional calculations for eight minutiae using the Pankanti parameters, and select values for the remaining models not included by Pankanti and colleagues (i.e., Champod, Amy, Meagher, and Kingston) can all be found in Table 14–2 and the accompanying footnotes.

Summary of Probability Models. There are two very important comments that must be made when one examines the previous proposed probability models for individuality. The first comment is that no matter which model is chosen (and among all the experts who have visited this topic, it is quite clear), one can fairly quickly reach staggeringly small probabilities that two individuals will share an arrangement of minutiae. All of these models demonstrate

Table 14–2

Pankati and colleagues' calculations (with chapter author's additions).

	Probability of Matching a Specific Configuration of:		
Author and Year	**36 Minutiae**	**12 Minutiae**	**8 Minutiae**
Galton (1892)	1.45×10^{-11}	9.54×10^{-7}	6.06×10^{-6}
Henry (1900)	1.32×10^{-23}	3.72×10^{-9}	9.54×10^{-7}
Balthazard (1911)	2.12×10^{-22}	5.96×10^{-8}	1.53×10^{-5}
Bose (1917)	2.12×10^{-22}	5.96×10^{-8}	1.53×10^{-5}
Wilder and Wentworth (1918)	6.87×10^{-62}	4.10×10^{-21}	2.56×10^{-14}
Pearson (1930)	1.09×10^{-41}	8.65×10^{-17}	1.22×10^{-12}
Roxburgh (1933)	3.75×10^{-47}	3.35×10^{-18}	2.24×10^{-14}
Cummins and Midlo (1943)	2.22×10^{-63}	1.32×10^{-22}	8.26×10^{-16}
Trauring (1963)	2.47×10^{-26}	2.91×10^{-9}	2.04×10^{-6}
Gupta (1968)	1.00×10^{-38}	1.00×10^{-14}	1.00×10^{-10}
Osterburg et al. (1977–1980)	1.33×10^{-27}	3.05×10^{-15}	3.50×10^{-13}
Stoney and Thornton (1985–1989)	1.20×10^{-80}	3.50×10^{-26}	7.50×10^{-17}
Pankanti et al. (2001)[a]	5.47×10^{-59}	1.22×10^{-20}	1.56×10^{-14}
Amy (1946–1948)[b]	$<<6.2 \times 10^{-18}$	3.4×10^{-14}	1.8×10^{-8}
Kingston (1964)[c]	3.90×10^{-97}	3.74×10^{-32}	1.97×10^{-20}
Champod (1995–1996)	Two configurations:	Configuration #1: five ridge endings and two bifurcations = a probability of 2.5×10^{-5} Configuration #2: three ridge endings, one enclosure, one spur, and one opposed bifurcation = a probability of 7.0×10^{-10}	
Meagher et al. (1999)	4 minutiae = 1×10^{-27} 18 or more minutiae = 1×10^{-97} fully rolled print = 1×10^{-97}		

Notes

Using data and equations provided by Pankanti et al. (2001, pp 805–812) and based on the previously listed models, additional calculations have been made to include all the models listed in this chapter and the probabilities for arrangements of eight minutiae. With the exception of Champod, these calculations were based on ridge ending and bifurcation arrangements only and do not include rarer ridge events. In addition, with the exception of Roxburgh's "Quality Factor", none of the models account for clarity or the presence of third-level detail.

[a] Eight-minutiae probability calculated using the parameters (M, m, n, q) equal to (57, 8, 8, 8). The value for M was arrived at by an estimate of A based on an exponential fit to the data, which included all tolerance adjustments, provided in the Pankanti calculations (Pankanti et al., 2001, pp 805–812).

[b] Based on specific arrangements of empty ridges, groupings of bifurcations and ridge endings, and whether they were oriented to the left or right. The specific arrangements for each case are described by Amy (1946b, p 194). Amy's calculations only went as high as 15 minutiae, thus the value provided for 36 minutiae would be significantly smaller than the 6.2×10^{-18} as listed in Table 14–2.

[c] The author could not obtain Kingston's Poisson estimator for the expected number of minutiae/area, as these were empirically derived from Kingston's samples. Therefore, the values given in Table 14–2 correspond to the assumption that the number of minutiae observed in a region was equal to the expected number of minutiae for that region. The calculations are also based on assuming exactly half of the minutiae are bifurcations and half are ridge endings and using values for M (area) similar to those in Pankanti et al. (2001, pp 805–812).

that fingerprint minutiae are highly discriminating features, and, generally, the more minutiae that are shared between impressions, the less likely it becomes to randomly observe these features elsewhere in the population. Although AFIS technology and access to larger databases of images make this possibility more likely, it is still a rare event. Exactly "how rare" is what must be fleshed out. The technology and databases currently exist to adequately estimate these events.

The second comment is that these models have not been validated. The staggeringly low probabilities proposed by the models have not been tested in real-world, large databases. These probabilities may be accurate or they may grossly underestimate or overestimate the truth. It is simply an unknown at this time. The models have value and are important to the development of the discipline, of course. But the fundamental steps of testing, validation, and then refinement, followed by further testing and validation—the very fabric of scientific testing that was outlined at the beginning of this chapter—is missing. Stoney has aptly noted (Stoney, 2001, p 383):[7]

> From a statistical viewpoint, the scientific foundation for fingerprint individuality is incredibly weak. Beginning with Galton and extending through Meagher et al., 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. None of the models has been subjected to testing, which is of course the basic element of the scientific approach. As our computer capabilities increase, we can expect that there will be the means to properly model and test hypotheses regarding the variability in fingerprints.

It is imperative that the field of fingerprint identification meets this challenge. Although the theory of biological formation certainly supports the notion of friction ridge skin individuality, it must be supported by further empirical testing. Statistical modeling is a crucial component to achieving this goal, and more research and study in this arena is needed.

[7] Pankanti et al. (2001) was published contemporaneously with Stoney's comment, and, therefore, exclusion of Pankanti et al. (2001) was not an oversight or error by Stoney.

All of the previous models dealt exclusively with minutiae configurations. With respect to sweat pore location, significant advances have occurred since Locard's time. Ashbaugh rekindled interest in pores with case examples of sweat pore use for individualization purposes (Ashbaugh, 1983, 1999). Ashbaugh described two methods for comparing pore location (Ashbaugh, 1999, pp 155–157). Significant contributions to sweat pore modeling have been advanced by Roddy and Stosz (Stosz and Alyea, 1994; Roddy and Stosz, 1997, 1999). Most recently, Parsons and colleagues reported further enhancements to pore modeling (Parsons et al., 2008). They concluded that sweat pore analysis can be automated and provide a quantitative measure of the strength of the evidence.

14.3.2 Persistence

14.3.2.1 Persistence of First- and Second-Level Detail. Although Herschel and Faulds were two of the most prominent early pioneers investigating the persistency of friction ridge skin, it was Galton who provided the first actual data and study. Herschel and Faulds claimed to have examined hundreds, perhaps thousands, of prints to reach this conclusion. Herschel had been employing fingerprints for identifications for approximately 20 years and he had noticed no apparent changes in the ridge formations.

Using a collection of inked prints provided by Herschel, Galton, on the other hand, conducted a very thorough investigation into every single minutiae present in the finger (and in some instances palmar) impressions from 15 individuals (Galton, 2005, pp 89–99). The longest interval between subjects was 31 years; the shortest interval was 9 years. Interestingly, Galton noted a single instance where a discrepancy existed (Galton, 2005, p 97). In this instance an inked impression taken from a young boy (age 2 1/2) was compared against an impression from the same finger when the boy was 15. In the earlier print, a bifurcation is visible that is not present in that region (that is, the ridge is continuous) in the later impression (Figure 14–1). Galton compared, in total, approximately 700 minutiae between these time intervals. He found only the one instance of a discrepancy. Misumi and Akiyoshi postulated that changes in the dermal substructure may have caused the anomaly observed by Galton (Misumi and Akiyoshi, 1984, p 53). They observed several changes with age in the dermal substructure (e.g., papillae proliferation and changes in adhesive forces between the epidermis and dermis) that may affect the appearance of the epidermal ridges and furrows.

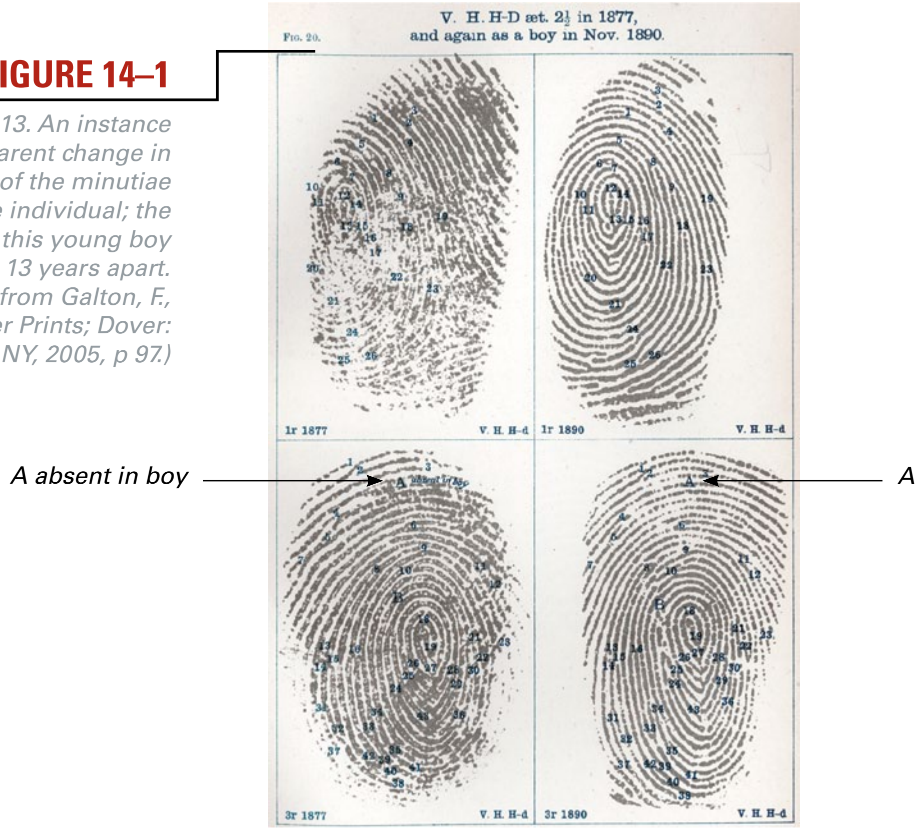


FIGURE 14–1

Galton's Plate 13. An instance of an apparent change in the appearance of the minutiae for one individual; the impressions of this young boy were taken 13 years apart. (Reprinted from Galton, F., Finger Prints; Dover: Mineola, NY, 2005, p 97.)

Wilder and Wentworth (1932, pp 126–131) performed a similar study on the minutiae of one subject, taking prints in approximately 2-year intervals from a young girl starting at 4 years and 11 months old until she was 14 years and 6 months old. Amongst these six time periods of collection, no change was observed in the minutiae of the subject. However, Wilder and Wentworth did note a proliferation of visible incipient ridges as the subject aged. This phenomenon has been observed and explored elsewhere (Stücker et al., 2001, 857–861).

Other instances where impressions have been examined for persistence after extended intervals have been noted in the literature. Herschel made successive impressions of his own fingerprints, starting at age 26, and throughout his life until age 83 (57 years in total) (Cummins and Midlo, 1943, p 40). No changes in minutiae were observed. Welcker (Cummins and Midlo, 1943, pp 40–41) made impressions of his fingers and palms at age 34 and then again later at age 75 (a 41-year interval). Another case is reported by Jennings (Cummins and Midlo, 1943, p 41) of palmprint impressions compared 50 years apart (taken at age 27 and then again at age 77). Finally, Galton continued to investigate the persistency of skin, increasing the number of individuals he compared to 25, with the longest time span being 37 years between prints (Wilder and Wentworth, 1932, p 128). With the exception of Galton's single instance, no other investigator reported any changes in minutiae.

14.3.2.2 Persistence of Third-Level Detail and Creases. With respect to pores, Locard (1913, pp 530–535) noted that the relative positions of the pores remain unchanged throughout life. Meagher, in a *Daubert* hearing, provided images of a latent print and an inked print, said to be from the same donor with an interval of 10 years (Figure 14–2). The images of the prints contained only two minutiae, but an extraordinary amount of clarity, clearly showing edges and pores. The third-level detail remained unchanged in that 10-year span.

However, the example provided by Meagher is anecdotal. The current literature lacks a comprehensive study demonstrating the persistence of third-level detail. More specifically, what is missing for latent print examiners is a comprehensive study, over a long period of time, demonstrating the persistence of third-level detail in impressions captured from the friction ridge skin.

Persistency of palmar flexion creases was observed by Herschel (Ashbaugh, 1999, p 190). Ashbaugh compared

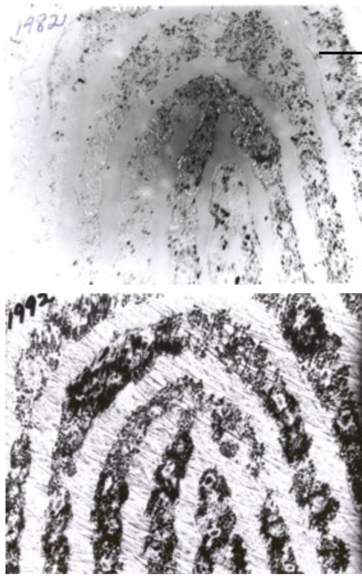


FIGURE 14–2

Exhibits 5-14 and 5-15 from U.S. v Mitchell, Daubert Hearing, July 8, 1999, testimony of Stephen Meagher. The image on the top is a perspiration print left on glass in 1982. The image on the bottom is an inked impression on paper from the same donor taken in 1992.

50 sets of palmprints taken from subjects at two different times, ranging from intervals of 1 to 60 months (Ashbaugh, 1999, p 189). Ashbaugh found that the flexion creases were in agreement, but noted some variation in appearance or prominence due to age, flexibility of skin, or other typical factors. Similarly, Evin and Luff (Ashbaugh, 1999, pp 193–194; Luff, 1993, p 3) reported persistency of palmar flexion creases after performing 600 comparisons (from roughly 100 individuals) with significant times between sample collection.

14.3.2.3 Theory Supporting Persistency of Friction Ridge Skin. The biological mechanisms for maintaining friction ridge skin persistency lie directly in the regenerating layer of skin found at the interface of the dermis and epidermis. This layer is known as the basal layer or stratum basale (germinativum). The persistency of the friction ridge skin is maintained by the basal layer and the connective relationship of these cells through desmosomes and hemidesmosomes. Wertheim and Maceo have reviewed and presented supporting pertinent medical research in this area (Wertheim and Maceo, 2002, pp 35–85; see Chapters 2 and 3).

14.3.3 Comparison Methodology

14.3.3.1 Overview of Comparison Methodologies. With respect to a *Daubert* challenge, at issue for admissibility of the evidence is whether the scientific principles or methodology upon which the conclusions are based are reliable. The previous sections have demonstrated core research supporting the basic principles of friction ridge skin science (i.e., uniqueness and persistence). The second half of this equation is the comparison methodology employed to compare two images, usually a latent print and a known exemplar.

It must first be noted that although ACE-V methodology is the generally accepted methodology in the United States (SWGFAST, 2002, p 2), Canada, Australia, and New Zealand, ACE-V methodology is not the *only* methodology available. For example, many European countries subscribe to the "Method for Fingerprint Identification" as described by the Interpol European Expert Group on Fingerprint Identification (IEEGFI) (IEEGFI-II, 2004). Although this methodology is very similar in most aspects to ACE-V methodology, it has some notable differences.[8] Additionally, probabilistic methodologies have been suggested by some authors (Locard, Stoney, Evett and Williams, Champod), but presently, this approach has been generally rejected as a viable methodology worldwide by examiners and professional bodies representing examiners (SWGFAST, 2002, p 4;

[8] For example, although creases, scars, and incipient ridges are completely acceptable features alone on which to make an individualization under the philosophy of ridgeology as applied during ACE-V, these features are not allowed as the sole basis for individualization under the IEEGFI methodology. These features may be used to add more weight to minutiae, depending on their relationship, but minutiae must be present. Additionally, under the IEEGFI-II, minutiae are subjectively weighted by the examiners on the basis of their frequency, location, and adjacent ridge features, and in this role, third-level detail and accidental features may be used to enhance the weight of minutiae. The IEEGFI method is quite innovative and thorough in its instructions for weighting minutiae. A weighting scheme based on the specificity of the features present is not explicit in the ACE-V methodology, though in fairness, may be applied by some examiners, knowingly or subconsciously, during the evaluation stage of ACE-V.

IEEGFI-II, 2004; Ashbaugh, 1999, p 147;[9] IAI, 1979, p 1). In fact, the penalty for using a probabilistic approach is so harsh that an expert found to give opinions of "probable, possible, or likely individualization" can be decertified and denied continued membership in the IAI (IAI, 1979, p 1). Academically speaking and from a perspective of evolving paradigm shifts in forensic science, exploring the viability of probabilistic evidence may have its benefits. Such efforts should not be summarily dismissed by the profession, because these methods may produce tools to aid or enhance current practices.

14.3.3.2 Research Pertaining to Fingerprint Comparison Methodology. Presently, there are few studies in the literature directly pertaining to the testing and validation of fingerprint comparison methodology. In fact, such works cannot be found prior to the 1993 *Daubert* decision.

Osterburg (1964). Osterburg conducted the first published survey of latent print examiner practices (Osterburg, 1964, pp 413–427). He sent surveys to 180 agencies throughout all 50 states. He received responses from 82 (46%). The surveys asked experts to subjectively rank the relative frequency of 10 types of minutiae characteristics (ending ridges, trifurcations, spurs, islands, etc.) based solely on the expert's training, experience, and personal recollection. Osterburg tabulated the ranked features. He also conducted a literature search to determine the minimum number of minutiae (points) needed to effect a positive identification (individualization). At the time, he found that individuals and agencies used between 6 and 18 minutiae to reach an individualization; the mean response was 12. He found that when experts were willing to reach an opinion below 12 minutiae, it was because they had "unusual characteristics". His study was an attempt to determine what an expert meant by "unusual". Years later, Osterburg and colleagues (1977) empirically measured the frequency of these features. The empirical counts of these features were very similar to the experts' intuitive assessment of rarity.

Evett and Williams (1996). The first actual study of fingerprint comparison methodology was performed by Evett and Williams (1996, pp 49–73). Their research, though conducted in 1988–1989, was not published until 1996, although it was presented at an international symposium in Ne'urim, Israel (Grieve, 1995, p 579). Their work predated the widespread knowledge, articulation, and general acceptance of ACE-V methodology among examiners. Evett and Williams investigated the basis for the 16-point threshold in place at the time in England and Wales. In their study, 10 sets of comparisons were provided to and returned by 130 experts from various bureaus in England and Wales. In addition, the researchers visited bureaus in the United States, Canada, Holland, France, and Germany. They provided experts in these countries with sets of comparisons as well, but did not include these results. They only reported the United Kingdom data, while giving the international results general commentary. The results of the United Kingdom data showed a surprisingly high level of variation among experts (Figure 14–3), not only in the reported number of corresponding minutiae that the expert saw, but also in whether the experts found sufficient agreement to determine an individualization. It is interesting to note that no expert reported an erroneous individualization. However, in one trial with two impressions that did originate from the same source, 8% of the United Kingdom experts erroneously excluded the images from having originated from the same source. Evett and Williams also found no statistical evidence that the number of individualizations reported by the United Kingdom experts was related to the years of experience of the examiner.

As a result of their research, the authors, while recommending standardization for training, certification testing, regular proficiency testing, regular audits of case files, and external blind proficiency testing, unequivocally stated that there is no need for a national predetermined numerical point standard if it can be demonstrated that each expert is operating above a minimum level of competence.

> Guidelines for individualization may be desirable, but these should be general recommendations and the expert should be allowed the freedom to exercise his/her own professional skills. In these circumstances, a rigid numerical point count is not only unnecessary, it is irrelevant. (Evett and Williams, 1996, p 72).

14.3.3.3 Error Rate Studies. With respect to the methodology, another testable *Daubert* factor is the known or potential rate of error (*Daubert,* 1993). In estimating latent print examiner error rates, some critics (Cole, 2005, pp 985–1078; Saks and Koehler, 2005, pp 892–895) have

[9] Ashbaugh notably does not specifically state that probability conclusions should not be produced. He merely states that "extensive study is necessary before this type of probability opinion can be expressed with some degree of confidence and consistency within the friction ridge identification science" (Ashbaugh, 1999, p 147).

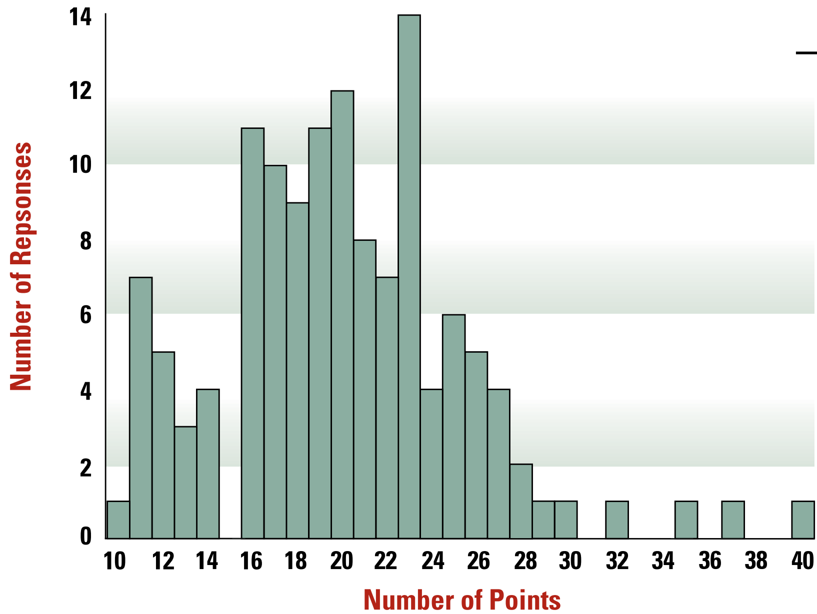


FIGURE 14–3

One graph (re-created) from the Evett and Williams (1996) study, depicting comparison of images marked "B". In reporting the number of minutiae found in agreement between the latent print and the known exemplar, respondents showed great variability. Most notable was the absence of any respondents reporting "15", which was one shy of the 16-point threshold to declare a positive match (for court) in the United Kingdom.

looked to performances of standardized latent print examiner proficiency tests administered through the external testing agency Collaborative Testing Services. Saks and Cole have also looked to anecdotal occurrences in case studies as indicators of a larger-than-reported error rate (Cole, 2005, pp 996–1034; Saks, 2005). Understandably, in the absence of any data produced from within the profession, they had little else to examine.

In an attempt to address the error rate issue, and thus provide the profession, the courts, and critics a better estimate of error than those previously available, Langenburg, Wertheim, and Moenssens conducted a two–stage error rate study (Langenburg et al., 2006, pp 55–92). During the first stage of the study, the researchers evaluated the comparison results of participants in a training course in which the participants compared friction ridge skin impressions (latent prints versus known exemplars). In the approximately 6000 comparisons performed by nearly 100 experts (as defined by the study, these experts possessed over one year of experience in comparing latent prints), the researchers found a total of 61 errors made at the highest level of confidence: 2 erroneous individualizations and 59 clerical errors. Although 59 errors were deemed clerical errors, 2 of these clerical errors wrongly associated the incorrect individual with the evidence; the other 57 were to the correct individual but listed the wrong finger or palm. Criteria were provided in the study for the determination of a clerical error versus an erroneous individualization. In the second stage of this study, 16 experts were asked to independently verify the results of a previous examiner. Each participant was provided with a packet that contained 10 comparisons and the stated results of a previous examiner. Eight of the individualizations for the verifier were accurate. Two of the results were errors and included one of the two erroneous individualizations from the previous stage. The other error would have been a clerical error or a second erroneous individualization, depending on which pack the participant randomly received. The verifier was not alerted that errors would be present in the verification packet. No expert verified any of the errors presented to them in this study. The study listed numerous limitations, most notably the absence of nonmatches (thus false negatives were not studied) and the fact that the experiments were not conducted under "casework" conditions.

Finally, it is important to note the empirical observations of forensic practitioners worldwide. Although these data cannot be readily seen in the literature, one must take into account the collective experiences of the tens of thousands of latent print examiners from around the globe during the last 100 years who have witnessed repeated success, application, and accuracy of the methodology during the training of new examiners, administration of internal competency tests, and other training tools (where the answers are known beforehand by the test administrator). Were the comparison methodology not very accurate, it would be commonplace to see errors frequently during the testing and measuring of examiner competency. This simply is not the case and has not been the author's experience in speaking with trainers here in the U.S. and abroad.

Although these empirical observations should not be dismissed, there are counterarguments to the weight of their support. The pros and cons of using proficiency testing data have been explored elsewhere (Saks and Koehler, 2005, pp 892–895; Langenburg et al., 2006; Cole, 2006b, pp 39–105; Gutowski, 2006). It has been argued that without the ground truth established for the comparison, anything else does not constitute a fair assessment of reliability (Cole, 2006a, pp 109–135). And even with the ground truth established in training exercises, without a standardized and validated model for comparison, the meaning of such results is questionable. For example, let us assume 10 experts all correctly individualize 10 latent prints to the correct 10 sources, for a grand total of 100 correct results and 0 errors. Presumably, these individualizations would exclude all other sources on the planet. The counterargument is that although these 100 conclusions were correct with respect to the ground truth, the relevant question becomes, Were there sufficiently discriminating features in agreement, and no observed differences, to *actually* exclude the world's population as the source of the latent prints? In other words, agreement among examiners is not necessarily de facto proof to support the strength of the evidence and the conclusion thus rendered.

14.3.3.4 Studies of Bias During Comparisons. A rising concern in the literature (Saks et al., 2003, pp 77–90; Steele, 2004, pp 213–240; Haber and Haber, 2004, pp 339–360), and in light of the Mayfield case (Stacey, 2004, pp 706–718), is the issue of whether biases affect the judgments and conclusions of forensic experts and specifically the judgments of the more subjective forensic comparative disciplines (i.e., handwriting, fingerprints, firearms examinations). Although there are many types of bias (e.g., culture, confirmation), some researchers are currently studying contextual information bias with respect to fingerprint examination.

The first study produced by Dror, Péron, Hind, and Charlton (2005, pp 799–809) found strong evidence that contextual information influenced the decision-making processes of nonexperts who participated in the study. Twenty-seven nonexperts (college student volunteers) were provided pairs of images (a latent print and a known exemplar) and asked whether the pair was a match. In addition to the images, the participants were exposed to varying levels of stimuli and contextual information. Dror and colleagues (2005) found that contextual information biased judgments when the matches were more ambiguous (i.e., had a lower quantity and quality of ridge detail or were look-alikes). They found that when the images were disparate in appearance and clear in detail, contextual information did not influence the participants. The group postulated that either fingerprint experts may be more resistant to these influences because of training and expertise or fingerprint experts may actually be more susceptible to these influences because of overconfidence and rationalization of differences.

A second study by Dror, Charton, and Péron (2006, pp 74–78) involved testing contextual information bias on five experts. For the study, the researchers selected five experts who were aware of the FBI's erroneous individualization in the Madrid Train Bombing case, but had not seen the actual images from the case. The experts were told that these images were from the Madrid Train Bombing case and had been incorrectly individualized by the FBI to Brandon Mayfield (Stacey, 2004, pp 706–718). The experts were asked whether they thought it was a valid match or was erroneous. However, the experts were not provided with images from the Mayfield case; rather, they were each provided with a pair of prints which that expert had personally individualized in casework 5 years prior to the study. Thus each expert was re-examining his own evidence. When provided with these images under the false contextual information, three of the five experts reversed their original opinions and stated the pair was not a match (exclusion), one expert changed his original opinion of a positive match to "inconclusive", and the final expert did not change his opinion but maintained a positive match, in spite of the strong contextual information. A number of concerns regarding the limitations of the study have been raised and discussed online (www.clpex.com), but the study suggests that experts are not immune to contextual information bias.

In the most recent study, Dror and colleagues (2006, pp 74–78) utilized a similar study design to the Madrid Train Bombing context-bias experiment. Six experts were presented their own previous work, but under less extreme circumstances of context bias than the previous study by Dror and colleagues (2005). Eight comparisons, on which the expert had previously provided conclusions several years prior to the study, were presented to each expert. Thus, there were 48 trials for the 6 experts. Twenty-four trials had no context bias and were control trials, 12 trials represented "easy" comparisons under routine bias, and 12 trials represented "difficult" comparisons under routine bias (see Table 14–3). Routine bias was represented by context bias that might be experienced by an expert in daily routine

Table 14–3

Results from Dror and colleagues' experiment.

	1	2	3	4	5	6	7	8
Past Decision	individual-ization	individual-ization	individual-ization	individual-ization	exclusion	exclusion	exclusion	exclusion
Level of Difficulty	difficult	difficult	not difficult	not difficult	difficult	difficult	not difficult	not difficult
Contextual Information	none	suggest exclusion	none	suggest exclusion	none	suggest individual-ization	none	suggest individual-ization
Expert A	consistent	consistent	consistent	consistent	consistent	consistent	consistent	consistent
Expert B	**change to exclusion**	consistent	consistent	consistent	consistent	consistent	consistent	consistent
Expert C	consistent	**change to exclusion**	consistent	consistent	consistent	consistent	consistent	consistent
Expert D	consistent	**change to exclusion**	consistent	**change to exclusion**	**change to individu-alization**	consistent	consistent	consistent
Expert E	consistent	**change to cannot decide**	consistent	consistent	consistent	consistent	consistent	consistent
Expert F	consistent	consistent	consistent	consistent	consistent	consistent	consistent	consistent

Note

Six experts were presented with eight comparisons on which they had previously rendered opinions. During the re-presentation, the comparisons were presented with context bias one might encounter in daily casework (knowledge of suspect confession, suspect criminal history, etc.).

(Reprinted from Dror et al., 2006, p 610.)

casework (a police officer's assertion of the suspect's guilt, knowledge of a confession, etc.). In the 48 trials, 6 trials resulted in responses that were not consistent with the original result provided by the expert. It is further interesting to note, of the six inconsistent results, two were in control trials (i.e., no context bias was provided). Dror and colleagues suggested two possible explanations for these inconsistencies in the control trials. The first possibility is that the experiment may not have been without bias even in the control conditions or, at a minimum, the conditions during the re-evaluation were not identical to the conditions under which the original decision was made. The second possibility is that there is less-than-ideal and less-than-expected reproducibility of expert results, even "within sample". In other words, the decision of an expert, when presented with the same evidence in multiple trials over time, may not be reproducible, and the expert is producing conflicting, inconsistent results. Dror and colleagues suggested further study of this phenomenon. With respect to the remaining four out of six trials of inconsistent responses, Dror and colleagues attributed these inconsistencies to the context bias in the trials, noting that three out of four inconsistencies reflected the bias prompt. However, as with the previous Madrid context-bias experiment, little to no information was provided about the experts or the presentation of the images to the experts, nor are the images available for review.

In contrast to the effect Dror and colleagues observed with respect to the *evaluation* of a latent print and an exemplar (i.e., the decision resulting in an individualization, exclusion, or inconclusive opinion), Schiffer and Champod (2007) reported no effect due to context bias in the *analysis* phase. Schiffer and Champod provided forensic science students at the University of Lausanne, Switzerland, with images of latent prints prior to the students' formal instruction series. Two experiments were conducted. The first experiment provided 39 students with 12 images of latent prints. The students were asked to annotate the minutiae in the images using a standard guideline. Upon completion of an intensive fingerprint instruction course, 29 of these students were provided with the same images to annotate again. Schiffer and Champod found a statistically significant increase in the number of minutiae reported and a decrease in the variation among student responses. Additionally, the number of reported instances declaring the print "exploitable" (i.e., "of value") and "identifiable" significantly increased after the training period. In the second study, 11 images of latent prints were provided to 2 groups of students (48 total students) after the fingerprint instruction course. The images were presented to the students under various context bias circumstances: no bias, presence of a matching exemplar, low-profile property crime case, high-profile terrorist case, and so forth. Students were asked to annotate the images and report the value of each print. Schiffer and Champod reported no difference for any of the factors examined between the two groups. They argued that not all stages of the ACE-V process are similarly vulnerable to bias, and their results supported the robustness of the analysis phase.

14.4 Future Directions for Research Related to Friction Ridge Examination

14.4.1 United States Government-Sponsored Research Available for Accepted Grant Applicants

Although some professional bodies (e.g., the Robert L. Johnson Foundation, created by the IAI) offer small stipends for research, these funds are generally not sufficient to conduct a large-scale study (e.g., a validation study) or a complicated study (e.g., the development of a quantitative model for measuring distortion), which would undoubtedly involve multiple experts and statisticians, a large computerized database, and software and hardware appropriate to the tests. Government agencies or academic institutions must properly fund this research. One agency which has supported open proposals for large-scale friction ridge research is the National Institute of Justice (NIJ). NIJ issued solicitations for Research and Development on Impression Evidence in 2009 and for Research and Development on Pattern and Impression Evidence in 2010. Both solicitations yielded a number of responsive friction ridge analysis project proposals, and multiple grant awards were made for both years (information is available at www.ojp.usdoj.gov/nij/awards/welcome.htm).

14.4.2 Recommended Topics for Research

The Scientific Working Group for Friction Ridge Analysis, Study, and Technology (SWGFAST) has posted on its Web site (www.swgfast.org) a list of recommended areas for study and research.

Another source for recommended research was provided by Budowle, Buscaglia, and Perlman (2006). Some of their notable "high-priority" recommendations include:

- Develop guidelines for describing the quality of ridge features in an image.
- Develop guidelines for sufficiency in declaring a positive match.
- Determine the minimum number of features (if any) that are needed pragmatically for an examiner to declare a positive match in casework.
- Rigorous testing (validation) of the ACE-V methodology as applied by experts.
- Testing for persistence of third-level features.

Many of their suggestions should be strongly considered by serious researchers, because the results of the work could be extremely beneficial and enlightening to the friction ridge identification discipline.

Another major area that needs to be addressed is an objective understanding of distortion and the development of an acceptable metric for tolerance. It was clear from statements made by the investigating bodies in the Brandon Mayfield case (Office of the Inspector General, 2006, pp 6–10) that the examiners had discounted dissimilarities between the latent print and Mayfield's exemplar. However, *a posteriori,* it was determined that the dissimilarities were outside of acceptable tolerance and an exclusion should have been the correct conclusion. Determining acceptable ranges of tolerances, or determining an appropriate weighting scheme for a feature based on the feature's departure from "normality" due to distortion, would be critical updates to any comparison methodology.

Finally, as previously discussed and highlighted by Stoney (see page 14–15), the development of a more complete probability model for fingerprint individuality is needed. The development of this model must be followed up by empirical testing of the model with real-world samples and large databases.

14.5 Conclusions

In a post-*Daubert* environment, there is a need for additional research in the field of friction ridge science. Certainly, any science wishes to expand the depth and breadth of knowledge of the discipline. We in the fingerprint expert community must attempt to challenge and study further the laws and theories that comprise our discipline. Specifically, we must focus our efforts to reevaluate the basic tenets of individualizing friction ridges using modern and enhanced technologies that were not available in Galton's day. There are many unanswered or partially answered questions regarding the individuality of friction ridge skin and the forensic comparison of friction ridge impressions. Although significant advances have been made, many of them in just the last two decades, this is really only the tip of the iceberg. With the advent of newer, more powerful technologies, software, and computer algorithms, we have opportunities to explore our vast fingerprint databases and quickly growing palmprint databases. We need to assess and quantify the full extent of variation of friction ridge features, starting with perhaps the most basic (patterns and minutiae—if one can truly call this "basic") and then attempt to assess and quantify other features such as creases, scars, edge shapes, and so forth.

It should be clear that there are aspects of this discipline that have been well-established and well-studied (particularly the biological theory of friction ridge formation and persistency). However, it should also be clear that there are areas of study that are woefully lacking (e.g., distortion, tolerance).

The absence of available published research into some aspects of the discipline speaks volumes about what our mission should be.

14.6 Reviewers

The reviewers critiquing this chapter were Leonard G. Butt, Christophe Champod, Deborah Friedman, Robert J. Garrett, Andre A. Moenssens, Michael Perkins, Jon T. Stimac, Michele Triplett, John R. Vanderkolk, and James L. Wayman.

14.7 References

Amy, L. Valeur de la preuve en dactyloscopie. *Journal de la Société de Statistique de Paris* 1946a, *88,* 80–87.

Amy, L. Valeur de la preuve en dactyloscopie II. *Journal de la Société de Statistique de Paris* 1946b, *88,* 188–195.

Amy, L. Recherches sur l'identification des traces papillaires. *Société de Médecine Légale* 1948, *28* (2), 96–101.

Ashbaugh, D. R. Poroscopy. *Royal Canadian Mounted Police Gazette* 1983, *45,* 12–17.

Ashbaugh, D. R. *Quantitative–Qualitative Friction Ridge Analysis: An Introduction to Basic and Advanced Ridgeology;* CRC Press: Boca Raton, FL, 1999.

Babler, W. J. Prenatal Selection and Dermatoglyphic Patterns. *Am. J. Phys. Anthropol.* 1978, *48* (1), 21–28.

Babler, W. J. Prenatal Communalities in Epidermal Ridge Development. In *Trends in Dermatoglyphic Research;* Durham, N., Plato, C., Eds.; Kluwer Academic Press: Dordrecht, Netherlands, 1990; pp 54–68.

Babler, W. J. Embryologic Development of Epidermal Ridges and Their Configurations. In *Dermatoglyphics: Science in Transition;* Plato, C., Garruto, R., Shaumann, B., Eds.; March of Dimes: New York, 1991; pp 95–112.

Balthazard, V. De l'identification par les empreintes digitales. *Comptes Rendus, des Academies des Sciences* 1911, *152,* 1862–1864.

Budowle, B.; Buscaglia, J.; Perlman, R. Review of the Scientific Basis for Friction Ridge Comparisons as a Means of Identification: Committee Findings and Recommendations. *Forensic Sci. Communic.* 2006, *8* (1).

Carroll, J. W. Laws of Nature; 2003; plato.stanford.edu/archives/fall2003/entries/laws-of-nature/ (accessed June 30, 2009).

Chacko, L. W.; Vaidya, M. C. The Dermal Papillae and Ridge Patterns in Human Volar Skin. *ACTA Anatomica (Basel)* 1968, *70* (1), 99–108.

Champod, C. Edmond Locard—Numerical Standards and "Probable" Identifications. *J. Forensic Ident.* 1995, *45* (2), 136–163.

Champod, C.; Margot, P. Analysis of Minutiae Occurrences on Fingerprints—The Search for Non-Combined Minutiae. In *Proceedings of the 14th Meeting of the International Association of Forensic Sciences,* Tokyo, Japan, 1996a.

Champod, C.; Margot, P. Computer Assisted Analysis of Minutiae Occurrences on Fingerprints. In *Proceedings of the International Symposium on Fingerprint Detection and Identification.* Almog, J., Springer, E., Eds.; Israel National Police: Jerusalem, 1996b; pp 305–318.

Champod, C.; Lennard, C.; Margot, P.; Stoilovic, M. *Fingerprints and Other Ridge Skin Impressions;* CRC Press: Boca Raton, FL, 2004.

Chatterjee, S. K. Edgeoscopy. *Finger Print Ident. Mag.* 1962, *44* (3).

Cole, S. A. More than Zero: Accounting for Error in Latent Print Identification. *J. Crim. L. and Criminol.* 2005, *95* (3), 985–1078.

Cole, S. A. Is Fingerprint Identification Valid? Rhetorics of Reliability in Fingerprint Proponents' Discourse. *Law and Policy* 2006a, *28* (1), 109–135.

Cole, S. A. The Prevalence and Potential Causes of Wrongful Conviction by Fingerprint Evidence. *Golden Gate Univ. L. Rev.* 2006b, *37* (1), 39–105.

Collaborative Testing Services. *Latent Print Examination Report # 9508;* Collaborative Testing Services, Inc.: Herndon, VA, 1996.

Cummins, H. Dermatoglyphics: Significant Patternings of the Body Surface. *Yale J. Bio. Med.* 2000, *73* (1–6), 79–90.

Cummins, H.; Midlo, C. *Finger Prints, Palms and Soles: An Introduction to Dermatoglyphics;* Dover: New York, 1943.

Cummins, H.; Waits, W. J.; McQuitty, J. T. The Breadth of Epidermal Ridges on the Fingertips and Palms: A Study of Variation. *Am. J. Anat.* 1941, *68,* 127–150.

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

Donnelly, P.; Friedman, R. DNA Database Searches and the Legal Consumption of Scientific Evidence. *Mich. L. Rev.* 1999, *97* (4), 931–984.

Dror, I. E.; Péron, A. E.; Hind, S.; Charlton, D. When Emotions Get the Better of Us: The Effect of Contextual Top-Down Processing on Matching Fingerprints. *Appl. Cognitive Psych.* 2005, *19* (6), 799–809.

Dror, I. E.; Charlton, D.; Péron, A. E. Contextual Information Renders Experts Vulnerable to Making Erroneous Identifications. *Forensic Sci. Int.* 2006, *156* (1), 74–78.

Eaton, J. Fingerprints are Still Unique: A Pilot AFIS Study. Presented at the International Association for Identification Conference, Dallas, TX, August 11, 2005.

Eaton, J. King County Sheriff's Office. Seattle, WA: Personal communication, 2006.

Evett, I.; Williams, R. L. A Review of the Sixteen Points Fingerprint Standard in England and Wales. *J. Forensic Ident.* 1996, *46* (1), 49–73.

Faulds, H. On the Skin—Furrows of the Hand. *Nature* 1880, *22,* 605.

Faulds, H. Finger Prints: A Chapter in the History of Their Use for Personal Identification. *Sci. Am. Suppl.* 1911, *1872,* 326–327.

Faulds, H. *Dactylography or the Study of Finger-Prints;* Milner and Company: Halifax, 1912.

Frishberg, D.; Yohe, T. I'm Just a Bill (Song lyrics); Schoolhouse Rock, American Broadcasting Company: 1975.

Galton, F. *Finger Prints;* Dover: Mineola, NY, 2005.

Grieve, D. L. Symposium Report Israel National Police: International Symposium on Fingerprint Detection and Identification. *J. Forensic Ident.* 1995, *45* (5), 578–584.

Grieve, D. L. Possession of Truth. *J. Forensic Ident.* 1996, *46* (5), 521–528.

Gupta, S. R. Statistical Survey of Ridge Characteristics. *Int. Crim. Police Rev.* 1968, *218* (130), 130–134.

Gutowski, S. Error Rates in Fingerprint Examination: The View in 2006. *Forensic Bulletin* 2006, *Autumn,* 18–19.

Haber, R.; Haber, L. Error Rates for Human Latent Fingerprint Examiners. In *Automatic Fingerprint Recognition Systems;* Ratha, N., Bolle, R., Eds.; Springer Verlag: New York, 2004; pp 339–360.

Hale, A. Morphogenesis of Volar Skin in the Human Fetus. *Am. J. Anat.* 1952, *91* (1), 147–173.

Hempel, C.; Oppenheim, P. Studies in the Logic of Explanation. *Philos.Sci.* 1948, *15* (2), 135–175.

Henry, E. R. *Classification and Uses of Fingerprints;* 1st ed.; Routledge & Sons: London, 1900.

Huber, R. A. Expert Witness. *Crim. L. Quarterly* 1959, *2,* 276–295.

Huber, R. A.; Headrick, A. M. *Handwriting Identification: Facts and Fundamentals;* CRC Press: Boca Raton, FL, 1999.

Hughes, G. Losing Sight of the Shore. *J. Forensic Ident.* 1998, *48* (5), 611–615.

IAI-Resolution VII. *Identification News* 1979, *29* (8), 1.

IEEGFI-II. *Method for Fingerprint Identification;* Interpol European Expert Group on Fingerprint Identification II: Lyon, France, 2004.

Kimura, S; Kitagawa, T. Embryological Development of Human Palmar, Plantar, and Digital Flexion Creases. *Anatom. Record* 1986, *216* (2), 191–197.

Kimura, S; Kitagawa, T. Embryological Development and Prevalence of Thumb Flexion Creases. *Anatom. Record* 1988, *222* (1), 83–89.

Kingston, C. R. Probabilistic Analysis of Partial Fingerprint Patterns, Ph.D. Thesis, University of California, 1964.

Langenburg, G.; Wertheim, K.; Moenssens, A. A. A Report of Latent Print Examiner Accuracy During Comparison Training Exercises. *J. Forensic Ident.* 2006, *56* (1), 55–92.

Lin, C. H.; Liu, J. H.; Osterburg, J. W.; Nicol, J. D. Fingerprint Comparison I: Similarity in Fingerprints. *J. Forensic Sci.* 1982, *27* (2), 290–304.

Locard, E. Les Pores et l'identification des Criminels. *Biologica* 1912, *2* (24), 357–365.

Locard, E. La poroscopie. *Archives d'anthropologie criminelle, de médicine légale et de psychologie normale et pathologique* 1913, *28,* 528–546.

Locard, E. La preuve judiciaire par les empreintes digitales. *Archives d'anthropologie criminelle, de médicine légale et de psychologie normale et pathologique* 1914, *29* (145), 321–348.

Luff, K. Letter to Fingerprint Society. *Fingerprint Whorld* 1993, *19* (71), 3.

McCann, P. Interim Report of the Standardization Committee of the International Association for Identification. *Ident. News* 1971, *21* (10), 10–13.

McCann, P. Report of the Standardization Committee of the International Association for Identification. *Identification News* 1973, *23* (8), 13–14.

McComas, W. Ten Myths of Science: Reexamining What We Think We Know. ... *School Sci. Math.* 1996, *96,* 10.

McRoberts, A. Nature Never Repeats Itself. *The Print* 1996, *12* (5), 1–3.

Misumi, Y.; Akiyoshi, T. Scanning Electron Microscopic Structure of the Finger Print as Related to the Dermal Surface. *Anatom. Record* 1984, *208* (1), 49–55.

Moenssens, Andre A. Columbia City, IN. Personal communication, 2006.

Montagna, W.; Parakkal, P. *The Structure and Function of Skin,* 3rd ed.; Academic Press: New York, 1974.

Montagna, W.; Kligman, A.; Carlisle, K. *Atlas of Normal Human Skin;* Springer-Verlag: New York, 1992.

Office of the Inspector General. *A Review of the FBI's Handling of the Brandon Mayfield Case—Unclassified Executive Summary.* U.S. Department of Justice: Washington, D.C., March 2006.

Okajima, M. Frequency of Epidermal-Ridge Minutiae in the Calcar Area of Japanese Twins. *Am. J. Hum. Genet.* 1967, *19* (5), 660–673.

Osterburg, J. W. An Inquiry into the Nature of Proof: The Identity of Fingerprints. *J. Forensic Sci.* 1964, *9* (4), 413–427.

Osterburg, J. W.; Parthasarathy, T.; Raghaven, T. E. S.; Sclove, S. L. Development of a Mathematical Formula for the Calculation of Fingerprint Probabilities Based on Individual Characteristics. *J. Am. Stat. Associ.* 1977, *72* (360), 772–778.

Pankanti, S.; Prabhakar, S.; Jain, A. On the Individuality of Fingerprints. In *Proceedings of Hawaii International Conference on System Sciences (IEEE),* Maui, HI, December 2001; pp 805–812.

Parsons, N. R.; Smith, J. Q.; Thönnes, E.; Wang, L.; Wilson, R. G. Rotationally Invariant Statistics for Examining the Evidence from the Pores in Fingerprints. *Law, Probabil. and Risk* 2008, *7* (1), 1–14.

Pearson, K. Person Identification. In The *Life, Letters, and Labours of Francis Galton,* vol. IIIA; Cambridge University Press: Cambridge, England, 1930; p 182.

Popich, G; Smith, D. The Genesis and Significance of Digital and Palmar Hand Creases: Preliminary Report. *J. Pediatr.* 1970, *77* (6), 1017–1023.

Popper, K. The *Logic of Scientific Discovery;* Hutchinson Education: London, 1959.

Popper, K. *Conjectures and Refutations: The Growth of Scientific Knowledge;* Routledge & Kegan Paul: London, 1972.

Roddy, A. R.; Stosz, J. D. Fingerprint Features—Statistical Analysis and System Performance Estimates. *Proceedings of the IEEE,* 1997, *85* (9), 1389–1421.

Roddy, A. R.; Stosz, J. D. Fingerprint Feature Processing Techniques and Poroscopy. In *Intelligent Biometric Techniques in Fingerprint and Face Recognition.* Jain, L. C., Halici, U., Hayashi, I., Lee, S. B., Tsutsui, S., Eds.; CRC Press: Boca Raton, FL, 1999, pp 37–105.

Roxburgh, T. On the Evidential Value of Finger Prints. *Sankhya: Indian J. Stat.* 1933, *1* (50), 189–214.

Saks, M. Legal Consideration of Fingerprints. Presented at the 1st International Web Conference on Human Identification E-Symposium, April 14, 2005.

Saks, M.; Koehler, J. The Coming Paradigm Shift in Forensic Identification Science. *Science* 2005, *309* (5736), 892–895.

Saks, M.; Risinger, D. M.; Rosenthal, R.; Thompson, W. C. Context Effects in Forensic Science: A Review and Application of the Science of Science to Crime Laboratory Practice in the United States. *Sci. and Justice* 2003, *43* (2), 77–90.

Schiffer, B.; Champod, C. The Potential (Negative) Influence of Observational Biases at the Analysis Stage of Fingermark Individualisation. *Forensic Sci. Int.,* 2007, *167* (2–3), 116–120.

Sclove, S. L. The Occurrence of Fingerprint Characteristics as a Two-Dimensional Process. *J. Am. Stati. Associ.* 1979, *74* (367), 588–595.

Sclove, S. L. The Occurrence of Fingerprint Characteristics as a Two-Dimensional Poisson Process. *Communic. in Stati. Theoret. Meth.* 1980, *A9* (7), 675–695.

Srihari, S. N.; Srinivasan, H.; Fang, G. Discriminability of Fingerprints of Twins. J. *Forensic Ident.* 2008, *58* (1), 109–127.

Stacey, R. Report on the Erroneous Fingerprint Individualization in the Madrid Train Bombing Case. *Journal of Forensic Identification* 2004, *54* (6), 706–718.

Steele, L. The Defense Challenge to Fingerprints. *Crim. L. Bull.* 2004, *40* (3), 213–240.

Stoney, D. A. Measurement of Fingerprint Individuality. In *Advances in Fingerprint Technology;* Lee, H. C., Gaensslen, R. E., Eds.; CRC Press: New York, 2001; pp 327–387.

Stoney, D. A.; Thornton, J. I. A Critical Analysis of Quantitative Fingerprint Individuality Models. *J. Forensic Sci.* 1986a, *31* (4), 1187–1216.

Stoney, D. A.; Thornton, J. I. A Method for the Description of Minutia Pairs in Epidermal Ridge Patterns. *J. Forensic Sci.* 1986b, *31* (4), 1217–1234.

Stoney, D. A.; Thornton, J. I. A Systematic Study of Epidermal Ridge Minutiae. *J. Forensic Sci.* 1987, *32* (5), 1182–1203.

Stosz, J. D.; Alyea, L. A. Automated System for Fingerprint Authentication Using Pores and Ridge Structure. In *Proceedings of SPIE, Automatic Systems for the Identification and Inspection of Humans;* Mammone, R. J., Murley, J. D., Eds.; 1994, Vol. 2277; pp 210–223.

Stücker, M.; Geil, M.; Kyeck, S.; Hoffman, K.; Rochling, A.; Memmel, U.; Altmeyer, P. Interpapillary Lines—The Variable Part of the Human Fingerprint. *J. Forensic Sci.* 2001, *46* (4), 857–861.

SWGFAST. Friction Ridge Examination Methodology for Latent Print Examiners, version 1.01; 2002; www.swgfast.org.

Thornton, J. I. The Snowflake Paradigm. *J. Forensic Sci.* 1986, *31* (2), 399–401.

Thornton, S. Karl Popper. *The Stanford Encyclopedia of Philosophy;* 2005; plato.stanford.edu/archives/sum2005/entries/popper (accessed June 30, 2009).

Tietze, S. Witthuhn, K. *Papillarleistenstruktur der menschlichen Handinnenfläche.* Luchterhand: Berlin, Germany, 2001.

Trauring, M. Automatic Comparison of Finger-Ridge Patterns. *Nature* 1963, *197,* 938–940.

U.S. v *Mitchell.* CA-No. 96-407, *Daubert* hearing transcript, July 7–9, 12–13, 1999.

Van Fraassen, B. *Law and Symmetry;* Clarendon Press: Oxford, 1989.

Wayman, J. L. *When Bad Science Leads to Good Law: The Disturbing Irony of the* Daubert *Hearing in the Case of* U.S. v. Byron C. Mitchell; Biometrics Publications: San Jose State University, San Jose, CA, 2000.

Wertheim, K.; Maceo, A. The Critical Stage of Friction Ridge Pattern Formation. *J. Forensic Ident.* 2002, *52* (1), 35–85.

Wertheim, P. A. Scientific Comparison and Identification of Fingerprint Evidence. *The Print* 2000, *16* (5).

Wilder, H. H.; Wentworth, B. *Personal Identification;* The Gorham Press: Boston, 1918.

Wilder, H. H.; Wentworth, B. *Personal Identification;* 2nd ed.; T. G. Cooke: Chicago, 1932.

Zynda, L. Lyle Zynda's Lectures on the Philosophy of Science;1994; www.soc.iastate.edu/sapp/phil_sci_lecture00.html (accessed June 30, 2009).

15.8 Additional Information

Champod, C. *Reconnaissance automatique et analyse statistique des minuties sur les empreintes digitales.* Ph.D., Thesis, Universite de Lausanne, Institut de Police Scientifique et de Criminologie, 1996a.

Epstein, R. Fingerprints Meet *Daubert:* The Myth of Fingerprint "Science" Revealed. *S. Cali. L. Rev.* 2002, *75,* 605–658.

Holt, S. B. *The Genetics of Dermal Ridges;* Charles C Thomas: Springfield, IL, 1968.

Loesch, D. *Quantitative Dermatoglyphics: Classification, Genetics, and Pathology;* Oxford University Press: New York, 1983.

National Institute of Justice. *Forensic Friction Ridge (Fingerprint) Examination Validation Studies;* National Institute of Justice, U.S. Department of Justice; U.S. Government Printing Office: Washington, DC, 2000.

Okajima, M. Frequency of Forks in Epidermal-Ridge Minutiae in the Finger Print. *Am. J. Phys. Anthropol.* 1970, *32* (1), 41–48.

Okajima, M. Epidermal-Ridge Minutiae in the Hallucal Area. *Mitteilungen der Anthropologischen Gesellschaft in Wien* 1977, *34,* 285–290.

Okajima, M. Quantitative and Genetic Features of Epidermal Ridge Minutiae on the Palm of Twins. *Hum. Heredity* 1984, *34* (5), 285–290.

Ökrös, S. T*he Heredity of Papillary Patterns;* Publishing House of Hungarian Academy of Sciences: Budapest, 1965.

Roberts, D. F. Dermatoglyphics and Human Genetics. In D*ermatoglyphics—Fifty Years Later;* Wertelecki, W., Plato, C., Paul, N. W., Eds.; Alan R. Liss Inc.: New York, 1979; pp 475–494.

Samuels, J. Letter regarding "Forensic Friction Ridge (Fingerprint) Examination Validation Studies". National Institute of Justice. U.S. Department of Justice: Washington, DC, June 20, 2000.