School of Applied Sciences

An Industrial Revolution for Fingerprint Science?
The Impact of Cognition and Human Factors on Fingerprint Examiners: Implications for the Use of Fingerprint Examiner Expertise and Administration within Law Enforcement

David Ashley James Charlton

Submitted in Partial Fulfilment of the Requirements for Award of Doctor of Philosophy by Bournemouth University

2011

An Industrial Revolution for Fingerprint Science? The Impact of Cognition and Human Factors on Fingerprint Examiners: Implications for the Use of Fingerprint Examiner Expertise and Administration within Law Enforcement

David Ashley James Charlton

'The mind sets the limits to everything that we do'
Nicholas de Cusa (1450)

Abstract

Fingerprint analysis has been a cornerstone of law enforcement investigation for well over 100 years (Block 1970, Duncan 1942, Holt 1936, Beavan 2001, Sengoopta 2003, Cooke 1932 and Charlton et al 2007). Fingerprint evidence has rarely been challenged by either the public or the judicial system. However, the people entrusted to perform the analysis of fingerprints are increasingly being seen as the weak link in the chain of evidence by some commentators (Schneier 2003).

Factors that affect the mind and its cognitive processes such as context and emotional state have an impact on decision making associated with a multitude of human endeavours including the medical and the military professions. This has been largely ignored by the forensic domain at large and by fingerprint examination specialists in particular. Errors in the analysis of fingerprint evidence in high profile cases such as Shirley McKie and Brandon Mayfield (Thompson et al 2005, Zeelenburg 2008, McKie 2007) have resulted latterly in legal counsel, media and the public asking whether fingerprint examination is valid and safe forensic science (Saks et al 2005). Psychologists have highlighted potential weaknesses in the policing domain and continue to make recommendations for improvement in order to minimise the risk of wrongful convictions (Adler 2004).

This thesis suggests that fingerprint examiners are not only emotionally driven and motivated to achieve results, but also that the motivations of examiners (Kruglanski et al 1983, Kruglanski et al 1987, Kruglanski 1989) exert leverage upon decision making thresholds. The consistency of fingerprint examiner observations during the analysis of fingerprints is also observed, in addition to the performance of fingerprint examiners during their interaction with colleagues and technology.

This thesis provides evidence of systemic practitioner inconsistency based on process and procedural weaknesses based on the cognitive realities that pervade the very nature of the work the human fingerprint examiners carry out. Through understanding of the human factors that impact upon the fingerprint examiner domain it has been possible in this thesis to offer insight and intervention that can mitigate against error and methodological breakdown of fingerprint analysis in the future, as well as to facilitate the design and implemetation of effective and robust recruitment, selection and training environments. In addition this thesis provides recommendations to facilitate the provision of fit for purpose technologies that are able to provide best practice for examiners and to satisfy public confidence in not only fingerprint examination but also other forensic and more generic expert policing domains.

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

List of Abbreviations

ACE-V	Assess Compare Evaluate Verify
ACPO	Association of Chief Police Officers
ASCLD	American Society of Crime Lab Directors
AFIS	Automated Fingerprint Identification System
ANSP	Air Navigation Service Providers
ATC	Air Traffic Control
BRC	Back Record Conversion
CAS	Competency Assessment Services
CRFP	Council for Recognition of Forensic Practitioners
CSI	Crime Scene Investigator
DNA	Deoxyribonucleic Acid
FBI	Federal Bureau of Investigation
FEAST	First European Air Traffic Controller Selection Test
FISH	Forensic Information Scanning Hub
FS	Fingerprint Society
FSS	Forensic Science Society
FSS	Forensic Science Service
GMCI	Generic Mark Camera Interface
HOSDB	Home Office Scientific Development Branch
IAI	International Association for Identification
ISO	International Organisation for Standardisation
MRI	Magnetic Response Imaging
NAFIS	National Automated Fingerprint Identification System
NIST	North American Institute of Science and Technology
NPIA	National Police Improvement Agency
PAF	Personal Appearance File
PCE	Phased Cognitive Engagement
PET	Positron Emission Tomography
SOCO	Scenes of Crime Officer
SWGFAST	Scientific Working Group Friction Ridge Analysis, Science and Technology

Preface

This thesis marks the culmination of seven years of research that has not only transformed my life but also has helped to re-shape the landscape of fingerprint examination within the judicial system on a global scale. As a fingerprint expert working for a UK Police Agency, my rationale for embarking on this journey was to answer questions around fingerprint examiner performance, motivation, cognitive vulnerabilities and the potential for group dynamic effects when making key decisions in fingerprint examination.

The findings within this thesis have invited professional abuse, even ridicule from the policing and judicial hierarchy. Many have sought to distance themselves from the implications of the research. Latterly, indifference has given way to acknowledgement, followed by pro-active embracing of the concepts behind bias, motivation and cognition within the forensic domain. The research in this thesis has been acknowledged by The British Psychological Society (Division of Forensic Psychology) who awarding me as author of this thesis, the Junior Prize for 2006 for the production of new forensic research of outstanding quality.

This research and thesis could not have been undertaken without the long term support of Dr Itiel Dror, who gave me the courage and conviction to keep going in the face of professional abuse. I also thank him for the mentorship and leadership that gave me the confidence to believe in myself and the research rationale and scope. I also thank Professor David Osselton for helping me through the final stages of this huge project. I also thank my senior management within Sussex Police, Scientific Support Managers Brian Cook and Louise Whiteoak, as well as the countless fingerprint examiners within law enforcement who allowed access to laboratories to gather data. Finally, and most importantly, I thank my wife, Hilary, and my children Heather and Alex, for their support, understanding and patience over the years as I could not have done this without them.

Since embarking on this project both of my parents, Jean and Charlie, have passed away and are not around to see the conclusion of this work. This thesis is dedicated in love and respect to my ‘Mum and Dad’ that provided the inspiration for me to realise my ambitions.

Declaration: This thesis is based on a large body of research in the domain of forensic science, specifically fingerprint examination. The research reported in this thesis represents the culmination of several years' long term personal research as well as collaboration with cognitive psychologists which has resulted in several shared authorship as well as numerous solo authorship peer reviewed publications. In all the research undertaken and reported on within this thesis I have conducted research design, ensured correct methodology, been the lead intermediary between academia and forensic practitioners. I have conducted extensive data collection as well as statistical interpretation of results and conclusions in the completion of this thesis. I have not knowingly copied (plagiarised) the work of others.

Chapter 1

The History of Fingerprint Science and a Review of Possible Vulnerabilities within the Domain:

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

This quote is made by John Massey (a senior fingerprint examiner at the FBI) who was referring to the Brandon Mayfield erroneous fingerprint identification that was perpetrated by him and others at the FBI as part of the investigation into the Madrid train bombing (Zabell 2005).

Introduction

To understand the climate within which fingerprint identification has evolved over the last 100 years or so, it is important to understand the relationship between law enforcement agencies and the examiners performing the analysis of fingerprints, and to review some of the legal and academic concerns on the validity of the science around fingerprint examination. By understanding the history and evolution of the science of fingerprint examination it will be possible to develop a framework upon which to introduce questions around the cognitive and human factors that impact upon the expertise associated with latent print examination that will form the basis of research within this thesis.

The Nature of Friction ridge skin

The majority of the skin on the human body is smooth and relatively featureless. This is not true of the palms of the hands, the fingers, toes and the soles of the feet. These areas of the human body are covered with a system of raised ridges, associated furrows, creases and sweat pores. The presence of such features is evidence of our evolutionary past (Charlton and Galloway 2007, Stocker 2007, Ashbaugh 1985, 1991, 1999, Vincent 1985, Faulds 1912, Galton 1892, Garn et al 1975). As tree dwelling apes we slowly evolved into the upright Homo sapiens we are today. Friction ridge skin gave our ancestors an evolutionary advantage for gripping objects such as rudimentary tools. To this day we still need to grasp items in order to maintain an everyday existence. Without friction ridge skin it would be difficult to hold on to a ladder, hold the steering wheel of a car, or even to hold a pen to write with.

Friction ridges form before birth (Babler 1977, 1978 and 1987, Vincent 1985, Garn et al 1975). The blueprint for friction ridge skin configuration is established very early on in the human embryo's development (Babler 1978 and 1987, Vincent 1985, Garn et al 1975). This blueprint, the configuration of the bifurcations and ridge endings and the distribution of the sweat pores and other microscopic features does not change except for injury, disease or decomposition after death. Only injury to the generating layer will affect the skin's ability to regenerate tissue.

The structure of friction ridge skin

Thick skin (which includes friction skin) has two principle layers. The epidermis is a layered, flat epithelial tissue five layers thick (Cummins 1929, Penrose 1980, Vincent 1985, Ashbaugh 1985, 1991 and 1999, Stocker 2007, Charlton and Galloway 2007). The dermis is much thicker than the epidermis and consists of two layers, the papillary layer, an area of loose connective tissue extending up into the epidermis as dermal pegs and the deeper reticular layer. The dermis sustains the epidermis. The boundary between the dermis and epidermis is a point of potential weakness. The two tissues can be separated from each other. This can be evident in cases of decomposition and long immersion in water where the outer layer of friction ridge skin can slough off.

The separation of the dermal layers was a feature of the bodies recovered in the Asian tsunami of 2004. Such was the length of time the bodies were immersed in water before eventual recovery, very often examiners were faced with hands that either had no epidermal ridge detail, leaving only sub-dermal ridges, or epidermal ridge detail that had sloughed off and separated from the dermis (see Figure 1).

Figure 1 Example of Dermal/Epidermal Separation Friction ridge growth (from personal collection)

Understanding friction skin growth during foetal development is important to the understanding of the persistency and uniqueness of friction skin. The biological uniqueness of fingerprints is core to the justification of the fingerprint profession in the scientific basis for the results of fingerprint comparisons. Scientific research in the field of embryology demonstrates why fingerprints are persistent and unique (W Babler 1977, 1978 and 1987, Galton 1892, Ashbaugh 1985, 1991 and 1999, Vincent 1985, Penrose et. al 1980, Faulds 1912, Samishenko 2001).

Similarities in fingerprint pattern can be seen quite often in siblings, especially identical twins, however, while identical twins have the same DNA, they will have different friction ridge configurations caused by external pressures and stresses (differential growth) which occur during foetal development (Babler 1987).

Chemical composition of latent prints

A latent finger mark comprises a mixture of natural secretions and contaminations (see Table 1) from the environment (Charlton and Galloway 2007, Stocker 2007, Kent 1993, Gerber 1983). Three types of glands are responsible for the natural secretions of the skin, the sudoriferous eccrine and apocrine glands, and the sebaceous glands. Eccrine sweat is approximately 98.5 per cent water, the remainder being principally made up of mineral salts, e.g., sodium chloride, organic acids, urea and sugars.

The palms of the hands and the soles of the feet produce only eccrine gland secretions, whereas the apocrine glands are located in the groin, in the arm pits, and in the perianal regions, where they generally open at the hair follicles. Sebaceous glands are found on the chest and the back, where they are associated with hair roots, and on the forehead, the lips of the vagina, and the glands of the penis. These glands secrete oil, the sebum, which serves to protect the skin and hair against water, to act as a lubricant, and also to help absorb fat, lipid, soluble substances (see Table 1).

Table 1 The major chemical constituents of the glandular secretions

The Major Chemical Constituents of the Glandular Secretions		
	CONSTITUENTS	
SOURCE	**INORGANIC**	**ORGANIC**
Eccrine Gland	Chlorides Metal ions Sulphates Phosphates Ammonia Water (>98%)	Amino Acids Urea Uric Acid Lactic Acid Sugars Creatinine Choline
Apocrine Glands	Iron Water	Proteins Carbohydrates Sterols
Sebaceous Glands		Fatty Acids Glycerides Hydrocarbons Alcohols

As the ridges of the hands are covered exclusively by eccrine glands, eccrine secretions are present to some degree in every latent fingerprint. Contamination by sebaceous gland secretions is also very common, whereas that from the apocrine glands is much rarer but may be important in certain crimes, e.g., crimes of a sexual nature. Sebaceous material is transferred onto the hands by contact between the hand and the forehead, the nose and the eye region of the head. Contaminants from the workplace which dirty the hands are also valuable when detecting latent prints.

Identification of common locations for prints

Police Officers are often disappointed with the lack of irrefutable fingerprint evidence that relies on the suspect's fingerprints being on a particular item of evidence. Unlike on television shows such as CSI, the suspect's prints don't always appear readily on some surfaces. Factors influencing the chances of obtaining prints assist us in understanding the fragile and elusive nature of latent impressions (Charlton and Galloway 2007). Each of the following factors independently or in combination can account for the lack of prints on any given surface:

- Individuals don't always have a sufficient quantity of perspiration and/or contaminates on their hands to be deposited.
- When someone touches something, they may handle it in a manner which causes the prints to smear.
- The surface may not be suitable for retaining the minute traces of moisture in a form representative of the ridge detail.
- The environment may cause the latent print to deteriorate.
- The perpetrator may be wearing gloves.

In considering a lack of fingerprints found at a crime scene, it should be noted that it does not suggest, imply, or establish necessarily that any person did or did not touch an item of evidence. Environmental conditions such as humidity will also impact upon the viability of latent finger mark survival at crime scenes. In the right humid conditions latent finger marks can survive for years before being powdered or treated with chemicals. However, a hot dry surface, such as you might find on a windshield of a car in Summer may degrade a viable finger mark in a matter of minutes or even seconds to a point where the material in question can never be recovered, by whatever means.

Early Beginnings of Friction Ridge Skin Analysis

There is evidence that society has been aware of the individuality of friction ridge skin for many centuries (Charlton and Galloway 2007, Ashbaugh 1999, Berry 1976). Neolithic carvings resembling friction ridge detail have been found in Brittany, France and similar carvings on granite slabs have been found in Ireland as well as carvings in Nova Scotia, Canada (see Figures 2 to 6).

Figure 2 Roman pot containing 2,000-year old cream or ointment, complete with finger prints. (Professor Vojtech Suk's collection, 1879–1967, Department of Anthropology, Faculty of Science, Masaryk University, Brno, Czech Republic, Journal of Ancient Fingerprints Number 1 2007).

Figure 3 6,000-year-old Coloured Paintings with Fingerprints Found in North China Caves (Professor Vojtech Suk's collection, 1879–1967, Department of Anthropology, Faculty of Science, Masaryk University, Brno, Czech Republic Journal of Ancient Fingerprints Number 1 2007).

Figure 4 Skin with epidermal ridges from the inner surface of second toe of the left leg – found on a mummy of a young female from Pompeii (Professor Vojtech Suk's collection, 1879–1967, Department of Anthropology, Faculty of Science, Masaryk University, Brno, Czech Republic Journal of Ancient Fingerprints Number 1 2007)

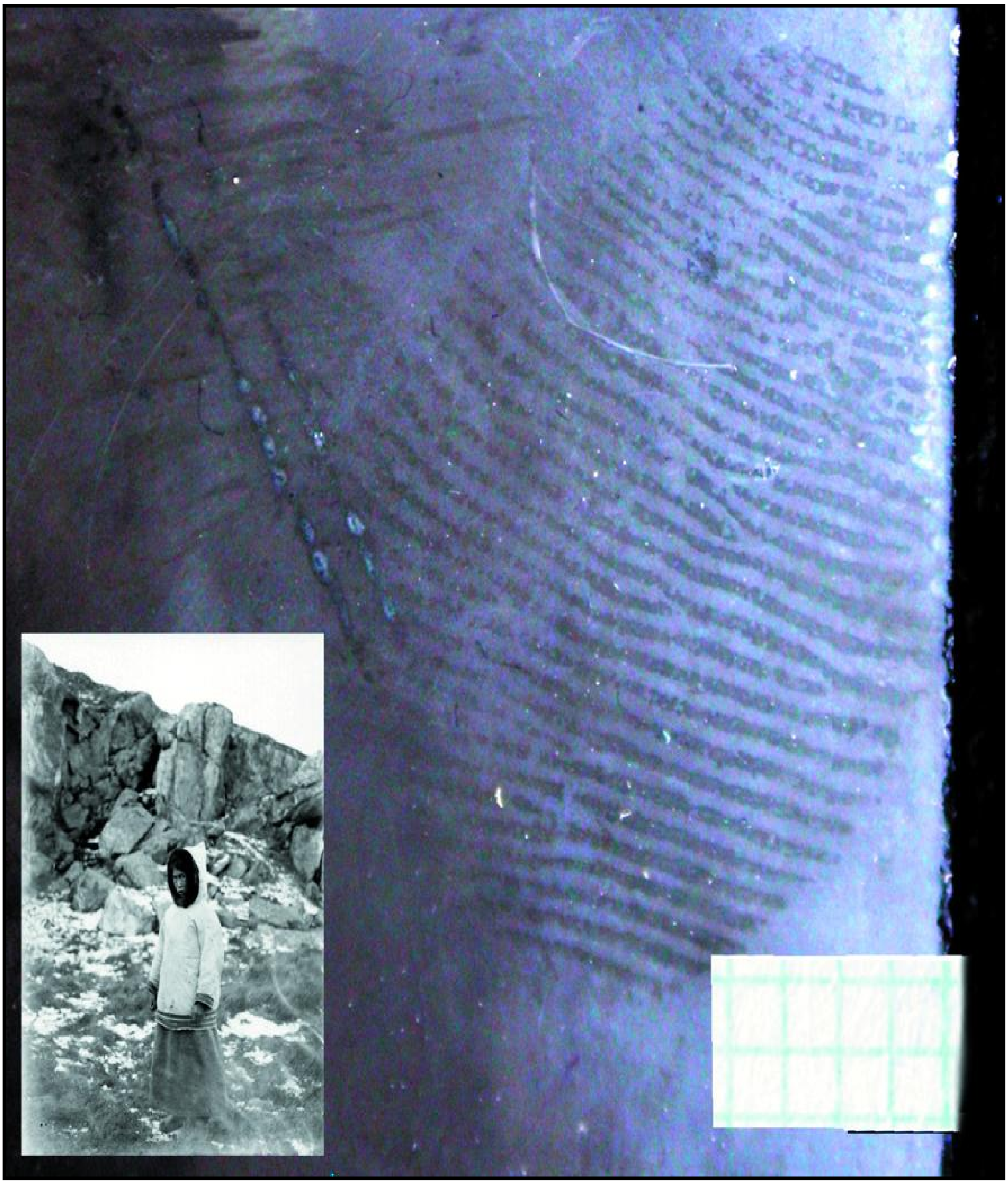


Figure 5 Fingerprint on a light-sensitive layer of a glass photographic plate (below left is the photograph - positive image). (Professor Vojtěch Suk's collection, 1879–1967, Department of Anthropology, Faculty of Science, Masaryk University, Brno, Czech Republic, Journal of Ancient Fingerprints Number 1 2007)

It is likely that these representations were made as recognition of physical attributes observed in the natural world, rather than any real knowledge of the power and individuality of friction ridges. Archaeological evidence from expeditions to the Middle East and Egypt discovered what were believed to be accidental artifacts (though some believed them to be deliberate markings) on items such as pottery and vases (see Figure 6).

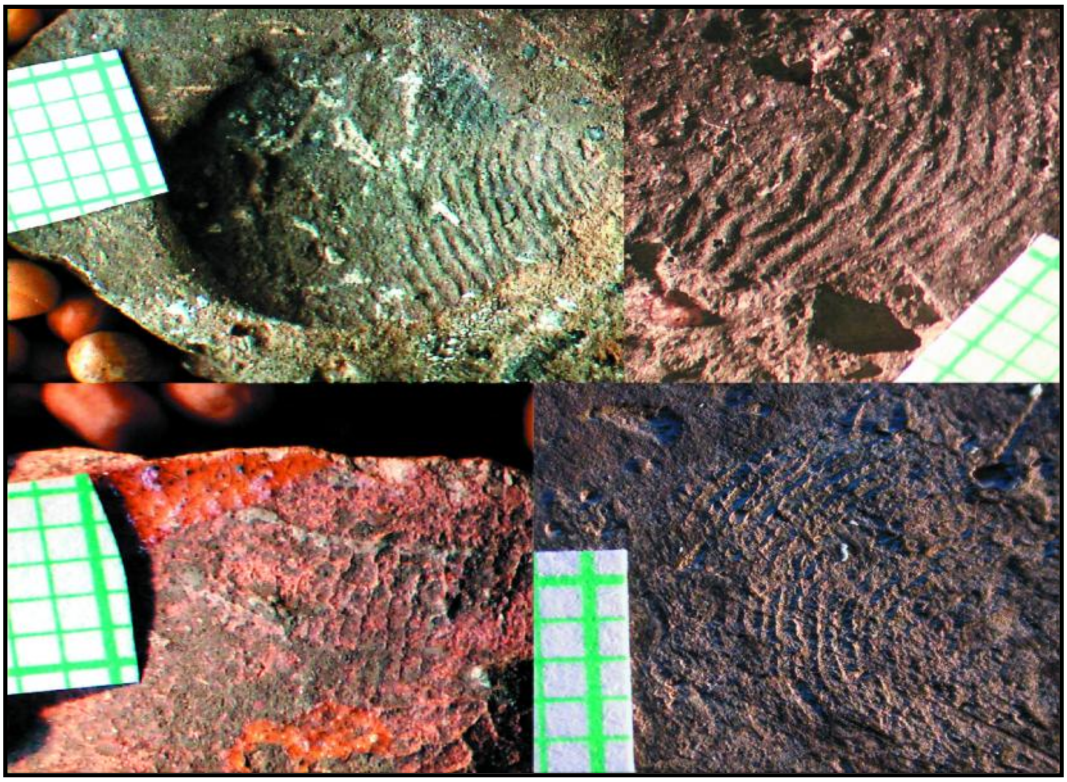


Figure 6 Examples of imprints common on ceramic pots from various periods (Martin Hložek Collection Journal of Ancient Fingerprints Number 1 2007)

It has been speculated that these markings were an attempt by the makers of such items to prevent forgery of individual designs and styles (see figure 6).

The earliest record of European literature associated with fingerprints is accredited to Neremiah Grew (Grew 1684, Charlton and Galloway 2007, Beavan 2001, Block 1970), fellow of the College of Physicians and of The Royal Society. Grew observed the patterns on the fingers and palms and described the ridge formation of pores and ridges. Italian anatomist Joannes Purkinje also described the formation of ridge detail (Cummins et al 1940, Charlton and Galloway 2007, Beavan 2001) and he was able to sub-divide the patterns observed into nine groups and to devise a rudimentary classification system.

Founders of Modern Friction Ridge Skin Analysis

Herschel and Faulds

Modern fingerprint science as we know it today only began to develop in the last few decades of the 19th century. While working as an administrator for the East India Company Sir William Herschel began the practice of recording handprints on contracts in the Hooghly District in India (Charlton and Galloway 2007, Sengoopta 2003, Beavan 2002). By his admission the taking of handprints of clients was totally unrelated to modern concepts of friction ridge skin and contained little usable ridge detail associated with the fingers for example. But as a method for repudiating fraud it was a system that worked efficiently.

Dr Henry Faulds' interest in fingerprints also developed in India, as well in Japan, while assessing the designs of ancient pottery in these regions. He also made comment on how fingerprints could be classified. He developed an understanding of the permanence of friction ridge skin and even went as far as to suggest that friction ridge skin analysis could be used for the detection of criminals at crime scenes. Both Herschel and Faulds asserted that they were the originators of modern fingerprint science. Faulds even went as far as to offer to test his methods at Scotland Yard, at his own expense; however, this offer was not accepted. Another system of human identification known as Bertillon Anthropometric measurement was introduced instead.

Galton and Bertillon

Sir Francis Galton is heralded as one of the greatest scientists of the 19th Century. An astronomer, explorer, inventor and relative of Charles Darwin, his primary obsession was anthropology, although he was also a student of psychology (Galton 1883). In 1888, Alphonse Bertillon invited Galton to Paris to prepare for Galton's lecture to the Royal Institution on 'Bertillonage' (see Figures 7 and 8), a system for identifying individuals by way of their body measurements (Charlton and Galloway 2007, Beavan 2002).

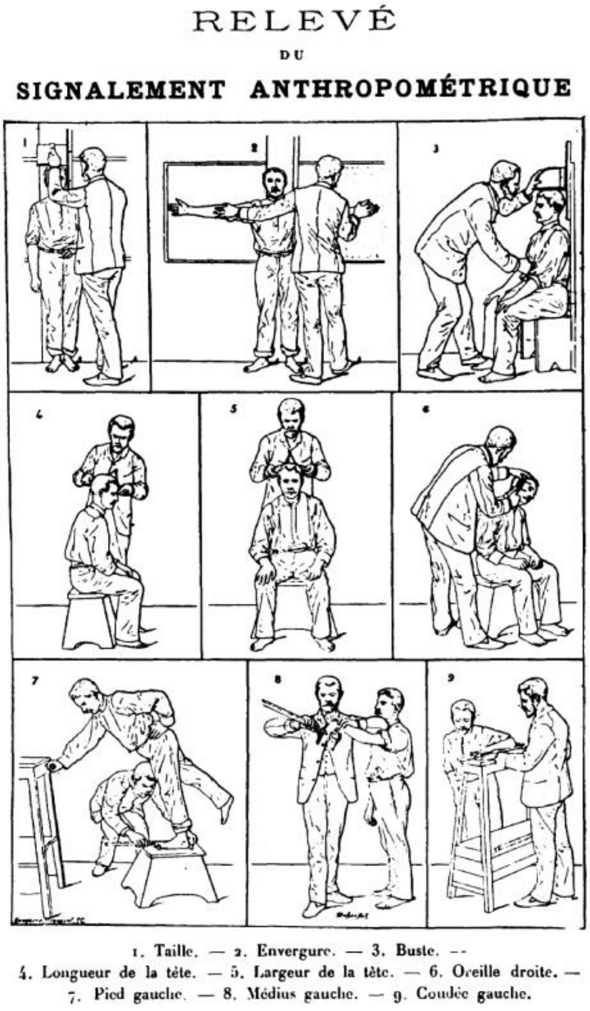


Figure 7 Alphonse Bertillon is seen by many as the father of modern biometric identification.

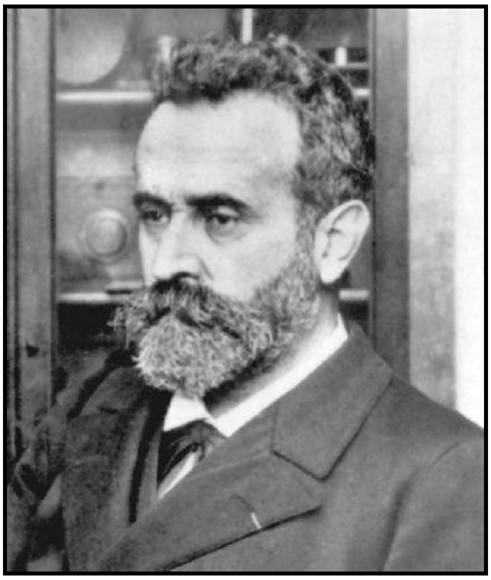
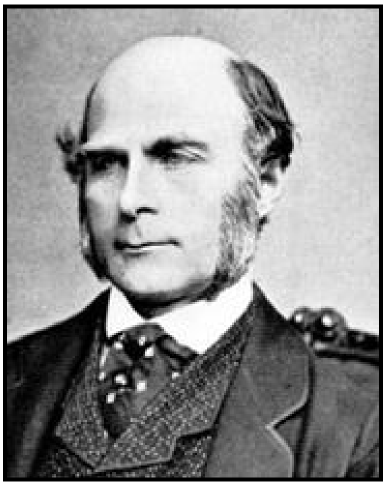


Figure 8 Alphonse Bertillon

(Figures 7 and 8 courtesy of the Fingerprint Society Archive)

Bertillon ignored fingerprints as a viable alternative because of the absence of a reliable classification system and the Bertillon method spread worldwide. Galton recalled his discussions with Hershel and gathered research on fingerprints as part of a presentation to the Royal Institution. At this meeting Galton (see Figure 9) was able to provide evidence of the persistence of friction ridge skin by showing the audience fingerprints taken from Hershel taken in 1860 and again in 1888.



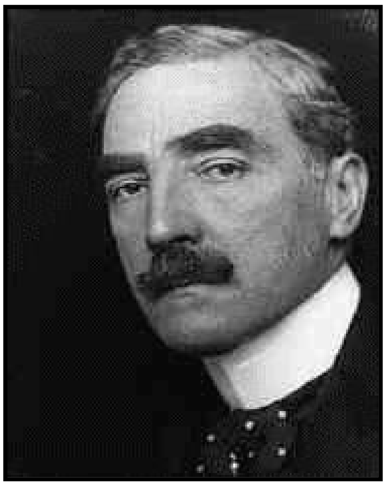


Figure 9 Francis Galton (Left)

Figure 10 Sir Edward Henry (Right)

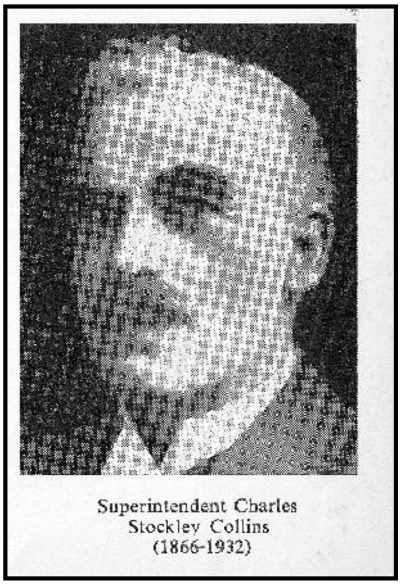


Figure 11 Juan Vucetich (Left)

Figure 12 Charles Stockley Collins (Right)

(Figures 9, 10, 11, 12 courtesy of the Fingerprint Society archive)

The prints showed the same characteristics and it was obvious to the audience that they came from the same person. From this time on fingerprints became a more popular human identifier. Bertillon even used fingerprints on the back of his anthropometric cards to add an extra biometric to the human measurements recorded.

In 1893 the English Home Secretary of the time appointed a committee chaired by Charles Troop to review the effectiveness of identification of the resident criminal population using both the Bertillon system of identification as well as fingerprints. The committee opting to stay faithful to the Bertillon system, employing Galton's fingerprint system only for subdivisions of criminal identification. Scotland Yard would maintain an anthropometric register but would fingerprint prisoners, classify them and use this data to ensure the prison population was representative of those who were supposed to be in custody and not impostors, which was a common occurrence at that time.

It was Galton (see Figure 9) who studied the heritability and racial differences in fingerprints. He wrote about the technique for identifying common patterns in fingerprints and devised a classification system that survives to this day (Galton 1892). This is why features such as ridge endings and bifurcations are known as 'Galton details'.

Henry and Vucetich

Sir Edward Henry (see Figure 10) had been working in India for many years and had developed an interest in fingerprints as a means of registration for Indian workers. He started experimenting with the anthropometric system, using ten measurements and noting eye colour and later refined the system using only six measurements but including a left thumbprint on the anthropometric record card (Henry 1900).

Henry had strong doubts about the accuracy of identification by measurement and became convinced that a record system based solely on fingerprints would solve many of the difficulties associated with Bertillionage. To increase his knowledge Henry wrote to Galton and in 1894 he returned to England where he visited Galton at his laboratory in the Convict Supervision Office. Galton gave Henry access to all that he had learned including the work of Herschel and Faulds. On his return to India Henry was determined to find a formula that could enable a fingerprint collection of several thousands to be classified, filed, searched and retrieved. His first task was to record all ten fingers of prisoners in addition to continuing to use the anthropometric system. He assigned two Bengali police officers Azizul Hacque and Chandra Bose, to study the classification problem (Sengoopta 2003). Their hard work was rewarded and a classification system, which allowed 1,024 primary positions with secondary and tertiary divisions, was developed. Both Hacque and Bose made a substantial contribution to discovering and setting up a classification system. It is ironic that today much of the credit is given to Henry for devising the classification system but most modern academics now give primary credit to Bose and Hacque (Sodhi et al 2002, Sengoopta 2003).

In 1897 Henry applied to the Indian Government for an independent evaluation of the classification system. As a result of the evaluation a resolution signed by the Governor-General directed that the official method for identifying criminals in British India by way of fingerprints be adopted (Garvie 1999). The anthropometric system continued along side the Henry system until gradually anthropometric cards were phased out and replaced by Henry classified fingerprint cards.

At the same time as Henry was experimenting with a fingerprint system in London, in Argentina, Juan Vucetich who was employed as a statistician in the Central Police Department at La Planta was also developing a fingerprint classification system (Beavan 2002, Charlton and Galloway 2007).

In 1891 Vucetich (see Figure 11) was appointed as head of the bureau of anthropometric identification. It did not take long for Vucetich to start experimenting with fingerprints and setting up his own equipment for taking prisoners fingerprints. Using Galton's basic material he formulated a ten-finger classification system.

In 1896 Argentina became the first country to abolish anthropometry and file criminal records by fingerprint classification alone. It was not until 1909 that Vucetich received recognition when he was awarded 'Perito Identificator' for his services to fingerprints by the President of the Argentine Republic.

In 1899 Henry was invited by the Association for the Advancement of Science to present a paper entitled 'Fingerprints and the Detection of Crime in India' (Henry 1900, Charlton and Galloway 2007, Beavan 2002). Giving due credit to Galton, the paper recalled the historical aspects of fingerprints and compared the anthropometric system with the fingerprint system. Henry's paper was well received.

Henry was very much in demand and while in England he had a request from the Indian Government to write a textbook on his method of classifying fingerprints (Henry 1900) as well as being reassigned to South Africa to organise a civil police force. At about the same time concern was being expressed once again about the ineffectiveness of the combined measurement and fingerprint system. The Home Secretary appointed a five-man committee under the chairmanship of Lord Belper with a remit to investigate the methods of identification of criminals by measurement and fingerprints (Charlton and Galloway 2007, Beavan 2002). The committee found that only 18,000 sets of anthropometric measurements had been registered since the recommendations of the Troop Committee and that some methods of identification in use before 1894 were still being employed. In

1900 Henry gave evidence to the Belper Committee, he gave an account of his system and a practical demonstration, using a fingerprint collection of about seven thousand persons. Shortly after Henry's presentation to the Belper Committee, his book entitled 'Classification and Uses of Fingerprint' (Henry 1900) was published. The Belper Committee finalised its report in December 1900 and recommended that the fingerprints of criminals should be taken and classified using the system that had been introduced into India.

In 1901 the Belper recommendations were implemented. In May the same year Henry was called back to London and appointed Assistant Commissioner of Police in charge of Criminal Identification at Scotland Yard. He selected three officers from the Anthropometric Office, Detective Inspector Charles H Steadman, Detective Sergeant Charles Stockley Collins and Detective Constable Frederick Hunt to form the Fingerprint Branch (Charlton and Galloway 2007, Beavan 2002). The officers soon mastered the new techniques and set about converting the fingerprints recorded on the anthropometric cards to the new classification system. Police forces from all over the world sent their officers for instruction in this new technique.

Detective Sergeant Collins' (see Figure 12) enthusiasm to extend the use of fingerprints led to him looking at the possibility of linking criminals not only to their past criminal records but also to their present and undetected crimes. He studied photography in order that when an opportunity arose he would be able to record fingerprints inadvertently left at a crime scene.

In 1902 a detective investigating the scene of a burglary at a house in Denmark Hill, London, noticed a number of dirty fingerprints on a newly painted windowsill at the point where the burglar had entered the house. The information was passed to the Fingerprint Branch whereupon Sergeant Collins visited the scene and examined the marks. The clearest mark he decided was a left thumb and on checking that the mark had not been left by a member of the household he photographed the mark and returned to Scotland Yard. Even though he knew that there were thousands of comparisons to be made he set about with the assistance of his colleagues to identify the mark. His efforts were rewarded when he identified a 41-year-old labourer named Henry Jackson who had several previous records. Fortunately, when Jackson was arrested he was found to be dealing in property from another burglary.

The identification was the easier part of the identification process, the hardest part was getting a court to accept the identification of a mark left inadvertently at the scene as the sole means of proving that Jackson had committed the crime. After some discussion between Henry and the Director of Public Prosecutions, Steadman and Collins it was arranged for the prosecution case to be heard by a well established Treasury counsel, Richard Muir. Muir spent many hours with Collins learning about the intricate details of this new investigative tool. Collins was called as a witness and explained in simple terms the basic principles of fingerprint identification, producing an enlargement of the left thumb mark on the window sill and the left thumb print of Jackson. He produced tracings of the ridges on both prints and indicated ten ridge characteristics, which were present in both prints and in the same coincident sequence. There was no attack by the defence on the fingerprint evidence, Jackson was subsequently found guilty and a precedent for fingerprint evidence was set.

The first UK case using fingerprints evidence in a murder was in 1905 (Charlton and Galloway 2007, Beavan 2002, Block 1970). Two brothers Alfred and Albert Stratton were charged with the murder of Thomas Farrow who was found dead in his paint shop in Deptford. During an examination of the premises a thumb mark was found on an empty cash box. Elimination fingerprints were taken from the deceased, staff and other persons who had access to the shop but the marks were not eliminated. A milkman and his assistant told police that they had seen two men leave the premises and another witness remembered two men running from the High street and recognised one of the men as Alfred Stratton but were unable to identify his companion. The Stratton brothers were arrested and were fingerprinted by Detective Inspector Collins, who later identified the thumbprint on the cash box tray as the right thumb of Alfred Stratton. The two brothers were eventually charged with willful murder. At their trial neither the milkman nor his assistant could identify the two brothers. The fingerprint identification became the vital evidence of the trial. Detective Inspector Collins described the method of identification by fingerprints and stated that he had never found two impressions taken from different people to agree in position and sequence. He produced a comparison chart (see Figure 13) to illustrate how the identification had been made and the twelve characteristics, which agreed in both prints.

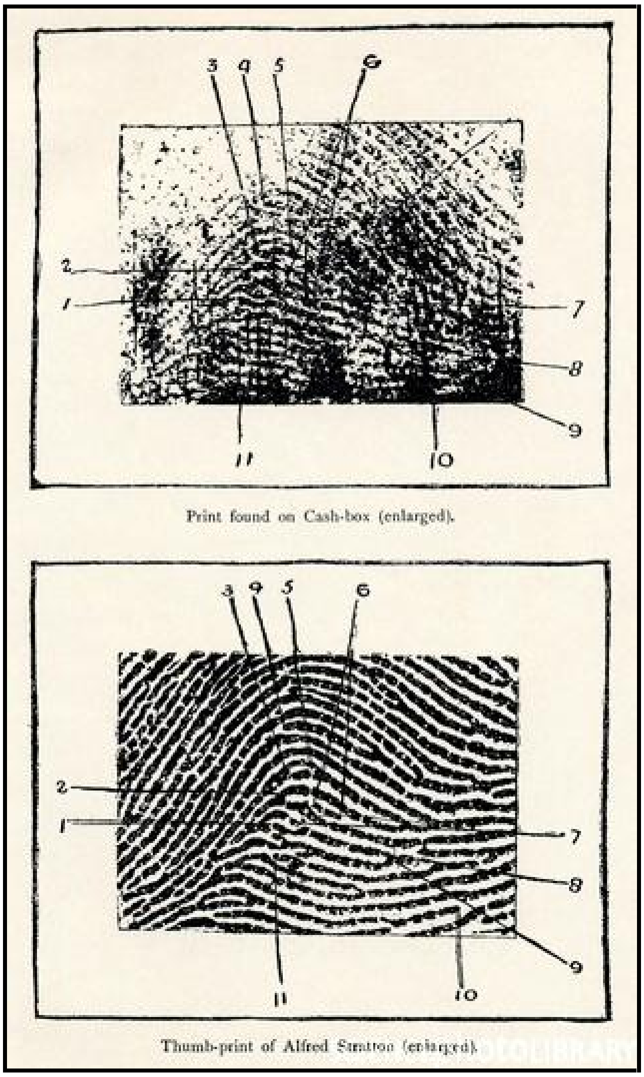


Figure 13 Marked up enlargement of comparison between crime scene mark and Alfred Stratton (Courtesy of the Fingerprint Society Archive)

It is difficult to establish what affect the thumbprint had on the jury but it is worthy of note that the jury accepted the mark as positive evidence of identification and after deliberating for two hours they returned a guilty verdict. Both brothers were sentenced to death by hanging.

In the early days of fingerprint identification there was no strict numerical standard for the number of characteristics in agreement required for identification; Detective Inspector Collins in the Stratton case presented his fingerprint identification with twelve areas of Galton detail in agreement. In 1920 the Fingerprint Department at Scotland Yard introduced the sixteen-point standard, but it was not until 1953 that it became a national standard (Charlton and Galloway 2007). Due to the varying standards used throughout the country fingerprint standards were re-examined at a meeting on 29th September 1953 between representatives from the Home Office, the Deputy Director of Public Prosecutions and five major police fingerprint Bureaux; Scotland Yard, Birmingham City Police Force, Lancashire

Constabulary, Manchester City Police force and West Riding Constabulary. It was agreed that it would be advantageous to have a national standard for officers giving fingerprint identification evidence in court and that the standard should be a minimum of sixteen points in agreement and that any other mark from the same scene with ten characteristics as a minimum could be mentioned. In 1983 the use of less than sixteen characteristics was accepted with the caveat that the evidence must be crucial and of such dire importance and that the evidence was given by an expert of long experience and standing in the profession.

In 1997 The National Fingerprint Evidence Standard Project Board was formed with a mandate from the Association of Chief Police Officers (ACPO) to replace the sixteen-point standard with a non-numerical standard. On 11th June 2001 the Police Forces of England and Wales adopted a non-numeric standard and fingerprint identification became firmly established on the principles of a methodology known as ACE-V, Analysis, Comparison, Evaluation and Verification of Friction Ridge Skin, using holistic analysis techniques (Charlton and Galloway 2007, Ashbaugh 1999).

Biometrics

Fingerprints have now been used as a means of identifying individuals within a law enforcement context for over 100 years. Today, automated systems are used to store, match and compare latent fingerprints found at crime scenes with those held on record (Charlton and Galloway 2007). Today, automated systems can achieve such accuracy in processing ten-prints that what is known as 'lights out processing' (where there is no human component in the identification process) is now viable. When systems search latent crime scene marks against a database of ten prints, it is now common place to see accuracy levels well above 75 to 80%. This by far exceeds what would have been possible in the days when fingerprint experts searched databases manually using an eyeglass and a thumb stool. Searches against the database that might have taken weeks and months in the past can now be carried out in minutes. Automated image retrieval systems have transformed the business processes for fingerprint examination and searching. But at this time, and for the foreseeable future automation will still require human intervention to both input data and to compare and verify results. AFIS systems (see Figure 14) deliver candidate lists from which identifications result, but any positive result must still be examined and subsequently verified by different fingerprint examiners.

Figure 14 Automated Fingerprint Matcher (personal collection)

An Automated Fingerprint Identification System (AFIS) is a storage, search and retrieval system for finger and palm print electronic images and demographic data. AFIS enhances the ability of the fingerprint examiner to search and identify crime scene evidence in an accurate and timely manner. AFIS systems have replaced outdated manual methods of fingerprint classification employed by law enforcement agencies over the past century. The Henry classification system is now largely redundant, and is not taught as a standard pre-requisite in many training establishments, something that many fingerprint experts around the world regret. Many examiners feel that the very process of learning the Henry classification enhances their knowledge and understanding of fingerprint patterns, friction ridge skin structure and elasticity.

AFIS systems utilise specialised software to create unique 'mathematical maps' (see Figure 15) or algorithms that are based upon the relationships between characteristics present within the finger or palm friction ridge skin structures (Pankanti et al 2001, Maltoni 2003).

Figure 15   Example of algorithm mapping techniques. (Handbook of Fingerprint Recognition 2003)

The use of such mathematical algorithms enables a fingerprint to be compared with millions of file prints in a matter of seconds.

Without doubt, fingerprints continue to play a huge role in the fight against crime. But fingerprints are becoming increasingly popular as a means of identification and verification in non-law enforcement domains also. Today, fingerprints are being considered as important biometrics identifiers in world-wide civil applications, such as National ID cards, passports, bankers cards, library cards as well as other less obvious applications, such as school access verification and even as a substitute for money in the school dinner line (Ashbourn 2000).

It is suggested that the manufacturers of fingerprint equipment have not solved the live detection problem (i.e. detect the difference between a live finger and a dummy). A Japanese cryptographer (Leyden 2002) demonstrated how fingerprint recognition devices can be fooled using a combination of low cunning, cheap kitchen supplies and a digital camera. Tsutomu Matsumoto used gelatine (as found in Gummi Bears and other sweets) and a plastic mould to create a fake finger, which he found fooled fingerprint detectors four times out of five (Leyden 2002).

Disaster Victim Identification

On December 26th 2004 a Tsunami devastated South East Asia. Many thousands were killed. The disaster had no respect for wealth, status or nationality. All peoples of the world succumbed to the terrible power of the ocean. Properties were taken away, hotels were levelled. Countries were ravaged.

Figure 16 The wall of remembrance in Phuket, Thailand (personal collection)

From Sweden to Sri Lanka, from the UK to Uganda, tourists (see Figure 16) and residents alike died at the hands of the killer wave (Charlton and Galloway 2007). Police agencies from around the world joined an international effort to use automated fingerprint recognition systems to try to identify the victims of the tsunami by comparing the fingerprints of the deceased with anti-mortem fingerprints from a victim’s property, house or identification documents.

Figure 17 Fingerprints being compared in Phuket, Thailand. Part of the victim identification effort (personal collection).

Biometric data was being used (see Figure 17), probably for the first time on this scale, in a civil application to identify victims, rather than to catch criminals, as is usually the case. It is clear that had many people not enrolled into biometric systems during their lives then many of the victims of the tsunami may still remain unidentified. DNA had a limited application in this particular disaster, since DNA degraded very quickly during the decomposition process. Many of the victims were found in an advanced state of decomposition (see Figure 18), so only fingerprints and dental records provided sufficient information to identify victims in many cases.

Figure 18 UK Forensic Specialists retrieving fingerprints from a tsunami victim in Phuket 2005 (personal collection).

This is the power of biometric science. The persistence of fingerprints, even after death, greatly aided the identification task set before the authorities in South East Asia.

Comparison Methodology - Principles of Friction Ridge Identification

A detailed examination of friction ridge skin under a magnifying glass will show that friction ridges are not uniform in configuration, nor do they necessarily flow in straight lines. Friction ridge skin has fractures and interruptions within the structure known as Galton details (Galton 1892, Stocker 2007, Henry 1900, Ashbaugh 1985, 1991, 1999, Charlton and Galloway 2007, Triplett 2006, Duncan 1942, Cooke 1932), also known as ridge characteristics or minutiae.

It is these characteristics (see Figure 19) that, when visible to the examiner, enable a determination of individualization of a latent crime scene finger mark against a suspect's ten-print card. There are two primary classifications of the characteristics to be found, namely, ridge endings (A), where ridges stop abruptly, and bifurcations, where a ridge divides into two (B).

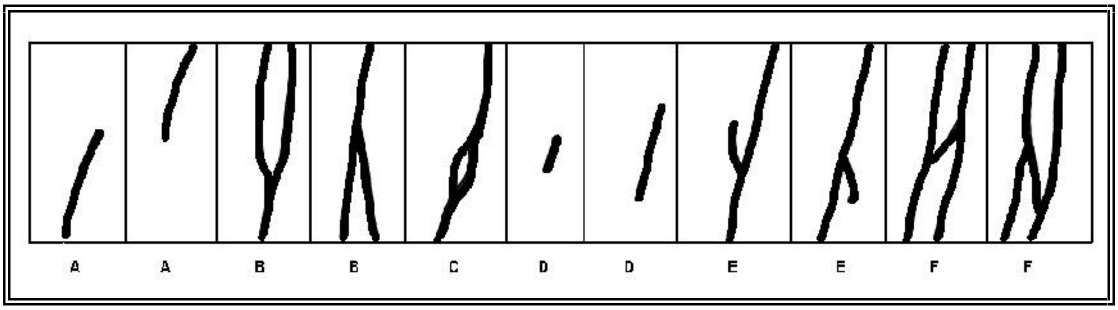


Figure 19 Types of Ridge Detail (Courtesy NPIA Training School)

In this instance the ridges running parallel to this feature diverge to accommodate the ridge division. There are variations on the theme for these primary characteristics. For example, a lake is where two bifurcations join together (C); an independent ridge is a short ridge that is divorced from any other ridge (D). A spur (E) is a combination of a small independent ridge and a bifurcation and a crossover (F) as the name suggests is a small ridge joined at each end to two parallel ridges.

1st Level **2nd Level** **3rd Level**

Figure 20 Different levels of ridge detail within a fingerprint (Courtesy NPIA Training School)

Latent print examiners should assess holistically all features (see Figure 20) within a fingerprint from which to make conclusions as to identity (Ashbaugh 1999). These features are broken down into three levels of detail. Level one refers to the basic pattern of the print, level two refers to the Galton details described earlier, and level three refers to the configuration of sweat pores and shapes of the ridge edges.

Figure 21 CSIs at work (Courtesy Sussex Police)

A wide range of optical, physical and chemical techniques are available for the detection and enhancement of latent finger marks (see Figure 21). The choice of the best sequence of techniques will depend on several factors that include:

- The nature of the surface and the presence of any particular contaminants
- Environmental factors
- The likely age of any evidential finger marks.

Figure 22 A SOCO examines a car using traditional brush and powder technique. (Courtesy Sussex Police)

Every feature and characteristic in a sample of friction ridge skin is readily recognizable and identifiable and identification can be made so long as there are sufficient characteristics of acceptable quality (see Figure 22). It is possible to match very small amounts of friction ridge skin from a crime scene against a suspect, so long as there is sufficient clarity upon which to rely upon the permanence and reproducibility of visible features.

When comparing detail between a crime scene mark and a ten-print exemplar, a generic methodology has been adopted known as ACE-V (Ashbaugh 1982, 1992, 1999). In order to make a value judgment on any identification, the examiner must first **assess** the friction ridge detail in the scene mark so that useful reference points can be identified in the next phase. Next, there must be a **comparison** of the mark features with that of features within the reference print. Then there must be an **evaluation** of the findings, from which conclusions as to identity must be derived. Finally, as in all scientific endeavours, there must be an element of peer review, known as **verification,** to repeat the ACE process independently to hopefully reach the same conclusions as that of a colleague.

During an assessment, the fingerprint examiner will look for:

- Distortion
- Development methods and medium used to visualize latent material (chemicals and powders etc)
- Deposition pressure to highlight areas of possible distortion
- Anatomical attributes (features)
- Clarity

During the Comparison, the examiner will look for similarities in:

- Pattern
- Ridge Path
- Ridge Shape
- Pore positioning

During the Evaluation, the examiner will look to form an opinion:

- Can the mark be eliminated?
- Is there sufficient information available to individualize?

During Verification, an independent assessment of the casework is undertaken to see whether the ACE process has been correctly carried out and that the conclusions reached are consistent with the original findings. In the UK, for a crime scene mark to be identified, the comparison must be undertaken three times. One check will be made by the original examiner, then verification by two further experts who must come to the same conclusions independently.

Scientific Validity of Fingerprint Examination and Human Error

Many observers have become increasingly uneasy with the apparent ease with which both the judiciary and the wider general public accept forensic evidence without detailed cross examination and interrogation (Charlton 2006). Some years ago I prepared a case for a murder trial that involved the identification; compilation and presentation of hundreds of latent finger marks from myriad different but associated crime scenes. Expecting many hours of court questioning by defence barristers I prepared tirelessly for many weeks to learn my case, fully familiarize myself with the evidence and generally prepare to answer any question put before me as the expert witness in the case. I was at the same instance relieved, and yet also disturbed, to spend less than four minutes under oath in the witness box. Both prosecution and defence, without testing the evidence in the court, accepted my evidence in chief. This troubled me. As an expert witness the duty to the court is to inform and to educate so that jurors can make informed decisions as to guilt or innocence. It can be hypothesised that the general demeanour of the expert witness, suit and tie, qualifications, the employer (the Police) and the very nature of the evidence was actually influencing the course of justice to a point where fingerprint evidence was accepted without question. Was the evidence really important at all? Or was it simply the status of certain individuals in the witness box that influences juries? Sir Bernard Spilsbury, the first and only Honorary Pathologist to the Home Office, gave crucial evidence in numerous murder cases between 1910 before his bizarre suicide in 1947. He made a major contribution to modern forensic pathology, explaining complex pathological findings in ways that a jury could understand. However, Spilsbury was a greatly revered forensic witness who was generally believed by juries even when he was wrong which may have led to the execution of innocent men and women (Rose 2007).

In light of the ease with which forensic experts can avoid cross examination and review by the judiciary, how far should we trust the abilities of fingerprint examiners to make correct judgments when comparing fingerprints if not tested or challenged? Fingerprints are obviously a useful tool for catching criminals, but the science suffers from a one major flaw, nobody really knows how often examiners make an erroneous decision (Bamber 2005). It is the examiners (and not necessarily the underlying biological science behind fingerprint identification) that are the potential weak link in the evidential chain (Schneier 2003). It is the examiner who is the instrument in determining a conclusion. It is the human examiner who must provide the decision.

Human error must exist in latent print examination as the fingerprint examiners are of course human beings, so intuitively it is illogical to assert that they are immune from making mistakes. For example, Steven Cowans was in prison for six years in the US before being released in 2004 after erroneous fingerprint evidence was overturned by conflicting DNA evidence (Coghlan 2005). This is not an isolated incident. Over many years numerous erroneous fingerprint matches have been made.
There is a similarity to be found in the methodology for comparing fingerprints to hypothesis testing in statistics, and the two types of errors which can arise in both domains. Both share four common elements.

The alternative hypothesis - This is the reason a criminal is arrested. This might result as a consequence of a positive fingerprint match. In statistics the alternative hypothesis is the hypothesis the researchers wish to evaluate (Lubin 1961).

The null hypothesis - In the fingerprint domain this is the presumption of a non-match. In both the judicial system and statistics the null hypothesis indicates that the suspect or treatment didn't do anything. In other words, nothing out of the ordinary happened. The null is the logical opposite of the alternative.

A standard of judgment - In the fingerprint domain, as in statistics there is no possibility of absolute proof. Given this fact, the methodology for fingerprint comparison is counterintuitive in that it permits a categorical decision (it is a match or it isn't). In the justice system overall, the standard is "a reasonable doubt". The null hypothesis has to be rejected beyond a reasonable doubt. In statistics the standard is the maximum acceptable probability that the effect is due to random variability in the data rather than the potential cause being investigated. This standard is often called the alpha level. It can be speculated that a fingerprint comparison methodology that does not have data to confirm error rate occurance is problematic as there is no account of the variability with either the evidence or the human. Thus, there is a debate to be had as to whether any fingerprint conclusion can be categorical 'beyond reasonable doubt' without significantly more data on error and similarity of evidence types. Perhaps, it can be argued, a statistical basis for fingerprint evidence might be worth investigating in more detail.

A data sample - This is the information evaluated in order to reach a conclusion (the fingerprint). Data is usually in numerical form for statistical analysis while it may be in a wide diversity of forms (eye-witness, fiber analysis, fingerprints, DNA analysis, etc) for the justice system. However in both cases there are standards for how the data must be collected and for what is admissible. Both statistical analysis and the justice system operate on samples of data or in other words partial information.

One good piece of evidence can negate a hypothesis (a Galton detail in a latent fingerprint that is absent in the exemplar being compared) but an endless amount of similar evidence to prove it correct. If the null is rejected then logically the alternative hypothesis is accepted. This is why both the justice system and statistics concentrate on disproving or rejecting the null hypothesis rather than proving the alternative. It's much easier to do. If a jury rejects the presumption of innocence, the defendant is pronounced guilty.

Type I errors: Unfortunately, neither the legal system nor statistical testing are perfect. A fingerprint examiner sometimes makes an error and an innocent person goes to jail. Statisticians call this a type I error.

Type II errors: Sometimes, the fingerprint examiner fails to correctly recognise a match between crime scene mark and exemplar. Statisticians give this error the name, type II error.

The fingerprint domain finds type II errors disturbing but not as horrifying as type I errors. A type I error means that not only has an innocent person been sent to jail but the truly guilty person has gone free. In a sense, a type I error in a trial is twice as bad as a type II error. Fingerprint laboratories puts a lot of emphasis on avoiding type I errors. Such errors can result in tarnishing the agency's reputation.

A way to minimise both the type I and type II errors is to increase the reliability of the data measurements. For example the Innocence Project has proposed reforms on how lineups are performed. These include blind administration, meaning that the police officer administering the lineup does not know who the suspect is. That way the officer cannot inadvertently give hints resulting in misidentification. Such blind techniques would be valuable within the fingerprint domain, and certainly the verification process would be an obvious part of the ACE-V process where such a process of blinding known results would make the evidence verification more relevant and powerful.
The value of unbiased forensic investigators with state of the art equipment should be obvious. There is no possibility of having a type I error if the examiner never makes a decision. But this is not desirable, and what is needed is rational and controlled procedure that aims to provide the most reliable evidence, in the most efficient and risk free environment possible.

The following examples relate to what is termed a false positive judgment (erroneous identification). However, there is also what is known as a false negative decision (a missed identification). While the fingerprint profession may prefer to weight the relative importance of such errors, preferring to insist that to have an 'erroneous Identification' is the most problematic mistake because a person has been wrongly sent into custody, one cannot ignore the ‘missed identification’ errors, as they represent the majority of all known errors in the profession (this is based on the author’s own experience in the workplace) and as such are just as relevant because they relate to a failure in the human actor to apply methodology and skill to arrive at a correct conclusion. Whether the mistake is to erroneously identify a perpetrator of a crime or to erroneously exclude another person, is potentially just as serious and no less a mistake. To wrongly exclude a perpetrator may leave him or her free to commit more crime, perhaps even to kill. One problem with knowing the true scale of the rate of error arising from a missed identification is that by the very nature of the error, a ‘miss’ is only known about when circumstances permit that error being found. For example a cold AFIS match as a result of a search of a crime scene mark may highlight that the AFIS match provided a potential perpetrator of the crime who had already been assessed against the crime scene mark as a result of being a nominated suspect in the case. This is the most common way such errors are brought to notice. In so far as the erroneous identification has more immediate and public ramifications potentially, has much to do with the rationale for the profession regarding errors in positive identification more serious than the missed identification. But in reality both are technical mistakes arising from the same methodology.

In the UK Neville Lee was arrested in 1991 in Nottinghamshire, for the rape of an eleven-year-old girl on the basis of a supposed fingerprint match. It is not known how many corresponding ridge characteristics were identified, but at that time a minimum requirement of sixteen matching ridge characteristics was in force in the England and Wales region of the Association of Chief Police Officers (ACPO). Lee's home was wrecked by vigilantes, and he was assaulted in jail. Another individual subsequently confessed to the crime, and Lee was released. The authorities admitted that the fingerprint match was erroneous. (Cole 2005)

Andrew Chiory was charged in London in 1996 for the burglary of the home of Miriam Stoppard, a writer and broadcaster who also happened to be the ex-wife of the well-known playwright Tom Stoppard (Nuffield Council on Bioethics 2007). Two separate latent prints from the crime scene were attributed to Chiory. Both matches were "allegedly triple-checked," and both were conducted under the requirement for sixteen corresponding ridge characteristics in force at that time.

'We tend to be blinded by science. It is time that fingerprint techniques were subjected to questioning'

Yard in fingerprint blunder

by Stephen Grey, Home Affairs Correspondent

Figure 23 'We tend to be blinded by science'…suggests this news headline about this case (Personal Collection).

Chiory served two months in prison before the match was exposed as erroneous. Despite an extensive external investigation of this miscarriage of justice (see Figure 23), no explanation for the misidentification has ever been made public (Cole 2005). It is known by the author of this thesis that the request to review this evidence was rejected at least once on the basis that 'we are Scotland Yard; we do not make mistakes'. It is this kind of corporate behaviour that is of particular interest to psychologists and has much to do with the relationship between employer, employee and the culture of the organisation. Joel Baken suggests (Bakan 2004) that corporations engage in a pathological pursuit of power. An eminent law professor and legal theorist, Bakan contends that the corporation (a law enforcement agency perhaps) is created by law to function much like a psychopathic personality whose destructive behavior, if left unchecked, leads to scandal and ruin.

It is suggested by Bakan that the corporation's legally defined mandate is to pursue relentlessly and without exception its own self-interest, regardless of the harmful consequences. To this extent it might be assumed that an organisation such as Scotland Yard may exhibit such tendancies as described earler to preserve its reputation and standing unless regulated effectively.

The McKie case in Scotland is perhaps one of the most notorious of all the recent controversies in latent print examination (Grieve 2000). Shirley McKie was arrested for perjury for stating under oath during a murder trial that a thumb print which was matched to her during routine elimination checks was in fact not hers (Zeelenburg 2008, McKie 2007). McKie was vindicated after latent print examiners from other agencies around the world challenged the validity of the identification made by the Scottish Criminal Records Office (SCRO). To this day there are experts in latent fingerprint analysis who still disagree as to whether it is a correct match or not. This inability to reach a consensus raises a number of interesting issues. As Cole (2005) suggests, is it possible to say with any certainty that Shirley McKie did not make the print in question given the need for, and lack of any consensus over the conclusions reached by opposing latent print examiners?

The erroneous fingerprint identification of Brandon Mayfield by FBI latent print examiners in response to the forensic investigation of the 2004 Madrid Bombings was another highly controversial case (Stacey 2004, Frieden et al 2006, and Thompson 2005). Mayfield’s fingerprints were alleged to have been identified against those found on a bag of detonators found in Spain after the bombings took place. However, subsequent re-analysis carried out after the Spanish authorities questioned the accuracy of the identification. The FBI fingerprint experts had to concede they had been incorrect in their original analysis.

Robert Loomis was convicted in 1920 for the murder of Bertha Myers during a burglary in 1918 in Easton, Pennsylvania. Two latent print experts gave evidence that a latent print found on a jewellery box was that of Loomis. Loomis successfully appealed on the basis of faulty jury instructions. At Loomis's second trial, the prosecution had to admit that Loomis was not the source of the latent print (Cole 2005).

A latent print found on a calling card at the scene of the Hall-Mills murders in New Brunswick, New Jersey in 1926 was attributed to William Stevens by three latent print examiners. Two latent print examiners later testified for the defense and claimed the attribution was erroneous (Cole 2005).

John "The Bug" Stoppelli was convicted in 1948 for the sale of drugs. After a raid, in which four other suspects were arrested, a latent print was recovered from an envelope containing heroin. The print did not match any of the four arrested suspects. However, after an extensive database search, agent W. Harold "Bucky" Greene attributed the latent to Stoppelli, a parolee in New York City. Greene found fourteen matching ridge characteristics. No other evidence linked Stoppelli to the crime. Stoppelli was convicted. Eventually, his attorney, Jake Ehrlich, convinced the arresting officer, Colonel White, to talk to Stoppelli. White became convinced of Stoppelli's innocence and had the print reviewed by the FBI laboratory. The FBI was able to exclude Stoppelli as the source of the print, and President Truman gave a full pardon. He had by this time already served two years in prison (Cole 2005).

Roger Caldwell was convicted of the murder of Elisabeth Congdon in Minnesota in 1978. Three latent print examiners attributed a latent print found on an envelope to Caldwell. The envelope was addressed to Caldwell and contained a gold coin believed to have been stolen from the victim's home. All three examiners were IAI-Certified latent print examiners. Agent Sedlacek testified that the latent print partial was identical with the inked impression on the fingerprint card bearing the name Roger Caldwell. This conclusion was based on eleven matching ridge characteristics and no unexplainable dissimilarities. The original negative of the latent print was reexamined for the trial of Caldwell's wife and supposed co-conspirator, Marjorie Caldwell. The forensic scientist Herbert MacDonell and the latent print examiners George Bonebrake and Walter Rhodes testified that Roger Caldwell could not have been the source of the latent print. (Cole 2005)

These cases, as well as other errors (Berry et al 1978, Berry 1977, 1991) and controversies (more than the fingerprint examiner community has publicised to date), have resulted in fingerprint analysis coming under attack from both the judiciary and academia. Some (Saks et al 2005) have questioned the very underlying scientific assumptions made by fingerprint experts. The rationale for the fingerprint profession remaining insular and secretive when considering erroneous identifications has much to do with the need of the profession to preserve the aura of infallibility around fingerprints analysis. The profession has built up over many decades a principle of human infallibility around the methodology used to compare fingerprints that has delivered into the public domain a mantra that training and peer review through verification of findings are enough to ensure that erroneous identifications will not only be extremely rare, but that such occurances will rarely if ever reach the courts. The forensic domain fears that any overt admission that the error rate for the examination of fingerprints is higher than publicised will undermine the position of power currently enjoyed by fingerprints in the criminal justice system.

Fingerprint examiners and law enforcement maintain that fingerprint evidence has unrivalled power within the judicial system. Fingerprint examiners also hold to a rigidity of thought around the clarity and unambiguousness of their work. In comparing fingerprints there are limited choices an examiner can make. The only possible conclusions are it is a match, not a match or there is insufficient detail upon which to make a conclusive decision (Bain 1985). However, the Evett/Williams report suggests that when many fingerprint experts use the term exact science (where conclusions are binary, it either is a match or it isn't) this is a contradiction in terms, science is not exact (Evett et al 1995). For example DNA analysis in the forensic domain, while founded on specific biological data researched over many decades, is reported in the courtroom in terms of a statistical probability since there is no way of empirically proving the absolute uniqueness (except in the case of identical twins) of DNA in the global population. The same is true of the uniqueness of fingerprints in that without assessing every set of human prints in the world either alive now, or have ever lived, there can be no way of truly knowing whether fingerprints can ever be unique to the degree of having a categorical belief that two prints match one another. So perhaps a statistical presentation of fingerprint evidence is also both valid and necessary?

Perhaps the apparent certainty and confidence displayed by fingerprint examiners is misplaced? Ellad identified that an overconfidence bias was present in interrogators. Interrogators believed they were much more accurate in their decisions than they actually were. This overconfidence in one's decisions is a common characteristic among fingerprint experts and appears in many other domains of expertise (Ellad 2003).

An example of the confidence expressed by practitioners of latent fingerprint examination can be found in the literature. Some suggest fingerprints ‘point the finger of guilt upon the suspect’ (Leadbetter 1984).

But is this what fingerprint evidence should really support? Rather, it can be argued, the evidence should be there to provide objective data that the investigator and ultimately the courts can use to arrive at informed conclusions about a case. Fingerprint examiners should be the agents for the truth, not the deciders of innocence over guilt. Huber states science should never be used to make a case, but rather it should be for the courts to use more wisely and profitably (Huber 1959).

There is some awareness of the dangers of dogma and arrogance in the science of fingerprints and a recognized desire by some practitioners to move away from tradition and dogma and to accept that errors are made on occasions. Graham Hughes (former editor of Fingerprint Whorld) stated that it is not heresy to abandon dogma that we have been used to, or to consider how errors may ultimately lead to identification sciences losing some integrity (Hughes 1998). What Hughes was implying was that the profession should not be afraid to recognise potential weaknesses and to improve upon current practice to the benefit of the science, even if that process is difficult and potentially creates short term embarrassment. Moreover, such self critique if it leads to changes in policy, training or working practices that leads to an increase in reliability and accuracy would be a positive step by the forensic community, not a negative action. It has been stated that training and procedures in fingerprint analysis give cause for concern (McCartney 2006). So such self critique by fingerprint examiners as well as academia within the domain seems both prudent and timely.

Fingerprint examination would be a unique science indeed if the human error rate were zero, which based on the evidence it is not. One only has to look at other forensic sciences where there are recorded human error rates. Human error rate data for other forensic sciences include 63% for voice identification, 40% for handwriting analysis, 64% for bite mark analysis, and 12% for hair analysis (Begley 2005). Given that these forensic disciplines have documented error rates does not invalidate the usefulness of the evidence in the court room. It should be remembered at this juncture that all evidence presented in the courts is provided to facilitate the jury in reaching a conclusion as to guilt or innocence. It is not the evidence or the expert presenting the evidence that necessarily provides absolute proof, rather, it is a culmination of manjy strands of evidence that adds weight either to the prosecution or defence arguments. To this end the fact that there is a known error rate for such forensic evidence helps the gatekeeper of such evidence, the Judge, to weigh up either its admissibility or relevance to a particular case. Fingerprint evidence should not hold itself up as an infallible science as this is to potentially overstate the power of the evidence and could, on occasions, misguide the pursuit of accurate and fair assessment of the case evidence by the jury. The fear of acknowledging any error rate in fingerprint examination has much to do with a perceived loss of power and standing of the science in the criminal justice system, where in fact, ironically, such knowledge may actually enpower the science and its standing in the 21st century courtroom. Critics of fingerprint examination have been calling for similar data to be made available for fingerprint examination too. Simon Cole asserts misidentifications are entirely predictable outputs of a normally functioning system of biometric identification relying on biological markers as perceived by experientially skilled human operators (Cole 2008).

Examiners explore several different rhetorical strategies to appear to be demonstrating validity, while in fact demonstrating other things. These include "the fingerprint examiner's fallacy" (if the ACE-V methodology is followed error cannot happen) and ‘the casework fallacy’ (over 100 years of fingerprint examination only a handful of errors have been recorded). In essence, whenever an error is highlighted in the profession, it is invariably the case that the hierarchy blames the practitioner. It is the fault of the person doing the analysis and that the methodology is sound, even if human factors contributing to the error proved to be unavoidable and to a large extent unconscious to the individual. To date it has been very rare for the profession to investigate possible factors other than practitioner incompetence for the occurrence of error (Ratha et al 2004). The 'inability of fingerprint proponents to refute the charge that validity studies are lacking is further evidence that the charge is, in fact, correct' (Cole 2006).

There have been a few rare attempts to measure error in latent print analysis over recent years. One recent study suggests that matching accuracy is a function of source finger type and image similarity (Vokey et al 2008). Dave Grieve states after a technical survey of latent print examiners, '156 answer sheets produced 48 erroneous identifications in total with erroneous calls made on all the latent prints submitted as part of the study'(Grieve 1996). He goes on to state that 'to assign clerical error to all these errors is wishful thinking. In casework 1 in 5 participants would have provided incorrect damning evidence against suspects'.

Perhaps one of the reasons why validity studies are so rare in the fingerprint domain is that there is a real fear of undermining the long held core values and beliefs that underpin the dogma within the profession, and that to question the reliability of the method or to assess wider cognitive issues impacting upon practitioners is deemed both dangerous and by some, unnecessary. Some senior practitioners resort to a defensive posture when challenged about fingerprint science and technique and often it is the practitioner that is blamed for error by their senior management rather than the underlying systems and processes that staff are asked to work within (Leadbetter 2007). Although Leadbetter acknowledges that ***an examiner might be influenced in his/her decision making process***, he categorically states that he never knew of cases of examiners who were influenced by context or emotion. However, this is only a reflection of his knowledge, rather than reflecting what represents reality (Dror et al 2007).

It has been suggested by Tiller (Tiller 1986) that an expert is an expert as long as his/her work can withstand an audit of peers. Any expert that is proven wrong loses the support of the profession. Infallibility of the profession must override individual rights (Tiller 1986). It appears it is easier to label the human practitioner as incompetent rather than to understand why the error has occurred. Working within an environment where the need to preserve the profession at the expense of all else prevails creates an oppressive culture (Johnson 2006) where dogma and tradition overrules genuine scientific thinking. Canteen socialising in law enforcement domains provides a fertile breeding ground for colleagues to view the world according to the dominant subculture within the organisation and results in new recruits looking to others to help make sense of the complicated environment in which they find themselves, which it can be hypothesised, would lead to stereotyping and prejudice (Ainsworth 2002). This seems to reaffirm what Cole was saying in 2006 in that rather than accept critique and attempt to provide answers as true scientists should, many examiners revert to a defensive posture that is, on face value, less than scientific in either rationale or logic. A culture may pervade within the forensic domain that discourages dissent and actively shuns data that complicates the environment within which forensic practitioners exist. Practitioners for the most part are simply not aware of the dearth of scientific support for their discipline's core assumptions (Saks et al 2008).

That there are so few studies of error rates or examiner vulnerabilities ‘owes a lot to the precarious legal situation in which latent print examiners find themselves. They have nothing to gain and everything to lose from validation studies’ (Cole 2006).

There have been legal challenges for many years in the US to the acceptance of fingerprint science as a valid methodology. In the landmark motion to exclude fingerprint testimony, (State of Maryland v. Bryan Rose K06-0545), the defendant contended that ACE-V was not a valid methodology which has been subjected to scientific testing and that the error rate in latent print identifications is unknown. It was asserted that in the absence of an error rate, the reliability of the methodology remained unproven. A fundamental problem, according to the defendant, was that the subjective comparisons in ACE-V involve psychological phenomena known as "confirmation bias." Further, the defendant argued that the "standards" for latent fingerprint identification are inadequately defined. The lack of error in certain fields of inquiry often derives from the nonexistence of methodological research into the problem and merely denotes a less advanced stage of that profession (Risinger et al 2002).

There are examiners who are prepared to look at the issues of error rate and to investigate the methodology as well as human factors, and to suggest ways of mitigating potential weaknesses. Wertheim and Langenburg (Wertheim et al 2005) showed in their research results that errors can and do occur. A study by Schiffer and Champod (2006) suggested that context may have an impact on practitioners dependent upon where in the ACE-V process the fingerprint analyser found themselves. Latent print examinations are currently performed by human examiners, and human examiners are fallible, no matter in what domain they work. Therefore, the examinations are only as valid as the examiner performing the task and the conditions under which the tasks were conducted. Measuring and enhancing fingerprint expert performance is a challenging task. However, for any expert domain, be it medical experts, military fighter pilots, police officers, or forensic experts, one must study and understand the underlying expertise (Dror et al 1993). Results obtained in law enforcement generally demonstrate that investigating police officers' judgments (this may include latent print examiners) are malleable and susceptible to influence from extralegal factors (Ask et al 2007, Ainsworth 2005).

Enthusiasm to seek answers to these questions is gaining in strength (National Academy of Sciences 2008). The fingerprint community must be willing to subject established beliefs to verification and peer review (Busey et al 2008).

Cognitive Issues in the Fingerprint Domain

While considering human error and cognitive vulnerabilities in the science of fingerprint examination, it would be useful to first undertake a brief review of science in general, and how cognition and bias generally have historically played a part in distorting scientific endeavour.

Perhaps the first recorded instance of a scientist recognizing that the cognitive traits of an observer were influencing the accuracy of particular observations occurred more than 200 years ago (Risinger et al 2002). In 1795, Nevil Maskelyne, Astronomer Royal at the Greenwich Observatory, realized that he and his assistant were obtaining different results for the times of stellar transits, even though they were using identical methods. These discrepancies reflected differences in judgments.

In the 1820s, Bessel, an astronomer in Konigsberg, studied the problem and found that such differences were not only common, but in astronomical measurements they reflected predictable individual tendencies. By the 1830s astronomers had developed a method for calculating "personal equations" that enabled them to measure these particular kinds of observer error, adjust for them, and remove the distorting effects from their findings.

Sir Isaac Newton failed to report absorption lines in the prismatic solar spectrum (Risinger et al 2002), though they would have been clearly visible with the apparatus he was using. The most likely explanation for his failure to see them is that he held theoretically based expectations that such phenomena should not exist. Because he believed they did not exist, he failed to see them, or at least to note their presence.

While Newton failed to see something that did exist, scientists of the early twentieth century saw something that did not exist. First reported by Rene Blondlot in 1903, ‘N-rays’ appeared to make reflected light more intense (Risinger et al 2002). So long as they were believed to exist, the effects of N-rays were ‘observed’ by many scientists. Once it was determined that N-rays did not exist, their effects ceased to be observed.

The FBI synopsis report on the Mayfield (Office of the Inspector General 2006) case (where there was an erroneous fingerprint identification) was significant in that it adopted the term "confirmation bias" to explain the cascading of the misattribution through the series of checks (verification, second verification, and independent review) designed to detect errors. This represented an implicit concession to the argument put forward around two years earlier (Risinger et al 2002) that forensic analysis is prone to confirmation bias, if for no other reason than that psychological research has found that nearly all, if not all, human observational measurements are prone to such bias (Cole 2006). Information guides people's search and evaluation of data, leading them astray (Dror 2008). Once people have a belief of what the data may suggest in terms of a theory or hypothesis, this has a powerful effect on how they perceive it, the way they process the information and the mental representations they form of this data (Dror 2008).

These effects can take many different forms and influence people in a variety of ways. For example, confirmation bias is when people notice and give more weight to information that is consistent with and supports certain interpretations and not others (Dror 2008, Nickerson 1998, Klayman 1995, Evans 1990, Kosnik 2005). Conversely, people do not notice, dismiss, or give less weight to other information that does not fit (or even contradicts) the interpretations they unconsciously support. Confirmation bias is only one example of the way people think that can diminish objectivity. Escalation of commitment and momentum, conformity, need for closure, prophecies that fulfill themselves are just a few other psychological and cognitive phenomena where experts unavoidably and unconsciously lose objectivity and can be selective and biased (Dror 2008).

It is surprising to note in the literature that the legal profession, the forensic community and the fingerprint domain in particular, have been aware of the influence of the mind in the analysis of fingerprints for many years, but have apparently done little to either investigate or to find solutions to such issues (Ashbaugh 1992). One UK examiner stated that fingerprint examination is a mental process; effective training and practice will improve innate abilities of observation, memory and recognition (Fairhurst 2003, 2005).

In more general policing domains racial stereotyping has been recognized to stimulate visual perceptions that either is inaccurate or absent altogether (Payne 2006, Payne et al 2004). Payne states that split-second decisions magnify the bias by limiting people's ability to control responses and that the bias requires no intentional racial animus, occurring even for those who are actively trying to avoid it. It is entirely conceivable that the shooting of Jean Charles de Menezes in London during a period of raised terrorist tension was in part, at least, due to such a stereotyping phenomenon. It has also been recognized that there may be a lack of protection to innocent people who confess due to psychologically coercive police interview techniques (Kassin 1997).

In the domain of fingerprints, some examiners have noted the need for caution in the relation of decision making when under the influence of investigative pressures. As Hughes states, pressure seeps through mental armory and can lead to mistakes (Hughes 1994). Such caution in guarding against inaccurate decision making is well founded. Wherever decisions are made by humans factors can conspire to influence and impact upon decision outcomes. As Dror states, we are different people, with different experiences, different views, and different brains and sensory mechanisms. This entails that we have different perceptions. This individualization of perception derives from the active nature of cognition and the wide range of factors that affect what and how we perceive (Dror 2005).

It has been shown that factors such as age can also have an impact on visual perception. It was found by Madden (Madden 2008) that older adults are often slower and less accurate than are younger adults in performing visual-search tasks, suggesting an age-related decline in attentional functioning. However, Madden disclosed that age-related decline in attention is not entirely pervasive and that visual search that is based on the observer's expectations (i.e., top-down attention) is relatively preserved as a function of adult age. This may represent an age-related increase in the role of top-down processing during visual tasks. The collective experience (and age) of the examiners in the Mayfield case was indicative of very experienced people in their profession. It would seem from Madden that these experienced officers could have lost some speed and accuracy in their visual search, but that also, their vulnerability to top down influences may have been consistent. This may have represented in this case a net increase in the risk of error through conformational and contextual biases.

It is not that the fingerprint examiners are incompetent, but it is that their perception and judgment as human beings is dependent on a variety of issues beyond the mere friction ridge detail present. For example some people will work at different levels of efficiency at different times of day and there is a performance curve in both morning and afternoon sessions either side of lunch which has a warm up, max, and decline phase (Mace 1932). As Dror states, the mind and the brain are dynamic systems that play active roles in how we perceive and construct realities (Dror 2005). The human mind is not a camera and we do not passively process information. Thus, to understand expert performance, and especially in a highly specialised domain such as human identification, one needs to examine the roles of the human mind and cognition (Dror, 2007).

In a study by New Scotland Yard (Hall et al 2008), it was reported that emotional context had an effect on experts' analyses and that they were affected by the information they were given. However, this study reported no effect on final opinions in high or low context conditions which has been heralded by some in the fingerprint profession as vindication for the sound methodology and science of fingerprint examination as perceived by the domain. The conclusions reached by Hall and Player could by observation be deemed inconclusive by stating that since the final decisions were accurate that context had no impact on decision making. This is to miss one fundamental point. The examiners themselves reported being influenced by the context. So by definition context played a part in the conclusions reached. The fact that participants made the correct decision and achieved the appropriate outcome may reflect a weakness in the experimental methodology of only using one point of stimuli. It could just as easily be the case that while the examiners reached the correct conclusion, they reached it for the wrong reasons, or, the context used was only able to shift decision thresholds to a certain point along a continuum, but not to a 'tipping' point of altering final conclusions. The most important conclusion of the paper was not that context effects did not impact upon accuracy of final decisions, but rather that they reported finding context effects that impacted upon decision processes, even if the outcome was not changed ultimately. Such studies have serious implications for the world of forensic science as Saks predicted in 2003 (Saks et al 2003).

Glenn Langenburg et al (2008) conducted experiments to assess if fingerprint specialists could be influenced by extraneous contextual information during a verification process. Prior to the experiment, participants were separated into three groups: a control group (no contextual information was given), a low bias group (minimal contextual information was given in the form of a latent print examination report prompting conclusions from an anonymous, but qualified, specialist), and a high bias group (an internationally recognized fingerprint expert provided conclusions and case information to deceive this group into believing that it was his case and conclusions). A similar experiment was later conducted with novices (laypersons with no experience in conducting fingerprint comparisons). The results, (as similarly showed by Hall et al 2008), demonstrated that fingerprint experts are influenced by contextual information during fingerprint comparisons, but not necessarily towards making errors in the final conclusions. Instead, fingerprint experts under the biasing conditions provided significantly fewer definitive and erroneous conclusions than the control group. They tended to provide opinions that were inconclusive. In contrast, the novice participants were more influenced by the bias conditions and did tend to make incorrect judgments, especially when prompted towards an incorrect response by the bias prompt. This was not the case with the fingerprint experts.

If one accepts subjectivity as fact in the fingerprint profession, the subjectivity is not whether one expert can make a correct decision over another but rather when one person will reach a point of decision in comparison with another (Charlton 2002). The complexity of comparisons may ultimately dictate accuracy as the threshold for decision making in such an instance becomes more important. In other words the ability for biasing factors to alter decision making thresholds may become more acute the more ambiguous a fingerprint comparison becomes. This hypothesis has potentially serious implications for how processes and procedures work in fingerprint bureau.

Courts in the US are already aware of bias associated with fingerprint analysis. However, the danger of confirmation bias affecting an examiner's subjective opinion is rarely discussed in the fingerprint examination methodological literature or in the court cases upholding admissibility of the technique. (Steele 2004).

Suggestions for addressing the issues of bias in forensic science have already been made. Koppl in 2005 suggested competitive self regulation would reduce group-serving bias. He advocated a system where each forensic laboratory would have to consider who else may be likely to examine the same evidence. A laboratory's reputation would suffer if it were found to have provided a false or sloppy analysis. Under competitive self regulation, every laboratory would have an incentive to give professional, scientific, and objective testing (Koppl 2005).

It has already been discussed that initial impressions and preconceptions can bias and can be detrimental to achieving high quality, evidence-based, decisions. Time pressure in particular, it seems, has a detrimental effect on the ability to ignore biasing factors (Kruglanski 1983). It is important to note that biasing effects are examples of honest mistakes brought about by cognitive nature which affects us all. Often, the motivation to help and solve a case, to provide justice, clouds our judgments and our ability to reach objective conclusions (Dror et al 2008).

People see what they want to see (Balcetis et al 2006). Within the domain of law enforcement where performance culture rewards positive outcomes and convictions, and where the public at large denigrate those who allow a felon to walk free from court, or worse, never get caught at all, then it is easy to see how those who work within the policing structure need to 'get a result'. As Sanders states it is nearly impossible to believe there is no bias in presentation, no shading of belief among a group of people who are vetted, hired, groomed, and rehearsed to present testimony often based primarily on materials provided by the party that employs them (Sanders 2007).

This thesis will probe in more detail the motivational influences on fingerprint examination as this will provide data to assess the interaction between certain motivations, performance and impact upon expertise in general.

Motivation can play a large part in decision making. Motivated decision makers can bias their judgments to support desired conclusions (Boiney et al 1997). Clarity and accuracy of object perception has been found to be related to an observer's 'interest' in perceiving them (Vernon 1962). Even whether someone is happy (Nettle 2005) or sad can influence decision making, a sad person is more likely to take risky decisions (Raghunathan et al 1999). Cognitive appraisal tendencies associated with sadness and anger has been shown to exert different influences on investigators' crime-related judgments (Ask et al 2006). Supporting evidence was found in an experiment with 61 experienced criminal investigators. First, when judging the reliability of a witness statement, ‘sad’ participants relied on their perception of both witness and situational variables, whereas ‘angry’ participants relied only on their perception of witness variables. Second, when making judgments of the case, ‘sad’ participants are sensitive to the consistency of a witness statement with the central hypothesis of the investigation, indicating substantive processing, whereas ‘angry’ participants were unaffected by statement-hypothesis consistency, indicating heuristic processing.

It can be hypothesized it is better to avoid making snap judgments in immediate connection to an anger-provoking event, such as the interrogation of an uncooperative suspect. In addition, it may be that investigators should consult the opinion of colleagues who are not personally involved in the particular case, since their judgments are less likely to be tainted by strong feelings. With further research, using forensic scientists and fingerprint experts in particular, the feasibility of these ideas and other motivational factors can be further established.

Motivational attributes which may impact on judgments and decision making by fingerprint examiners, include the motivation termed need for closure (Webster et al 1994). Need for closure is an expression that represents a desire for firm knowledge, a need to see things through to a conclusion (Kruglanski 1989, 2002, Mayseless et al 1987). It is hypothesised latent fingerprint examiners are motivated to make certain judgments and decisions. Those motivations may include need for closure and may determine how and when examiners use technology processes and methodology, such as ACE-V, and therefore it may be possible to predict the possible extent to which examiners are prone to biases. For example, if fingerprint examiners exhibit a need for closure, this would manifest itself through a tendency to either get frustrated should evidence not be accounted for and identified, but also they will exhibit tendencies in these circumstances to consider heightened risk taking and bias in their conclusions, leading to potential error. Just as importantly it will be important to understand when these biases may manifest themselves in certain contexts and situations (Tversky et al 1974, Hogarth 1981, Kruglanski et al 1983, Kruglanski 1989, Chaiken et al 1989, Pronin 2006).

Although fingerprint expertise has gone unchallenged over potential error rates for more than a century, the issue of susceptibility to human errors must be considered within the domain of fingerprint expertise. Most domains of expertise are susceptible to some kind of human error so it seems unlikely, almost unprecedented, that the domain of fingerprint expertise has a 100% accuracy rate. Jurors must know that 'certainty is not always desirable and sometimes hazardous to the innocent' (DiFonzo et al 2007).

The discussion within this thesis will highlight some interesting theoretical concepts. While some research has been conducted in the real world of policing (Ask et al 2005), it will be useful to place under scrutiny the domain of fingerprint examiners to look at examiner susceptibility to bias, motivational factors, interaction with technology as well as observational consistency that might offer some further understanding of the influences that may exist.

By understanding the cognitive attributes of examiners in expert domains it may be possible to mitigate against future error and methodological breakdown of fingerprint analysis, as well as design and implement effective and robust recruitment, selection and training (Brown 2005) environments that are able to provide best practice for examiners and to satisfy public confidence in not only fingerprint examination but also other forensic and expert domains.

Chapter 2

Emotional Experiences and Motivating Factors Associated with Fingerprint Analysis

Fingerprint analysis has been a cornerstone of forensic investigation for well over 100 years. Such is the traditional trust placed in fingerprint evidence that it is rarely questioned by the public or the wider judicial system. The normal reaction of the courts in deciding upon the admissibility of fingerprint evidence is often one of acceptance, based in a large part on the assertions of the practitioners within the domain. In the case of the People v Jennings in 1911, fingerprint evidence was accepted, based on the rationale that the fingerprinting system to that time in England, had been used in thousands of cases without error. It was stated that there was a scientific basis for the system of finger print identification, and that the courts were justified in admitting this class of evidence (People v Jennings 1911).

In the UK today the vast majority of cases involving latent print examiners' findings and conclusions are for the most part still unchallenged and accepted in the courts as unquestionable scientific fact, much as they were 100 years ago. Admissibility of fingerprint evidence is still largely based on the historical successes of fingerprints and experts in the domain still have an aura of experience and authority. Courts still assume that there is sound justification for accepting fingerprint evidence. Indeed, many people arrested and charged as a result of fingerprint evidence will often admit to crimes based solely on the knowledge a fingerprint match is confirmed.

The fact that fingerprint evidence holds such a strong position within the global judicial framework fuels a sense of heightened confidence on the part of practitioners in the fingerprint domain in not just their own abilities, but also confidence in the accuracy and provenance of the fingerprint evidence itself. Such confidence in the value of fingerprints and their reliability requires the collective protection on the part of the practitioners to ensure that the reputation the evidence enjoys in courts is maintained. This in itself provides ample motivation for those in the domain to act in ways that protect fingerprint science and maintain the privileged position which it enjoys.

Recently the judiciary has begun to assess forensic evidence in a more critical light. In a keynote speech made to the Forensic Science Society in 2010, the Rt Hon Lord Justice Leveson made clear and unambiguous comment on the state and reliability of fingerprint evidence. He stated that first, expert opinion is just that, an opinion. Second, experts need to know the limits of their expertise and have the integrity to inform the Court of those limits. Third, expert evidence should, indeed must, be submitted to robust testing, either by another witness in the same field or in relation to accepted scientific methodology.

Lord Levenson indicated that in order to accept fingerprint evidence within the 21st century legal framework, then it was necessary that the science of fingerprint analysis should be backed up by empirical research and understanding of the human element, that is, the human actor and their vulnerabilities to bias and other human factors.

If the comments made by Lord Levenson are compared with the judicial comments made in the People v Jennings in 1911, it would appear that over the course of a century little research has been done to either validate the science of fingerprint analysis or to assess it's reliability, and only now are questions being asked of fingerprint examination and its place as forensic evidence in the court room.

The acceptance of fingerprint evidence is the culmination of 100 years of reverence on the part of the public and judiciary alike of the fingerprint profession that is based at best on the assumptions of applied science and educated guess work, and at worst, is representative of the ultimate self-fulfilling prophecy. For example, the profession argues that fingerprint examination is reliable because the error rate is very low. It can be argued the error rate is very low because there appears to be little motivation amongst practitioners to assess the true basis of this assertion. The judiciary, without reference to sound academic scientific data to back up the notion of reliability around fingerprint examination, must at this time rely upon the 'best guess' that fingerprint examination is reliable, safe and trustworthy. A best guess provided by the fingerprint profession itself.

Not only is the historical rationale for the safety of fingerprint examination reliability questionable, but now such assertions of reliability by the profession are further clouded by the advent of digital technology associated with the scanning of fingerprints of detainees in custody. Many detainees, who would have traditionally challenged the identification methods used in the hope of convincing their custodians that they had the wrong person, now no longer challenge the digital processes deployed to identify them.

Much of the validity of fingerprint science and the trust placed in the evidence over many decades are based on the biological uniqueness of friction ridge skin and the methodology of fingerprint identification that is considered to produce correct matching with zero error rates (Cole 2005). While the fingerprint profession asserts errors are extremely rare, there are notable instances where such errors do occur, often with huge implications. Errors in the analysis of fingerprint evidence in high profile cases around the world including the erroneous identifications of Brandon Mayfield and Shirley McKie (Thompson et al 2005, Zeelenburg 2008) have resulted in legal counsel, media and public attention focusing on the core issue of whether fingerprint examination is to be considered one of the most reliable and valid forensic sciences (Saks et al 2005), or a pseudo science, based on ill thought out dogma.

The erroneous fingerprint identification of Brandon Mayfield by FBI latent print examiners as part of the forensic investigation of the 2004 Madrid Bombings was one such contentious case (Frieden et al 2006). Mayfield's fingerprints were alleged to have been identified against those found on a bag of detonators found in Spain after the bombings took place. However, after subsequent re-analysis carried out after the Spanish authorities questioned the accuracy of the identification, the FBI fingerprint experts conceded they had been incorrect in their original analysis. The cognitive and motivational factors may have conspired to provide the fertile ground for such an error to occur.

The McKie case is perhaps one of the most notorious of all the recent controversies in latent print examination. Shirley McKie was arrested for perjury for stating under oath during a murder trial that a thumb print which was matched to her was not hers (Zeelenburg 2008). McKie was only vindicated after latent print examiners from other agencies around the world challenged the validity of the identification. In contrast to the Mayfield case where the error has been acknowledged, to this day there are still experts in latent fingerprint analysis who still disagree whether it is, or is not the thumb mark of McKie. This inability to reach an agreed consensus raises a number of issues. For example, is it possible to be certain that Shirley McKie didn't make the print in question given the need for, and lack of any consensus over the conclusions reached by opposing opinions of latent print examiners? (Cole 2005).

The McKie case in particular raises questions around the core motivations of fingerprint examiners and what it is that drives them to make decisions, and why examiners maintain those decision outcomes (even if erroneous) in the most stressful and controversial of circumstances. For example, albeit there is no ground truth for knowing whether Shirley McKie really did leave her fingerprints at the crime scene, intuitively, the print in question is either hers, or isn't hers. Logically, the two polarised groups who either confirm or deny the print belongs to Shirley McKie are in a position where one of these groups of expert examiners must be wrong in their conclusions. The lack of consensus has its roots in not just cognitive biases, including context (Dror et al 2005 and 2006), where the case assumes such a high profile that the ability to assess the evidence objectively is hindered by the top down context of the case, but will also include other phenomenon such as escalation of commitment (Staw 2004), where the ability to step back from a decision becomes harder, the more entrenched you become in a particular direction of thought. In addition, hindsight bias (Blank et al 2007) can manifest itself, where there is the inclination to see events that have already occurred as being more predictable than they were before they took place. This can be seen in the way in which experts around the world were so keen to offer their support to those who asserted the McKie mark to be erroneous by suggesting it should never, and could never be a correct identification. However, such a ‘knew it all along’ way of thinking in relation to the McKie case should not be surprising as all humans look back on events in such a way as to bolster their own understanding and validate prior decisions.

Aligned with cognitive factors, there are also the cultural and professional motivations for supporting one side or the other where no real technical consensus can be reached. Such motivations might include a fear of being made to look foolish, a fear of admitting a technical failing, leading to a loss of reputation and standing in the forensic community. Such motivational factors are important to understand as they will inevitably contribute to the overall decision making process.

Fingerprint identification involves visual search and processing of information. Scientific research in other domains shows that the interpretation and selection of visual information can be greatly affected by emotional state. Examples of this include biases of cognition resulting in preferential processing of visually threatening stimuli (Eysenck 1997), processing of facial expressions corresponding to the emotional state of the perceiver (Niedenthal et al 2000), and even ambiguous sounds can be processed and interpreted in a way that correlates to the person’s emotional state (Pincus et al 1996). These top-down processing effects enable context and background knowledge to influence the selection and processing of information (Ashbaugh 1999, Eysenck 1997).

Research described in more detail later in this thesis examined whether fingerprint examiners can objectively focus solely on feature information in fingerprints without being misled by extraneous information such as context (Dror et al 2005). Other research has also demonstrated that fingerprint experts are vulnerable to biasing information even when they are presented with relatively subtle and routine day-to-day contexts, such as corroborative (or conflicting) evidence of confession to the crime. The results obtained consistently highlight that context influences the judgment of the experts. Thus, it appears that contextual information does not need to be extreme and unique to influence experts in their fingerprint examination and judgment (Dror et al 2006).

Contextual information is only one of many cognitive influences that may affect fingerprint expert performance. Other influences may also include need for cognitive closure, emotional rewards, belief perseverance, escalation of commitment, conformity, motivated perception, self-fulfilling prophecies, cognitive dissonance, wishful thinking, diffused responsibility, framing, and a whole set of established cognitive and psychological phenomena (Dror et al 2008, Dror et al 1999, Basola 1999).

To summarise, the McKie and Brandon Mayfield cases, as well as other errors and controversies, have resulted in fingerprint analysis coming under attack from both the judiciary and academia in a way that was unprecedented only a few years ago. Some (Cole 2005, Saks et al 2005) have questioned the very underlying scientific assumptions made by fingerprint experts. Fingerprint examination is now a topic of scientific inquiry in academia, the criminal justice system and the forensic science community (Wertheim et al 2006, Dror et al 2005 and 2006). The US National Academy of Science has created an independent forensic science committee to assess the present and future needs of forensics science (National Academia of Sciences 2008).

This chapter focuses on investigating two motivational influences on fingerprint examination in more detail: need for cognitive closure (the psychological need to bring a decision making process to a definitive conclusion and termination so as to avoid ambiguity or unresolved issues), and emotional experiences (simplistically stated, the feelings (or expected feelings) associated with fingerprint analysis, including reward when one finds a match as well as fears associated with the possibility of making an error).

Kruglanski (Kruglanski et al 1996, Kruglanski et al 2006) found that participants motivated to avoid closure generate the largest number of hypotheses, in contrast to those motivated for a need for closure that produced the fewest alternative hypotheses. As the need for closure rises, quicker judgments are attained with higher confidence associated with those decisions. High need for closure leads to the ‘unfounded confidence paradox’. This paradox arises when there is reduced information processing but at the same time higher confidence in those judgments and conclusions. Thus, a need for cognitive closure may lead to lower decision thresholds, but increased confidence. However, Ask (Ask et al 2005) found only partial support to the hypothesis that investigators with a high need for closure are less likely to acknowledge observations that are inconsistent with their belief of guilt of a suspect. It is therefore important to investigate if the need for closure plays a role in fingerprint analysis.

The need for closure enhances a desire for consensus (Kruglanski et al 2006), thus adopting the ‘path of least resistance’ to achieving agreement with others. This may entail, for example, derogating those who hinder consensus and complementing those who facilitate it (Kruglanski et al 2006) This is important and relevant to the area of fingerprint examination because the need for closure will not only potentially affect the initial analysis, but may be critical to the verification stage at ACE-V (Ashbaugh 1999) as well as arbitration of contentious fingerprint decisions.

Given that decision making can be affected by extraneous influences such as emotional response, context and motivations, then the apparent presence of emotions or motivations in fingerprint examiners will be indicative that these influences may play a role in fingerprint analysis and potentially influence decision outcomes. Conversely, if emotions and motivations were found to be absent from the day to day work of fingerprint examiners then that would support the counter view that fingerprint examiners are immune to such influences on decision making. As yet there has been no investigation into the emotional experiences of operational fingerprint examiners. This chapter examines these issues and will yield valuable insights into the potential role of emotions and motivations on decision making. The qualitative research reported here facilitated an examination of the views of fingerprint examiners without the restraint of preconceived theory or experimental restrictions. The aim of the research in this chapter is to highlight and understand the work of fingerprint experts from a new perspective.

A Qualitative Investigation of Fingerprint Examiner Emotions and Motivations

Although qualitative research has been used in other domains, there has never been an investigation into emotional and motivational experiences within the fingerprint domain. Prior to this research, it was only possible to speculate about the emotional experiences of fingerprint examiners.

The top-down contextual and motivational (and many other cognitive mechanisms) effects often occur without consciousness (Ask et al 2005, Dror et al 2009). As a result, it was not realistic to expect participants to be aware of any information processing effects. Therefore individual interviews were conducted in an attempt to obtain a broad range of views and then perform thematic analysis on the findings in an attempt to uncover trends in the responses of the participants. Thus determining themes and underlying similarities in the experience of the decision makers. Similar research methods were used in Hermsen and Have's (Hermsen et al 2005) research in which semi-structured individual interviews were conducted. They attempted to determine the specific moral and emotional considerations and arguments that might arise from people who must decide whether or not to withhold treatment in a palliative care scenario. They similarly studied a relatively small participant base and as such looked across various care giving environments to get broad, underlying characteristics.

A qualitative investigation was conducted to assess whether there were any emotional and motivational factors that relate to the latent fingerprint examiner's experience. The aim of this research was to find broad themes that occur across the whole sample and not differences between participants. By observing themes that were discussed by each and every participant it was possible to derive themes that could be generalized to the larger forensic community.

Participants

Thirteen latent fingerprint examiners were interviewed from a variety of law enforcement agencies; each had at least 7 years experience. The participants included those involved in the investigation of daily 'volume' crime such as burglary and vehicle theft, as well as others who dealt with the rarer and more serious investigations of rape, murder or armed robbery, in addition to senior officers and managers with a number of years experience. The broad range of the participants' experiences decreased the chance of deriving participant or role specific themes. All participants were fully trained latent print examiners, and performed fingerprint comparison analysis daily. Each interview lasted approximately 30 minutes.

Design and Materials

A semi-structured interview technique was employed, which involved the use of an interview guide. This method was used in preference to a fully structured interview as heavily structured interviews tend to constrain participants' responses towards the researchers preconceived ideas. Rather, the more open-ended structure allowed participants to respond in a naturally ambiguous way. When participants are offered multiple options they tend to constrain their responses between options and as a result it is possible to miss some important areas of internal conflict (Wilkinson et al 2004). Furthermore, structured questions can impose ideas making it possible to lose vital areas of interest that would otherwise be missed, for example how people make sense of their experience.

The interview questions encouraged participants to talk about the various different aspects of their work from the mundane to the more serious and disturbing casework where potentially emotional feelings are engendered, and to attempt to probe any expression of affect that arose. All participants were asked to talk about three aspects of their work:

- Day to day fingerprint analysis processes.
- Particularly harrowing or difficult cases.
- What it meant to them as individuals to be involved in latent print identification.

During these interviews, more direct and probing questions were asked in order to gain further information and resolve any potential misinterpretations by the interviewer. Probing also facilitated the interviewee's own understanding of the framework of meaning without imposing the researcher's assumptions (Britten 1995). It is also important that the questions were not so non-directive that they actually led to more constraint on the participants, as they may spend more time and energy guessing what the interviewer wanted to know (Walker 1985). An attempt was made to find a balance between non-directive general questions that might elicit emotional responses and more direct prompts to examine discussion points in more detail.

Procedure

The interviews were conducted at various operational fingerprint laboratories away from the noise and distraction of the operational environment. Before the interviews, participants were given information and consent forms, and were informed of the general nature of the study. All participants were guaranteed anonymity. All the interviews were recorded so the interviewer could be sure that all information was captured. The interviews were subsequently transcribed verbatim. Also, notes were taken during the interview about how the interviewee appeared, how the interview was progressing, and other appropriate events during the interview.

Analysis

Using thematic analysis protocol codes were assigned to various segments of the text. It was not clear what the findings would be and there was little in the way of guidance from past research. As a result, specifics were inductively coded, whereby individual categories were generated from the interview text itself rather than from specific theory (Joffe et al 2004). Initially, the coding was very broad to encompass anything which had emotional content. Then, as subsequent themes appeared these were broken down into separate codes; for example, distinguishing positive emotions from negative ones. It was important to be cautious not to generate too many categories. Consequently, a few, broad, general themes were chosen as this allowed much greater generalization. It was also important to this particular research that both latent and manifest content was coded. Although this involves a certain amount of interpretation by the researcher, it was hoped that any clarifications made during the interview and the concurrent notes made would avoid inaccurate interpretation.

Reliability is vital in any qualitative study. It was important that the coding was both stable and consistent and that it had good reproducibility (Krippendorff 1980). There is another reliability measure, accuracy, which refers to the extent that the coding corresponds to a previously generated standard or norm, which provides the strongest form of reliability (Weber 1985). However, as this is a new area of study there are no standards to compare it against. As a result it was not possible to measure accuracy reliability. However, it is hoped that the themes from this research might be used to gauge further qualitative studies investigating the emotional or motivational experiences of forensic or criminal investigative personnel. Inter-rater reliability was tested by two independent analyses. Only agreements above .68 were considered (Howel 2002). If sufficient reliability was not apparent, further refinements to the coding were performed in order to increase reliability. To ensure the stability of the initial codes, they were rated twice on two separate occasions. This ensured test-retest reliability and inter-rater reliability.

After assessment of the percentage level of agreement between the raters, it was found that there was an overall coding agreement between raters of .74. Allowing for chance coding agreements a further statistical analysis was performed. An overall free marginal (where raters are not forced to assign a certain number of cases to each themed category) Cohen's Kappa of .69 was achieved (Cohen 1960, Fleiss 1981). It was decided that this met the level of reliability required to ensure the themes highlighted in this study were representative and reliable.

Results

The data revealed five main themes associated with emotion and fingerprint analysis: Reward, motivation, satisfaction, fear and need for closure. These themes were broken down into separately coded categories; job satisfaction and pride associated with the use of skill; motivation, satisfaction and hope, associated with catching criminals and solving crime; the expression of satisfaction and motivation associated with working on more serious or long running cases; the feelings directly associated with searching for and finding matches; expressions indicating a need for closure on casework and emotional feelings associated with making mistakes.

As a precursor to the discussions held with the fingerprint examiners, each participant was given an opportunity to explain the identification process of fingerprint analysis and to go into some detail regarding the methodology of comparing fingerprints. It is important that this be reported within this research. As will be highlighted later, there was objectivity in their methodological description (see table 2) which was in stark contrast to the emotive language and motivations observed later.

Table 2 Examiner quotes associated with methodology

Examiner Theme	Quotes
Methodology	You will run through a series of questions in your mind; quality over quantity, clarity within the mark, the tolerance that you will give to it, what volume of detail have you got? Can you see ridge flow, can you see ridge pattern, what features are there i.e. are there any ridge characteristics? Can you see a scar? Can you break it down even further because in relations to the clarity can you see a particular shape or ridges or pores within there? Anything else which is detail which is going to be useful you to enable you to make a comparison.
Methodology	…assess the value of the marks…we got three basic categories…no value…which means there is so little information in the mark you can never individualise it…then you move on to a mark that is suitable for a direct comparison.
Methodology	…have a look at that mark to see what available information is on that mark….to see if there is any idea of which finger that mark came from…any other indications…left hand…right hand…. Then I would look at the fingerprint form then if I don't know which finger straight away to compare it to…If there was no clear indication, I would compare all ten fingers and analyse the mark from the mark to the fingerprint form looking for anything that looks similar…any points or characteristics that show in both impressions…I build up in my mind what characteristics are similar and I will keep going until I have identified…or not identified.

Job Satisfaction Related to Skills:

There was a great deal of pride and ‘job satisfaction’ (see table 3) exhibited by the fingerprint examiners interviewed associated with the process of fingerprint analysis and the science associated with friction ridge skin. There was a sense of pride in the skills they had learned and a real sense of civic duty and making a difference to society.

Table 3 Examiner quotes associated with satisfaction

Examiner Theme	Quote
Satisfaction	“…from a personal perspective I thoroughly enjoy it [the job] because it involves patience, it involves tenacity, and it involves you really having to concentrate and focus, and the reward comes at the end of the day when you can actually walk away saying ‘I’ve done absolutely everything I can’. ”
Satisfaction	“…using your skill and expertise gives you that bit of drive and feeling, ‘yeah I’ve done something no one else can do’, and makes you feel worthwhile and feel, you know, you can’t be replaced [laugh].”
Satisfaction	“…the thing I still like about this job is the idea that when I get home and I have had a frustrating day and things aren’t going right… you at least know that all my efforts are going to have a tiny but important part of improving society… improving life generally for people a little bit…”
Satisfaction	“…you are doing something useful and you have developed a skill or a talent that is being used and that gives you a sense of satisfaction…”
Satisfaction	“…you are believed in…you are in a position of importance…it’s a nice feeling…”
Satisfaction	“…I am proud of my position in it…proud of what I have achieved.”
Satisfaction	“I am very proud of the service that we do for the public.”

Use of Technology

The increase in computer technology in fingerprint analysis has resulted in some fingerprint examiners feeling under valued as specialists, which could be seen an obvious drain on morale. However, there was not a consensus on the true value of technology in the domain. Indeed, some of the comments from examiners were contradictory in nature, some feeling technology de-skilled them, whilst others got heightened feelings of pleasure from using technology to search for cold cases (see table 4).

Table 4 Examiner quotes associated with technology

Examiner Theme	Quote
Technology	"…which is a shame because we use computer technology more and more and more so it removes the ability of fingerprint officers to use their brain and actually use their skill."
Technology	"Searching has always given more satisfaction…One to one suspects doesn't give you the same buzz I suppose as a search on ident 1."

Satisfaction with Crime Solving:

The sub-category of satisfaction demonstrates how the clinical and scientific job of matching details and patterns within fingerprints has a very human element associated with a personal interest in solving crime and catching criminals (see table 5).

Table 5 Examiner quotes associated with crime solving

Examiner Theme	Quote
Crime Solving	"…I think there will generally be a … very…a lot of …pleasure about it if the case is resolved to a successful conclusion, with a successful conviction, I think that would be a natural thing. If the case remains open then there will always be work to be done. There is always a potential of finding someone."
Crime Solving	"I mean, I was beginning to give up hope of ever matching this fingerprint, I thought, 'oh they'll get it in DNA' they'll, they'll, they'll, find someone and say, 'there we go, that, that's the…the perpetrator', and it won't even match this fingerprint and all my time would be wasted."
Crime Solving	"We catch more in here than the Police officers do on the street and the Police officers are praised and get more money and things [laugh]"
Crime Solving	"…they don't realise the work that's gone on behind the scenes and its nice, it, it is really satisfying, it sounds really sad, but catching people. You don't really see the name or the person you just see that fact that you're hopefully solving a crime."
Crime Solving	"….the whole case was identified to people they wanted it identified to…it was a good result…. That sticks in my mind because I got good feedback from the police officer and the OIC."

These comments demonstrate that matching fingerprints is not just a laborious task of visual search and comparison of details. It appears that examiners feel a direct link between finding matches and actually solving crime. This has significant importance because it suggests that the frame of mind of the examiner is variable in different cases being processed depending on the importance of catching the perpetrator.

Satisfaction Associated with Case Importance:

There were specific comments concerning the experience of reward linked to working on more serious, or longer running cases (see table 6).

Table 6 Examiner quotes associated with case importance

Examiner Theme	Quote
Case Importance	"…for me personally working a long protracted case it is rewarding because you know you are working towards an end goal."
Case Importance	"…that [the feeling] was, that was great, I mean, to be involved in such a high profile case and finally get a match."
Case Importance	"…well it depends on the type of work that you do. Print to print analysis, just so many of those going on it [feelings of satisfaction] doesn't really happen. But again it depends on the severity of the crime. If you're getting volume crime like car theft, or shop lifting, or whatever, and then you get identifications on that, then it's okay. If you haven't' had any in a while in a week or for a month, then it's really good.[…] If it's a more serious offence then I makes you feel even better, er, even more, um, happy with your job,…"
Case Importance	"…the scale of the crime that they were doing was very significant and to actually be a part of that was great, it was really nice, all the benefits are you're actually catching some one up, quite high up the food chain so, it is, really, a really nice feeling."

There were some counters to these statements where the severity of the case was said to be irrelevant and that both volume crime, and the more serious crimes like murder, were treated the same way (see table 7).

Table 7 Examiner quotes associated with objectiveness

Examiner Theme	Quote
Objectiveness	"The fact of the matter is, it doesn't matter what the offence is. What we are focusing on are the crime scene marks and the end result, again, is to complete your analysis, your comparison, and verification to the best of your ability, using your skill."
Objectiveness	"It doesn't matter, really, the size of the case, you know."

In contrast, participants also made comment on their subjective influences (see table 8).

Table 8 Examiner quotes associated with case influence and subjectivity

Examiner Theme	Quote
Case influence	"That's not to say that the same commitment doesn't go with each job but, you, you know, everyone will do more work for a murder than they would for a shop lifting"
Case influence	"…you know, especially with a serious case you are liable to get a lot more suspects and you are still going to have to look at that piece even if you have a strong feeling that it belongs to somebody else other than the person you are looking at…You still have to look just in case because no one wants to be in the position of ignoring something.
Case influence	"Major crime sounds glamorous but you don't actually get out there and see much of the major crime…it's just a pile of work and I actually enjoy the small cases better…the day to day volume crime I actually enjoy better."
Case influence	"…these big cases that start off…I think the worst ones are the drug related which create masses and masses of work and often you are not involved on the investigation side…You might get a few dribbles of information but often you don't get that much…So you are ploughing through great piles of work…It's a job…it's what you are there for…but it's not as exciting as people might think'.
Case Influence	"I don't argue that volume and other crimes are consciously treated differently and indeed the actual process of matching fingerprints is, as stated, identical, whether it has come from a murder or from a house burglary"

This discrepancy may not have been found under a structured or questionnaire research methodology and demonstrated the effectiveness of an open ended interview technique. One participant responded to direct probing of the differences between volume and other crime. It led to a direct sounding answer whereby the examiner stated that there is a conscious decision to treat volume and other types of crime differently.

However, it does appear that although they are treated the same with a consistent level of "commitment", and using the same comparison techniques, the end result has different impact, and the desire to find a match appears to be stronger depending on the crime severity.

Feelings Associated with the Act of Finding a Match:

Apart from direct job satisfaction and desires, there were lower level emotional responses associated with finding small areas of similarity within two fingerprints that corresponded or in the determination that two prints matched. There was a wide range of responses, from descriptions of feeling a "buzz" as a direct response of matching the prints, to an emotional outburst in one case (see table 9).

Table 9 Examiner quotes associated with finding a fingerprint match

Examiner Theme	Quote
Finding a match	"…that feeling when you know you've identified someone because all the features correlate"
Finding a match	"…oh it's a buzz. It's a definite buzz. […]. When you get one, especially from the search, the buzz is there."
Finding a match	"…I was getting used to turning over every set of fingerprints I saw because the palm prints are on the back and thought "heyup, what's that?", and it was like "WAHEY!" and a really, really good, really good feeling".
Finding a match	"I was just beginning to get the feeling that it was a match…"
Finding a match	"…You pick your initial target, you know you're first feature you're going to look for, and then you look through you prints and you recognize it. That gives you a little, encouragement, you know, I've got something to focus on, somewhere to start, um, […] and…you know, every time you see something you recognize your confidence builds in the fact that it's a match, and the end point is "can I build my confidence to absolute confidence". You know, "can I eliminate all doubt in my mind whether these two prints came from the same finger, and it's a process of eliminating doubt."
Finding a match	"…then all these recognition events pile up in you brain until you, you, in a way you've got no choice but to come to the conclusion that they were made by the same source. It just becomes overwhelming and it's just like seeing your friend down the pub…'I know who that is' ".
Finding a match	"Some people are just naturally slow where everything has to be done perfectly, they have to dot every I and cross every T, check every little bit of scrap. Where other people would be a lot more cavalier about it, but be quicker and perhaps get more idents,… I don't know..."

Finding a match	"… everyone perceives things differently …the levels of information they are looking at varies between person to person…confidence levels vary…I tend to think I am pretty much middle of the road…I won't go over the top and count every single characteristic in a palm impression but at the same time I am looking to find a suitable amount to satisfy myself."

This suggests that not only are there motivational factors associated with solving crime but there are direct emotional feelings associated with finding fingerprint matches. Furthermore there are indications of emotional responses during the process of matching prints as well, i.e. before a definitive conclusion has been reached. There are descriptions of a build up of "recognition", and increases in "confidence" and "encouragement", which appear to enhance the "feeling of a match".

This finding describes minor positive emotional responses of recognition as a result of seeing areas of agreement during a comparison. There are small emotional rewards of matching individual targets within the whole fingerprint before a tipping point is reached.

These reports by our examiners of the feeling of accumulating evidence until a specific level of confidence is met confirms the theoretical decision making threshold 'winner take all model' (Busemeyer et al 1993). This suggests that fingerprint examiners use this type of decision mechanism to make their judgments. Specifically that evidence accumulates over time to a specific, but malleable, level where a decision can be made, rather than a normative model of evidence deduction and an objective judgment. Therefore, prints are not said to match because logical deduction has proven them to, but that for the examiner concerned, their subjective level of confidence or their 'decision threshold' has been met by the accumulation of evidence.

Furthermore it was reported that different examiners appeared to have different decision thresholds. A decision maker with a lower decision threshold will result in faster decisions as they require less evidence before the same degree of confidence is met. Whereas, a higher decision threshold results in slower decisions as more evidence must be accumulated.

This explains how a system that is supposed to be objective can result in a difference of opinion between examiners. The primary concern is whether or not contextual biases and external pressures might influence these threshold levels for the finger print examiner, i.e. to what extent are these threshold levels determined by normative, objective prescription, or by subjective, context dependant, mechanisms.

Fear

There was a strong sense of fear associated with making mistakes in fingerprint examination. When asked, examiners for the most part asserted that to make an erroneous match was the very worst thing an examiner could do in so far as a person might be wrongly arrested. While there was also an expression of fear in making a false negative call, there seemed to be less emphasis placed on this type of error. In fact, some suggested that misses were just a part of being human. There also appeared to be a value placed on the fear associated with either a false negative over a false positive. For example, because generally there seemed to be a primary fear of making a false positive judgment, this appeared to weight the attitude of the examiner toward a more conservative demeanour (see table 10).

Table 10 Examiner quotes associated with fear

Examiner Theme	Quote
Fear	"I know everyone is human and you can make errors but I would probably feel awful like I can't do my job properly."
Fear	"…I think 'is it my judgment that's wrong…or someone else's?'…but then you have to remember that fingerprints is opinion…it's not an exact science…it's our opinion" "a wrong ident,…you are doing something badly wrong…That's what I would be more worried of doing."
Fear	"You should not miss,…should not have a wrong ident…A wrong ident is out of the question…I don't think it should happen…It happens…Unfortunately it happens"
Fear	"To actually miss an identification could hurt the individual as much as making an erroneous identification…But obviously the implications behind the two will be slightly different…I suppose there is a tendency to believe that the cardinal sin is a wrong identification…Missed identifications may not necessarily lead to problems…"

Fear	“Fear? Only fear of making a wrong decision…I think. I think that’s the fear. So you just wanna be sure that you have made the right decision so you will probably err on the side of safety because it’s better to let the ident go than to make the wrong ident”
Fear	“The management, certainly when I was training, would make it quite an official thing…You have to sign a sheet saying you missed it and put any comments down…The manager who had started it off would put some comments down and would go down on file so if you did another one soon afterwards it would be brought out and it would be a far more serious thing…You weren’t allowed many misses before it got serious and that’s as a trainee’.

Closure

It was apparent from the data recorded that the examiners interviewed expressed, in general terms, a desire to avoid ambiguity and to see cases through to conclusion. Some examiners displayed feelings of frustration at not being able to finish tasks. In addition, there was a desire to account for all the evidence and to seek out a definitive solution to the casework. There was strong evidence of a need for closure (see table 11).

Table 11 Examiner quotes associated with a need for closure

Examiner Theme	Quote
Need for closure	“…its annoying…It’s like ohhhhh…I got that one little bit left’….
Need for closure	“I would like to finish it up…maybe I’m a bit of a perfectionist occasionally …I need to complete everything”
Need for closure	“Once I start something I like to finish it…and it’s nice to finish it…and as a fingerprint expert it’s nice to have a case wrapped up’.
Need for closure	“You would like to have a result in a case…i.e.…that the mark’s been identified to a suspect or the mark has been eliminated…Whereas, the mark’s not being identified or eliminated is hanging in the air…you would like a result either way’….
Need for closure	“The chances of being able to account for every single piece are slim…its nice when you can do it”

Need for closure	"You like to…because it clears the job up…if you identify the eliminee…they could have a record…have their prints previously on file…at least that is the job cleared up…and that's the important thing…"
Need for closure	"It's nice if there aren't unnecessary marks on the database…because they are being searched against unnecessarily …it's nice to know that that job is finished…all the marks have been checked and assigned to whatever outcome and you know you don't have to revisit that job."
Need for closure	"It gives you a better sense of closure."

Discussion

There were a number of significant findings as a result of this investigation. Some important motivational and emotional factors appeared to be an integral part of the working life of fingerprint examiners. There was a considerable amount of pride and satisfaction associated with the skills they had learnt and used daily. In addition, there was a significant personal interest in catching criminals or solving cases, especially when it related to high profile, long running, or serious crimes. Experts described the process of looking for matches in emotional terms and, specifically, described matching in terms of 'feeling' and reaching a specific threshold at which they can make a final determination. Scientific analysis of fingerprints and the comparison and evaluation of such material has always been assumed to be an objective process, yet clearly there are subjective elements introduced by human factors and their interaction with the ACE-V methodology for comparing fingerprints.

These findings are very important as they indicate specific cognitive mechanisms. For example, the Madrid bombing case (Thompson et al 2005, Frieden et al 2006), was a very high profile and important event. The investigators working on that case were highly motivated to get 'a result' and close the case. That is not to say that they would have intentionally falsified matching the two prints. Rather, they may have reached the decision that the two prints matched sooner, based on less evidence than they may have ordinarily required. The combination of a strong underlying motivation to find a match or close the case, as well as smaller emotional feedback, when finding small similarities between the two prints, might have had an effect on subsequent information selection and processing which may have resulted in the decision threshold being lowered. The experts would have felt that they had performed their job accurately and correctly since a subjective feeling of heightened confidence would have been experienced.

Given that participants generally viewed major crime as being more rewarding; this may act as an emotional amplifier by increasing the potency of the emotional rewards and moving them closer to the threshold at which a conclusion is reached. This would result in decision thresholds being met with different levels of evidence depending on the context and the type of crime. If this was to be the case, then the chance of erroneously matching prints might increase as a result of context, such as case severity or drive for closure.

There was also an expression of fear and consequence in making an erroneous match. This fear of error may result in more conservative decision thresholds which would entail incorrect non-identification conclusions (misses). Examiners seemed to feel that missing an identification was less important than falsely identifying an individual. Some examiners acknowledged however that to have too many false negatives would be detrimental to the professional standing of any examiner. It is just as important to understand why examiners miss identifications as it is to understand how erroneous identifications arise. Both are incorrect conclusions.

Another key area of interest in this investigation was the apparent need for fingerprint examiners to achieve closure on casework. Many of those interviewed appeared to make comment suggestive that they possessed a high need for closure. Such examiners would exhibit a stronger desire to obtain a definitive answer, as opposed to uncertainty and ambiguity. These examiners would tend to prefer the company of those with similar attitudes and philosophies and feel positively disposed toward those who allow for consensus. Similarly those examiners who require need for closure may feel negatively toward those who deviate or jeopardise consensus.

Examiners with high need for closure may make correct judgments so long as the cues initially seized upon were correct. However, examiners with a high need to avoid closure may also commit errors if they too readily unfreeze correct judgments through excessive openness to misleading or irrelevant information. In other words, fingerprint examiners might be vulnerable to error through a heightened need for closure that may impact upon verification and arbitration discussions by arriving at an erroneous consensus through associating with other examiners who are likely to agree with them. Or, conversely, may miss identifications because they were unable to come to the right conclusion because they looked at the mark for too long and effectively talked themselves out of it.

Examiners under a heightened need for closure may seize on information appearing early in a sequence during a fingerprint comparison and freeze on it, ignoring or unfairly weighting subsequent information within the fingerprints that may offer an alternative hypothesis. Examiners with high need for closure may process less information within the fingerprint before committing to a judgment and generate fewer competing hypotheses to account for the available data.

To put it in terms of the threshold theory, high need for closure results in a lower decision threshold and therefore less information is required before the decision maker can close the case and make a judgment. It is the 'seizing and freezing' (Kruglanski 1996), which may be central to the notion that contextual information can bias decision making.

The ability to generate alternative hypotheses depends on cognitive capacity, environmental conditions as well as contextual information and how tasks are framed (Sommers et al 2007). Validation of hypotheses is accomplished through deductive logic, where there is confidence that the hypothesis fits in with known facts (Kruglanski et al 1983).

The process of generating a hypothesis in fingerprint examination may be prompted by an interest in acquiring knowledge and is as a result a motivational act or behaviour (Kruglanski 1980). Such motivational forces may dictate when the knowledge search both begins and ends, resulting in the validation or invalidation of the hypotheses. The ability of the examiner to create these hypotheses may be dictated by various motivational forces such as a need to achieve specific conclusions, a fear of invalidity and a need for closure (Kruglanski et al 1987).

A need for specific conclusions can either encourage or discourage hypothesis generation; hence it is a directional motivation. If the hypothesis falls in line with the wishes of the individual then they are more likely to accept the hypothesis and halt the search for other hypotheses. If the hypothesis generated is undesirable to the individual for whatever reason (in latent print verification this may arise when the hypothesis generated may indicate a colleague has made a false positive match) then a search for an alternative hypothesis will continue until a more plausible hypothesis turns up (That is to say to find information that will confirm a colleague's findings). This will inevitably lead to what has been termed 'wishful thinking' bias (Kruglanski et al 1983) and could lead to catastrophic error. It is often said in the fingerprint domain that errors are very rare and that verification processes underpin this assertion. A need for a specific conclusion in addition to a shut down of alternative hypotheses would clearly be detrimental to effective verification of results and conclusions.

A fear of invalidity is a consequence of an individual's perceived cost of making an error of judgment. A fingerprint examiner with a high fear of invalidity would, in theory, generate a greater number of hypotheses during the decision making process and be sensitive to information that did not comply with currently held beliefs (Mayseless et al 1987). Roger Koppl advocates competitive self regulation (Koppl 2005) to ensure every forensic laboratory has an incentive to give professional, scientific, and objective tests. Competitive self regulation requires casework be assessed by independent agencies as part of the verification process. The act of such validation outside the home agency will give rise to greater impartiality and objectivity since the checks are blinded (no contextual information), taking away much of the emotional context. In effect Koppl advocates a form of induced need to avoid closure in examiners by encouraging through this process an environment where there is no motivation to close down discussion and no motivation to shut down hypothesis generation since there is little context or emotional reward generated because of the 'blinding' process.

However, while procedures and processes within fingerprint bureau have been designed in the sincere belief that such processes as Koppl advocates will protect against bias and other cognitive shortfalls, it is by no means certain that this has been universally achieved at this time and certainly there has never been any research to confirm the effectiveness of these procedures when applied to human behaviour. When considering a laboratory environment, one cannot divorce the individual from the group when examining the likely effectiveness of decision making. It is important to note that it is the external environmental conditions such as social interaction and other life experiences, as well as more personal tendencies such as lifestyle and beliefs that also shape who we are and how we perform (Asch 1955, Deutsch et al 1955, Hardin et al 1996).

It has been shown that individuals who demonstrate a high need for closure will endorse autocratic decision making (Kruglanski 1989, 2002), and be intolerant of group diversity. They may reject and fail to absorb the opinions of those who deviate from their own thoughts (Ask et al 2003) and philosophies leading to group pressures for conformity (Back 1951). Such individuals will have favourites and will take steps to marginalize those who disagree (out group derivation). I have personally observed this in the real world of latent print analysis and it is true to say that verification processes in latent print analysis are often blighted by what has long been termed 'the buddy system' whereby examiners will go to other examiners who they believe will best be able to confirm their own conclusions. Those individuals with a high need for closure will invariably display a mentality of intransigence and dogma (Livi et al 2006), resulting in critical derogation of groups that disagree with their own thoughts (Dechesne et al 2000, Jackson 2002). If these traits were present in fingerprint examiners, the implications could be serious.

A need for definite knowledge can lead to what is known as the urgency tendency to attain closure as soon as possible (Wei et al 2005, Kruglanski et al 1996). Also, the permanence tendency is an individual's inclination to maintain closure for as long as possible. So not only might it be possible for fingerprint examiners to make decisions too hastily should they be prone to a need for closure, but also, they will hang on to the conclusions, perhaps irrationally as a result of the need to maintain closure. In short the need to close the case down and draw early conclusions based on data observed.

Need for specific closure implies the desirability of a particular answer to a question. Need for non specific closure implies the desirability of any answer as long as it is definite (Kruglanski et al 1996). Need for closure may bias the fingerprint examiners' choices and preferences toward closure bound pursuits inducing negative affects when closure is threatened and positive affects when it is facilitated. Need for closure may be heightened under noisy conditions, when the task is unpleasant or dull or when the individual is fatigued. It may also be heightened when closure is valued by others because possessing closure may earn their esteem and appreciation. In a police culture where reputation and esteem is all important, it is hypothesised a high need for closure may be evidence of more cultural investigative motivations that may lead to institutionalised biasing conditions.

People under a heightened need for closure will often 'seize' on information appearing early in a sequence and freeze on it (Kruglanski et al 1996), ignoring subsequent data impacting on information processing. People with high need for closure process less information before committing to a judgment and generate fewer competing hypotheses to account for the available data. The fewer the competing hypotheses the more confidence is exhibited by the person in the ones that remain (McKay et al 2006). In the methodology of latent print analysis, ACE-V (Ashbaugh 1999), this may have profound implications to the effectiveness of the methodology should need for closure be observed amongst examiners.

People with high need for closure often prefer those with similar mindsets. They will feel positively disposed toward those who facilitate consensus and feel negatively toward those who deviate or jeopardise consensus. A person with high need for closure may make correct judgments so long as the cues initially seized upon and froze on was correct (Kruglanski et al 1996). Conversely, those with a high need to avoid closure may commit errors if they are too readily willing to unfreeze correct judgments through excessive openness to misleading or irrelevant information. Earlier, I discussed the relative importance placed on types of error in latent print analysis. It is hypothesised 'seizing and freezing' phenomenon may provide contributory factors leading to erroneous matches, but just as importantly, may help to explain how type two, false negative matches occur. Understanding seizing and freezing may help us to more fully understand how forensic errors come about. Individuals high in need for closure limit information processing activities. Kruglanski suggests the relationship between intelligence and need for closure is not significant (Kruglanski et al 1996). It may be interesting to further investigate this concept to see whether there is a correlation between need for closure and expertise.

For those high in need for closure, thinking for its own sake is undesirable. This is especially true for ambiguous information where one might expect ambiguity to lead to more deliberate and considered decision making, but where in fact need for closure impedes this process (Van Hiel et al 2002). Kruglanski showed through experiments that those motivated to avoid closure generated the largest number of hypotheses, participants conditioned to a need for closure produced the fewest hypotheses and were quicker to attain high judgmental confidence (Kruglanski et al 1996, Kruglanski 1983). High need for closure leads to the 'unfounded confidence paradox' of reduced information processing against a backdrop of heightened confidence in a given hypothesis. In addition, a high need for closure will impact upon the type of information processed, for example, acceptance of prototypical information over more diagnostic interpretation leading to false assumptions. It is possible that high need for closure individuals change their mind abruptly and completely if later evidence is particularly compelling however and they are motivated to accept what they see. Need for closure also appears to enhance theories around anchoring effects (Englich et al 2001) in that those high in need for closure tend to overestimate the likelihood of conjunctive events and underestimate disjunctive events.

Need for closure tends to enhance a desire for consensus. When information present requires people to crystalise a prior opinion, Kruglanski states that people profess a desire for an easily persuadable partner so that such a partner can be readily won over. By contrast, lack of information, such as in an ambiguous latent print analysis, will illicit a desire on the part of an individual to seek out a persuasive partner to employ a 'change self' strategy, and to aspire to have some other person effectively 'make up your mind for you'. If these two strategies fail then it can be possible to reach consensus by 'rejecting the deviate', by excluding dissenters (Kruglanski et al 2002 and 2006), and by associating with those who it is felt will allow for a quick conclusion.

Kruglanski suggests people are more open minded in group consensus situations when time pressure is perceived to be minimal (Kruglanski et al 1996). Where time pressure is enforced then need for closure conditions might be amplified (Shah et al 1998). This can become an acute problem when complex information, such as that which might need to be exchanged between two examiners about an ambiguous latent print is conducted in a time pressurized environment. In such circumstances, should examiners exhibit traits associated with need for closure then communication of findings to an experienced examiner may be limited in detail compared to discussions with examiners who are less experienced.

Kurt Lewin coined the term ‘group dynamics’ to describe the way groups and individuals act in a given set of circumstances (Lewin 1947). Lewin suggested that neither nature (inborn tendencies) nor nurture (how experiences in life shape individuals) alone can account for individuals' behaviour and personalities. When considering a laboratory environment, one cannot divorce the individual from the group when examining the likely effectiveness of decision making of one examiner over another. It is important to note that it is both the external environmental conditions such as social interaction and other life experiences, as well as more personal tendencies such as lifestyle and beliefs that shape who we are and how we perform. If social, as well as personal behavioural traits were combined with a need to please in order to gain promotion (Moran et al 2008) then the consequences for accurate and objective investigation could be impeded.

The authority of the group and as individuals depends on the ability to construct a factual consensus. This is especially true of the fingerprint profession (Cole 2002, Festinger 1950, 1954). It is hypothesised examiners may join certain groups of individuals to reduce uncertainty (Hogg 2000). These groups will form through the use of inappropriate discussion strategies that may lead to a further cultural conformity of the entire organization (Fu et al 2007). Motivated choice of language as evidenced in this research may affect how people act and feel toward each other within the workplace (Webster et al 1997). For example if someone is asked why he or she liked a particular film, the answer is predominantly around the nature of the good acting and directing. In contrast, if a causal question is asked in concrete terms the answer will follow the logical subject of the sentence, for example, 'why did you go to see the film’; the answer may be ‘I needed a break from work’. In fingerprint examination during discussion of ambiguous marks, the use of language and how discussions are framed is vitally important to maintain objectivity. Discussions along the lines of 'why did you make it a match' could be followed with the answer ’I found enough in agreement’, as opposed to 'explain how you came to this conclusion on this comparison' where the answer will be more detailed. If one adds other personality traits such as general rudeness (Chiu 2003), which can also impact upon communication processes, then discussion can be further degraded between individuals.

There is an association between internal security and information processing (Mikulinser 1997). Secure persons hold internalised surety, or self confidence that enables them to take risks in exploring novel stimuli. In contrast, insecure people lack a sense of mastery and optimism in dealing with actual or potential threats. It leads them to reject evidence that can create confusion and uncertainty by closing their schemata to this information. Insecure persons perceive more threats in incorporating new data and more benefits in maintaining stable knowledge than secure persons. This raises questions at a time of fiscal constraint where forensic practitioners are under threat of losing their jobs and when budgets are forcing curtailment of forensic processes and procedures in favour of cost effective strategies. It could be that those who exhibit tendencies associated with insecurity in the fingerprint domain have these feelings amplified at times of uncertainty. For example, if job security is threatened and performance culture based on the number of identifications achieved is important in determining job security, then it is possible examiners who lack surety will be less willing to accept new data to contradict findings, preferring a status of safety, maintaining the identification is valid, closing down acceptance of new information that may lead to a negation of the fingerprint comparison and conclusions reached. In short, the examiner may risk error to maintain the status quo, rather than be open to novel stimuli that may counter their original hypothesis, for fear of looking sub standard or inefficient in the face of competition for jobs and livelihoods.

Need for closure and the fear of getting it wrong (invalidity) have opposite motivational effects on hypothesis generation, it is possible for a person to be highly prone to both states, or less prone in both states, or high on one and low on another etc (Kruglanski et al 1983). Both need for closure and fear of invalidity will depend upon situations. In addition, both constructs will influence knowledge acquisition in a stable and consistent way across different situations (Kruglanski 1989). Thus, it can be hypothesised that fear of failure in fingerprint analysis should, in theory, negate any need for closure traits. Further investigation in this area is warranted.

All human judgments contain a motivational component arising not only from personality dispositions but also from contextual influences (Webster et al 1994). One of the most obvious features of experts (such as fingerprint examiners) is that they are often superior in their performance, compared with novices and are able to reason more efficiently through experience. However, as already discussed, most experts make errors in their domain at some time or other. These errors often result from a disruption in the processing of the very cognitive abilities that serve to facilitate expert performance. Conditions, including context and motivation, time pressure and group centrism may disrupt the processing of these mental abilities, degrade expert performance and may lead to human errors.

It is possible that decision making thresholds of fingerprint examiners is dynamic along an elastic continuum that is dependent upon certain factors including; the cost of error, motivation to be accurate, time pressure, the importance of the case, the context in which evidence is framed, the individual traits of the examiners themselves, such as need for closure, as well as the environmental conditions and culture within which fingerprint examinations take place, such as background noise and interference. It is this theoretical framework of decision threshold which might explain how biasing factors affect the decision process in some scenarios but not in others.

So while contextual influences may be broadly observed in fingerprint analysis and other scientific disciplines, it will be important to understand at what point contextual bias impacts upon the actual conclusions of the examiners. Bias and cognitive influences affect the decision process but not necessarily the decision outcome.

The aim of the above discussion is to stimulate further discussion and research rather than to deliver final conclusions. However, this study presents some exciting questions about the nature of top-down motivational effects and contextual influences and the possible catalysts that may exacerbate such phenomenon. In future research it would be valuable to look in more detail at the concept of need for closure in fingerprint examiners. For example, need for closure may bias the fingerprint examiners' choices and preferences to facilitate attaining closure. Need for closure may or may not be a generic feature of fingerprint examiners and this will need to be investigated. If the phenomenon is present, may it be mitigated or amplified under environmental noise, when the task is unpleasant or dull or when the individual is fatigued?

As with any research there are potential weaknesses as well as strengths in the investigation reported in this chapter. For example, the lines of questioning could have gone into more detail about the correlation between methodological objectivity and how participants felt this process was affected by the emotions and motivations highlighted. This could be an area for further investigation.

What is certain is that fingerprint examiners not only are emotionally driven and motivated to achieve results for themselves, their employees the police and wider society, but also that there are more subtle psychological factors such as need for closure that may exert leverage upon the decision making thresholds of examiners that may, in the right circumstances lead to erroneous conclusions should the context and the motivation be strong enough. Only by understanding this phenomenon will it be possible to guard against future error and methodological breakdown of fingerprint analysis, as well as design and implement effective and robust recruitment, selection and training environments that are able to provide best practice for examiners and to satisfy public confidence in not only fingerprint examination but also other forensic domains as well.

This chapter provides preliminary understanding of how motivation and emotion interacts with attempts at objective scientific endeavour. This investigation concentrated on the domain of fingerprint examination, but it could just as easily apply to other pattern recognition disciplines in law enforcement such as tyre tread interpretation, shoe wear analysis, and even archaeology, where there might be strong motivation to preserve a theory or notion instead of objective interpretation of the evidence. For example, when is a striation mark on skeletal remains evidence of battle scars (if the archaeologist wanted to confirm a battle took place at a particular venue), as opposed to normal wear and tear of a bone as a result of burial at a deposition site over many centuries?

Having established that fingerprint examiner objectivity is superficial in that such ‘objectivity’ is clouded by the motivations and emotional traits of individual examiners, it will now be important to quantify the level to which other influences on decision making may impact upon fingerprint examiner accuracy and consistency, namely top-down cognitive influences such as context.

Chapter 3:

When Emotions Get the Better of Us: The Effect of Contextual Top-down Processing on Matching Fingerprints

The need to identify people accurately and quickly is important in both the civil and law enforcement domains. With advances in science and technology, which include automated fingerprint matching tools as well as remote transmission facilities, a variety of methods are now available for identifying people more easily and quickly, which maximises the potential to recover lost property and prevent the perpetration of further criminality.

Figure 24 Automated fingerprint matchers are now commonplace.(Fingerprint Society archives)

Fingerprints continue to be the major method used for identification in forensic and other domains (Alam, Akhteruzzaman and Cherri, 2004). Fingerprints are comparatively easy to find, collect, and process; and they are also relatively non-intrusive. With the development and increased use of computer technology (see figure 24) in searching very large amounts of fingerprints held in databases, fingerprints are likely to continue to be the major method for biometric identification.

The strength of fingerprint identification also derives from perceived reliability. The use of fingerprints has evolved over a long period of time and for over 100 years fingerprints have been used quite successfully as a means of identification. The reliability of fingerprint evidence stems from applied scientific knowledge of the uniqueness of friction ridge skin within the fields of biology, embryology, and genetics (Babler 1977, 1978 and 1987, Galton 1892, Ashbaugh 1985, 1991, 1999 Vincent 1985, Penrose et. al 1980, Faulds 1912, Samishenko 2001). There has not been a single reported case of two different people having identical fingerprints (even identical twins have differentiated fingerprints).

Fingerprint identification seems to have withstood the test of time and proven itself as a sound and authoritative tool. Consistent with the above, whereby there is overt trust placed in fingerprint evidence, research has shown that fingerprint evidence affects the perceived innocence or guilt of defendants (Bregman & McAllister, 1987).

In recent years the reliability of fingerprint identification has come into question and is under close scrutiny (Cole 1999, Moenssens 2003). A publicly exposed erroneous identification in the US has highlighted a weakness in fingerprint identification. In this case an individual (Brandon Mayfield see figures 25 and 26) was wrongly linked to the Madrid bombing based on a fingerprint match found by the Latent Print Unit (LPU) of the Federal Bureau of Investigation (FBI).

Figure 25 Brandon Mayfield (Internet Image)

A number of experts, including an independent examiner appointed by the court upon the request of the defendant, all confirmed the initial finding of a match. However, a few weeks after arresting the suspect, this match was proved to be false and he was released (for more details, see full report on this erroneous identification, Stacey, 2004).

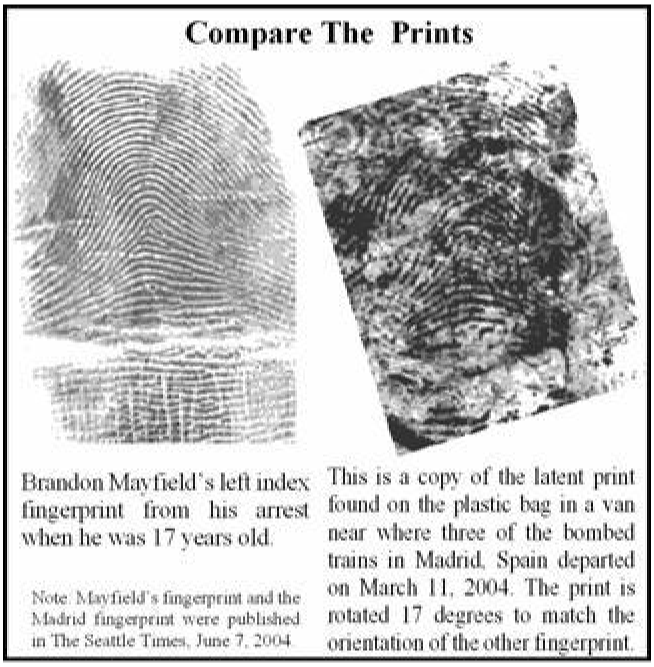


Figure 26 Brandon Mayfield erroneous comparisons (Courtesy of Kasey Wertheim CLPEX collection)

This is only one of a handful of cases that has been publicly exposed and acknowledged. Such cases are only found under extremely unique and rare circumstances, and it is unreasonable to believe that other erroneous identifications have not occurred.

The research reported here attempts to examine some of the processes involved in fingerprint identification and factors that may interfere with these processes. Fingerprint identification involves a decision making process. This requires making a decision as to whether or not pairs of fingerprints match (for example, whether a fingerprint lifted off a crime scene matches that of a potential suspect). Such decisions, as with many other cognitive processes, are composed from two main components: First, the bottom-up component which is purely data driven (see for example, Ashworth & Dror, 2000); and the second is the top-down component in which contextual effects mediate how the input is processed, evaluated, and a final decision is derived (see for example, Dror, Busemeyer, & Basola, 1999; Levy, Ashman, & Dror, 2000).

As per the bottom-up component, each fingerprint is composed of a pattern which is believed to be individually unique.

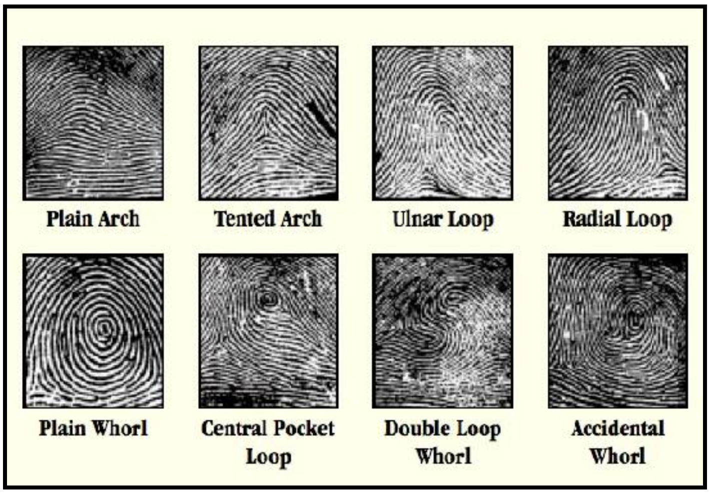


Figure 27 Fingerprint Patterns (NPIA Training Manual)

A close examination at three different levels (see figure 27) can help decide if fingerprints match or not. The first level examines the overall pattern of friction ridges; the second level examines the characteristics of specific ridges; and the third level (see figure 28) zooms in to examine things such as locations and distribution of sweat pores, individual ridge topology, and other uniquely identifiable features.

Figure 28 3rd level detail including sweat pores (NPIA Training Manual).

The examination of such bottom-up information means that if a decision can be made, it would constitute compelling identification evidence (or lack thereof). In an ideal world, such decisions would seem to be rather simple and relatively easy to make: either there is, or there is not, a match. However, in the real world many fingerprints are far from perfect. They are often degraded and partially missing and are often distorted by the substrate upon which the latent print was deposited (Maceo 2005) as well as the constituents that make up the essence of the latent print, such as sweat, oil, grease, and other contaminants.

Figure 29 Degraded and partial fingerprint image on left with the exemplar on the right….this is not a match (Courtesy of Kasey Wertheim CLPEX collection).

In fact, even in an ideal world, should an individual provide many sets of fingerprints one after the other, even then they would not be 100% identical (the varying pressure on the skin's elasticity, among other things, would produce slightly different prints). Since no pair of fingerprints can ever be 100% identical, one needs to decide if they are similar enough to determine that they originated from the same individual (see figure 29). In many cases the information contained in the prints (especially those collected at a crime scene) is not enough to enable a sound decision (Ashbaugh, 1999). Although fingerprint matching is a complex and challenging pattern recognition problem, it is important, if not imperative, that decisions are accurate. This becomes an even more monumental task when you take into account top-down processes involved in pattern recognition.

The other component involved in deciding whether there is a fingerprint match is top-down processing. A top-down component occurs when the processing of incoming bottom-up information is mediated by a variety of factors, such as motivation, as we found in the previous chapter, prior experience and knowledge, as well as the person's expectations and emotional state. Top-down processing can facilitate the processing of information by making it more efficient and faster (for example, help direct attention to important features in object recognition, e.g. Dror & Kosslyn, 1998). It can also help interpret ambiguous information (Selfridge, 1955) or fill in missing information (for example, the phenomenon restoration affects e.g. Warren, 1970).

However, in some cases top-down influences are so pronounced that they can even override the 'objective' information coming in as input from the bottom-up component (for example, different top-down information leads to contradicting judgments on the same bottom-up data, e.g. Darley & Gross, 1983). Thus, top-down components can interfere with and distort the 'objective' processing and evaluation of incoming data. Top-down is a term that encompasses a very wide range of phenomena, such as expectation, hope, context, knowledge, emotional state, and mind set, to name but a few. Indeed, 'mind set' has been identified as one of the main contributors to the FBI erroneous identification (Stacey, 2004).

A large body of research demonstrates that the emotional state of the individual plays a critical role in how they interpret information, and specifically that their interpretations correspond to their emotional state (Byrne & Eysenck, 1993; Eysenck, Macleod, & Mathews, 1987; Halberstadt, Niedenthal, & Kushner, 1995; Niedenthal, Halberstadt, & Setterlund, 1997; Pincus, Pearce, & Perrott, 1996; Richards, Reynolds, & French, 1993). Many of these studies involved looking at lexical ambiguity and revealed that both state anxiety and trait anxiety were linked with increased tendencies to adopt negative interpretations (Byrne & Eysenck, 1993; Richards et al., 1993; Russo, Patterson, Roberson, Stevenson, & Upward, 1996). In addition to verbal stimuli, similar findings linking emotional states to associated interpretations of stimuli have been found using facial expressions (Niedenthal, Halberstadt, Margolin, & Innes-Ker, 2000; Richards et al. 2002), interpersonal situations (Hirsch and Matthews, 1997), and even physical sensations (Calvo and Eysenck, 1998).

Given that fingerprint matching occurs frequently within a highly emotional context of forensic evidence associated with finding those who committed crimes, it is important to examine how emotional states may affect fingerprint identification. There has not been any research examining top-down influences on fingerprint matching. However, research has shown that presenting gruesome evidence does influence the verdicts of mock jurors (e.g. Bright & Goodman-Delahunty, 2004).

The research reported here examined people's decisions on matching fingerprints. Observations assessed if and how their decisions were influenced by top-down information. Emotional states were manipulated and participants were motivated to find a fingerprint match by providing background information and through subliminal priming. This investigation was conducted to see if such top down manipulations can affect their decisions, and to what extent.

The level of top-down influence was manipulated as well as the actual difficulty of the task. The main manipulation of the top-down component was achieved by introducing information about the background of the crime, where the fingerprint was collected and by including explicit and emotionally provoking photographs. To further strengthen the top down effect subliminal priming was introduced by introducing the words 'guilty' and 'same'. This was to examine the general vulnerability of the matching process to top-down external influences. The task difficulty was manipulated to examine the possible strength of the top-down component. As the match becomes more difficult, so the more room there is for varying levels of ambiguity in the bottom-up information that was provided to the decision makers.

The following investigation assessed whether top-down contextual information would impact upon decision making during the examination of fingerprints.

Participants:

The participants in this study were made up of 27 university student volunteers, with a mean age of 23 (nine were males and 18 were females).

Materials and apparatus: Fingerprints:

Ninety-six pairs of fingerprints were selected from a large fingerprint database (Maltoni, Maio, Jain, & Prabhakar, 2003). This database facilitated the use of an established set of stimuli from which to construct experimental conditions: half of the stimuli (48 pairs of the fingerprints) provided clear and detailed bottom-up information and hence were relatively easy to decide (see figure 30); 24 of them presented a perfect match whereas in the other 24 pairs of fingerprints it was clear that they did not match.

The other half of the stimuli (the remaining 48 pairs of fingerprints) was not as complete and detailed and hence did not provide sufficient bottom-up information from which to make a clear decision.

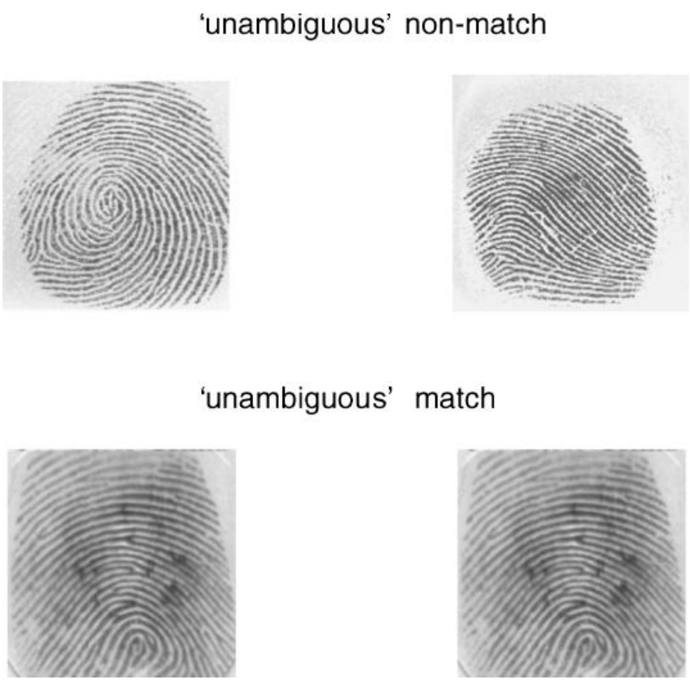


Figure 30 Example of 'Unambiguous' pairs of fingerprints…match and non-match (personal collection)

The ambiguous pairs (see figure 31) of fingerprints were introduced for two main reasons. First, it was for ecological validity; many fingerprints in real world applied settings are far from perfect. Second, weakening the bottom-up information may allow the top-down component more room to influence the process.

Figure 31 Example of an ambiguous pair (personal collection)

Background information

Two emotional states were invoked (low and high) by exposing the participants to background stories and photographs. The background for the low emotional state included stories about bicycle theft, burglary, and other relatively common crimes that do not include physical harm to a person. The high emotional evoking stories included a variety of crimes, such as murder, personal attacks, and other cases where there is a victim who is seriously hurt.

To further induce the emotional state photographs from the crime scene were included. For the low emotional state there were photos of the items that were stolen. For the high emotional states highly emotional photographs of victims were included (see Figure 32 for examples). The photographs were obtained from a standardized set of photos (the Affective Photographic Gallery (Lang, Bradley, & Cuthbert, 1997). These photos have been widely used in research and have been established as evoking emotional states.

To further increase the strength of top-down bias to find a match subliminal messages were introduced. Following established paradigms (Levy, Ashman, & Dror, 2000) 'guilty' and 'same' messages were presented to participants to try and induce them to find a match to convict a suspect.

The experiment was programmed using the experimental software Cedrus Superlab Pro. Participants were tested on an IBM computer with a 17 inch monitor.

All participants were tested in all conditions. There were two levels of stimuli difficulty, (ambiguous vs. unambiguous) and four levels of top-down influence (a control with no emotional influence, low emotion stories and photographs, high emotion stories and photographs, and the highest level of top-down influence that included the high emotion (see figure 32) provoking photographs stories and photographs as well as the subliminal messages. The dependent variables were the number of matches made in each condition, i.e. number of times the participant made a ‘same’ decision.

Figure 32 Examples of low emotion (top panel) and high emotion (bottom panel) (pictures courtesy of Dr Itiel Dror)

Each participant was tested individually and was presented with 96 trials. Each trial contained a pair of fingerprints. For the blocks of trials that included the top-down manipulations, participants were presented with the stories and photographs prior to showing them the pair of fingerprints. For the trials that included subliminal priming, the words ‘guilty’ and ‘same’ were flashed on the screen for 88 ms after the emotional stories and photographs were presented and right before the fingerprints were presented.

For each pair of fingerprints, the participants were asked to decide if the fingerprints in the pair were the same or different, and to respond as quickly as possible. The participants responded by pressing the appropriate keys on the computer keyboard (either the ‘b’ key which was labelled ‘same’ or the ‘n’ key which was labelled ‘different’). Each pair of fingerprints was presented simultaneously, and remained on the computer screen until the participant made a ‘same’ or ‘different’ decision, at which point the next trial appeared on screen.

Initially participants were given practice trials consisting of six pairs of fingerprints, and asked to decide if the prints in the pair were the ‘same’ or ‘different.’ Once the participants responded to a practice trial, the correct answer (i.e. ‘same’ or ‘different’) would appear on the screen before the next pair was displayed.

After the practice trials, the participants began the actual experiment. No feedback was given during the actual experiment. The 96 pairs of fingerprints were divided to four blocks of trials. Each block contained 12 unambiguous pairs (six were a ‘match’ and six were a ‘no match’) and 12 ambiguous pairs of fingerprints. Within each block, the fingerprints were presented in random order. The experimental blocks themselves were not randomized, to avoid emotions crossing between trials whereby highly emotional states will transfer and affect low emotional states and control trials. Thus, it was unwarranted to randomize the order of the blocks, but the trials within the blocks were randomized. To summarize, the first block of 24 trials included all the control decisions that had no top down component. Then the second block of 24 trials included the low emotional manipulation of top-down influence, followed by the third block of 24 trials with the high emotional manipulation. Then finally the fourth block of 24 trials was presented which included the high emotional manipulation along with the subliminal messages of ‘guilty’ and ‘same’.

Results

A two-way analysis of variance (ANOVA) was carried out with Stimuli Type (ambiguous vs. non-ambiguous) and Top-Down Manipulation (control, low emotion, high emotion, and high emotion subliminal) as within variables. Of main interest was a significant interaction found between Stimuli Type and Top-Down Manipulation ($F(3, 78) = 8.172$, $p < 0.001$). This interaction reflected that the Top-Down Manipulation affected decisions on matching fingerprints, but that this effect varied with the different Stimuli Type. There were no significant main effects for Stimuli Type and Top-Down Manipulation ($p > 0.05$) by themselves.

In order to understand better the source of the interaction, each one of the Stimuli Type was subjected to a separate ANOVA. The analysis of the unambiguous fingerprints revealed that participants found matches in 50%, 54%, 51%, and 46% of the time, respectively for the control, low emotion, high emotion, and high emotion subliminal Top-Down Manipulations (and identifying the non-matching pairs in the remaining trials, respectively). Thus participants correctly distinguished between the match and non-match fingerprints. Furthermore, the unambiguous data did not differ across these experimental manipulations, and hence was of no further interest by itself.

When examining the ambiguous Stimuli Type, where there was no objectively correct response, a different picture emerged. A significant effect was found ($F(3, 78) = 6.247$, $p < 0.001$), with matches differing across the experimental manipulations. Participants found a match in 47%, 49%, 58%, and 66% of the trials, respectively for the control, low emotion, high emotion, and high emotion subliminal Top-Down Manipulations. These analyses together reflected the source of the interaction found in the overall analysis; namely, that decisions only varied with the Top-Down Manipulation when judgments were made on the ambiguous Stimuli Type.

In order to determine the nature of the significant differences within the ambiguous Stimuli Type, repeated t-tests were carried out across the four Top-Down Manipulation conditions. The t-tests revealed that the low emotion condition did not affect decisions, as no significant differences were found between decisions made in the control condition and those made in the low emotion Top-Down condition, $p > 0.05$ (47% was comparable to 49%). However, there was a significant difference between the control condition and the high emotion condition ($t = -2.057$, $df = 26$, two-tailed = 0.050), reflecting that participants were more likely to find a match when subjected to the high emotion Top-Down Manipulation (47% vs. 58%). Furthermore, the addition of subliminal messages to the high emotion condition produced even higher levels of matches (66%), which was also reflected in the focused t-test, ($t = -2.687$, $df = 26$, two-tailed = 0.012).

Discussion

The aim of this investigation was to assess the possible effects of top-down processing in interfering with the bottom-up identification of fingerprints. There was a focus on emotion and subliminal messages for the top-down manipulations for three reasons. First, both are present in many of the applied real world forensic settings where fingerprints are matched. In these settings background information is available to the decision makers. This information may have emotional impact (such as the nature of the crime and the victims) and may also include subliminal messages (such as non-verbal biases communicated by colleagues and superiors, as well as additional evidence that points towards the suspect).

Second, past cognitive research has shown that both manipulations may affect decision making. Emotion-congruent effects and subliminal messages have shown in a number of domains (but not in the context of fingerprint matching) that they can alter how we process information, what we see, and our decision making process (e.g. Byrne & Eysenck, 1993; Darley & Gross, 1983; Hirsch & Mathews, 1997; Murphy & Zajonc, 1993). Third, they are relatively easy to control, quantify, and administer in a laboratory condition. Hence, manipulations of emotion and subliminal messages as our top-down conditions seemed to capture a number of important applied, theoretical, and practical considerations.

Using the top-down manipulations it was hoped to examine whether they can affect decisions in matching fingerprints. Furthermore, if they are able to make such an impact, the interest would focus on whether it was possible to see the potential strength and scope of this affect.

The results of the research in this chapter demonstrate that emotion and subliminal messages did influence decision making. Specifically that top-down influences can interfere with people's decisions in matching fingerprints. However, the findings show that this top-down effect is limited in scope and strength. When the fingerprints were a clear match (or no match) then the top-down component was not able to override the bottom-up input information (see Pylyshyn, 1984, for a full discussion of cognitive penetrability).

The findings did, however, show that when the fingerprints to be matched were ambiguous, the top-down component had effects on the decisions being made. Thus, the top-down component was able to bias how gaps are filled but did not have the power to override clear bottom-up incoming information. Top-down components may well be able to override and contradict clear bottom-up information, but this may only occur under very specific circumstances.

With the growing use of technology in fingerprint identification, some claim that such human biases and weakness will be reduced, if not eliminated altogether. Although technology is an important ally in fingerprint matching, the issues addressed in this chapter, as well as other psychological/cognitive issues, will continue to exist and even increase. Automated Fingerprint Identification Systems (AFIS) are able to judge whether a pair of fingerprints match or not. But computers will only be able to make good judgments with confident high levels of accuracy when both prints are of high quality and in very good condition. Human experts, the fingerprint examiners will continue to be needed in the foreseeable future to deal with prints that are partial, distorted, not clear, contaminated, or not-ideal in any respect. Furthermore, the growing use of computer technology in fingerprint matching gives rise to giant databases that contain larger and larger samples of fingerprints (e.g. approaching 10 million in the UK system (Ident 1) and 100 million in the USA system). With such large samples, the relative similarity (and hence difficulty in matching fingerprints) will increase (Ashworth & Dror, 2000). The data reported in this study demonstrates that with greater difficulty in the bottom-up matching of the prints, greater opportunity and vulnerability are created for the top-down contextual components to influence and interfere.

The research reported in this chapter represents the first step in examining how top-down influences may interfere with (or enhance) fingerprint matching. These influences need to be further understood in order to develop better ways to be found to either avoid them (or utilize them). If their existence is denied, rather than acknowledge and study them, then it will not be possible to deal with or manage them appropriately. In this chapter it was important to establish that top-down processing can indeed influence decisions about fingerprint identification. Now that this has been demonstrated further research will need to address two main lines of research. First, a more careful scrutiny of the interaction between top-down and bottom-up information in the domain of fingerprint identification is required. This experiment included the more extreme manipulations and design to see if such influences have any effect on how fingerprints are matched. For example, subliminal messages were combined with emotional state, without scrutinizing the possible effects of subliminal messages on their own.

In the experiment reported in this chapter there was no opportunity for the participants to respond 'cannot decide, 'rather, they had to state either a 'match' or a 'no match'. This, of course, may yield still further findings and have important implications to real world fingerprint matching. What this study did was to demonstrate the existence of an effect in which top-down components interfere with fingerprint identification and future research is needed to further elucidate this effect. Second, the findings need to be examined within the context of routine everyday work of fingerprint experts. The training, experience, and work procedures of fingerprint experts may play an interesting and crucial role in if and how top-down components play a role in fingerprint identification. On the one hand, fingerprint experts may be less susceptible to top-down interference, perhaps even immune, to such effects, though this is unlikely. However, given their highly specialized skills, they may be able to focus solely on the bottom-up component and be data driven without the external influences that has been observed in the research reported here. On the other hand, and in contrast, fingerprint experts may be even more susceptible to such top-down components. Their vast knowledge and experience may provide them with extra degrees of freedom to rationalize and justify what they are biased to find by the top-down components.

Research has demonstrated that professional police officers are susceptible to attentional biases caused by top-down influences as much as novices (Eberhardt, Goff, Purdie, & Davis, 2004). Further research can address these theoretical and applied issues.

Chapter 4:

Contextual information renders experts vulnerable to making erroneous identifications

Being a scientist or forensic expert is rooted in the ability to examine evidence reliably and objectively. To do this, these professionals must be able to dissociate themselves from extraneous contexts and other influences that may interfere with their ability to examine, evaluate, and judge the relevant information. Their decisions should be based on the information relevant to the task at hand and its unbiased interpretation. This involves independent thought that ignores to a large extent extraneous pressures and influences. External pressures and influences are many and varied. The history of science is full of examples of extraneous influences, and today too, scientists work within, and are influenced by, political, economical and other agendas (e.g., global warming, genetically modified crops, and measles mumps rubella vaccine).

Terrorism has brought about a wave of contextual influences. These include, among others, heightened suspicion of ethnic minorities, fear, anger, helplessness, as well as pressure on governments to control (or at least appear to control) such threats. Such contextual influences provide strong and ample opportunities to contaminate objectivity, leading to distortions and errors of judgment beyond the unavoidable. Indeed, within this context we have witnessed major misevaluations and misjudgments by intelligence experts.

Within a similar extraneous context the United States Federal Bureau of Investigation (FBI) positively, but erroneously, identified a Muslim as the Madrid bomber (see figure 33). This incorrect identification was further verified by a number of FBI and other fingerprint experts and led to the arrest of an innocent person. It was only due to rare and exceptional circumstances that this error was revealed and eventually acknowledged by the FBI (Stacey 2004). Errors can occur across forensic science evidence, including DNA (Thompson 1995).

Empirical cognitive research in these areas has been largely neglected (if not ignored), partially because professional expert assessment of evidence (as in the criminal justice system) is believed to be relatively objective.

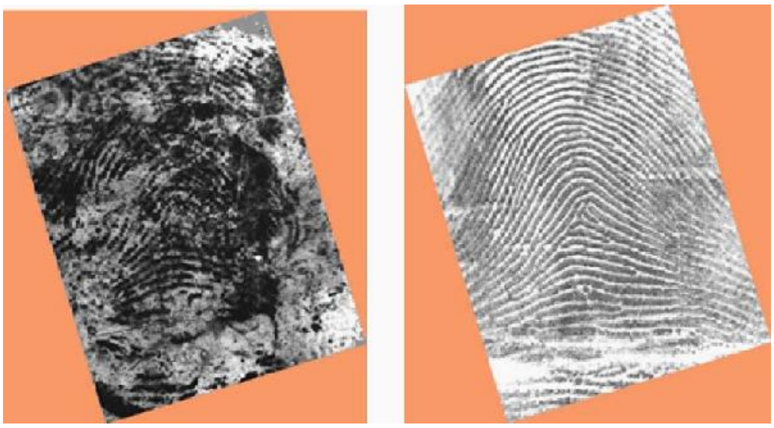


Figure 33 The FBI's erroneous identification of the Madrid bomber. The latent print from the crime scene (left panel) and the fingerprint of the innocent suspect who was positively identified by a number of fingerprint experts (right panel) (courtesy of Kasey Wertheim CLPEX collection).

With the growing number of anecdotal cases that question this belief and suggest that forensic assessment is far from being as objective as it can and should be, it is important to conduct cognitive scientific research in this area. Laboratory experiments performed thus far in this thesis have already suggested that motivational factors, as well as emotional context may bias fingerprint identification. These investigations found that fingerprint examiners are highly motivated to 'get results', and that university students were more likely to judge that there was a positive match between pairs of fingerprints that were presented within an emotional context than those presented within an emotionally neutral control context (Dror et al 2005). However, this vulnerability was apparent only when the prints were ambiguous and lacked clarity. The emotional context had minimal effect when there was a clearly matching pair (or a clearly non-matching pair). The research reported in the last chapter was, however, based on non-experts and conducted in a laboratory setting.

The experiment reported in this chapter presents empirical data on whether real world fingerprint experts in their normal everyday working routines and environment are susceptible to extraneous contextual influences. A within-subject design was employed in which the same experts made judgments on identical pairs of fingerprints, but in different contexts. The aim was to focus on and to examine the contextual influences themselves rather than reveal possible individual differences between experts. Accordingly, I collected and used pairs of fingerprints from archives that the same experts had examined and judged approximately 5 years earlier as a clear and definite match. These previous identification matches were taken from real criminal investigations.

In this experiment, these very same pairs of fingerprints were re-presented to the same experts who had originally evaluated them as a match, but now the examiners were provided the fingerprints within an extraneous context that might bias them to evaluate the prints as a non-match. This was to test whether their decisions were independent and relatively objective, and thus consistent regardless of extraneous influences. Alternatively, if they contradicted their previous decisions, this would demonstrate vulnerability to bias.

Method

Participants

The participants consisted of five fingerprint experts. Together they represented over 85 years of experience in examining fingerprints (mean of 17 years). The participants were taken from an international fingerprint expert pool of volunteers. This pool of participants includes fingerprint experts from a variety of Fingerprint Bureaux, Agencies, and Laboratories from across the world (including the USA, UK, Israel, The Netherlands and Australia). Only experts were used who were not familiar with Brandon Mayfield's fingerprint and from whom we could covertly access past archival identification matches that they themselves had assessed.

Materials

A different pair of fingerprints was prepared for each of the expert participants. Each pair of prints had been previously identified as a match by that same expert in the year 2000, within the normal course of their work. The latent fingerprints had been obtained from the crime scenes and were all presented again to the experts in their original format.

It was further established that all of the pairs of fingerprints were a match by submitting them for verification, 'context free' to two experienced fingerprint experts who were not involved in or aware of our study (each had over 20 years of experience). Both experts independently verified that all five pairs of fingerprints were matches.

Procedure

Participants signed a consent form a few months prior to the experiment. In this form they consented to being tested sometime within the next 12 months without their knowledge. Thus, it was possible to obtain consent but yet test the experts within their normal working environment without them knowing that they were in an experimental situation. Participants were pre-screened and only participants that were not familiar with the fingerprint of Mayfield were used.

Participants were asked by one of their colleagues to examine a set of fingerprints, composed of a latent print (from the crime scene) and a print exemplar (a print obtained from a suspect). They were told that the pair of prints was the one that was erroneously matched by the FBI as the Madrid bomber, thus creating an extraneous context that the prints were a non-match.

The fingerprint experts were asked to decide whether there was sufficient information available in the pair of prints to make a definite and sound decision, and if so, what that judgment was (a match or non-match). They were allowed to evaluate the prints as they would do routinely: handling of the prints, magnifying, lighting equipment, and so forth. The experts were allowed an unlimited amount of time to make their evaluation. The fingerprint experts were further instructed to ignore the context and background information, and to just focus solely on the actual print in their evaluation and decision-making.

Results

Only one participant (20%) judged the prints to be a match (see figure 34), thus making a consistent identification regardless of the extraneous context. The other four participants (80%) changed their identification decision from the original decision they themselves had made five years earlier. Three of these four participants directly contradicted their previous decision and now judged the fingerprints as definite non-matches, whereas, the fourth participant now judged that there was insufficient information to make a definite decision (either a match or a non-match).

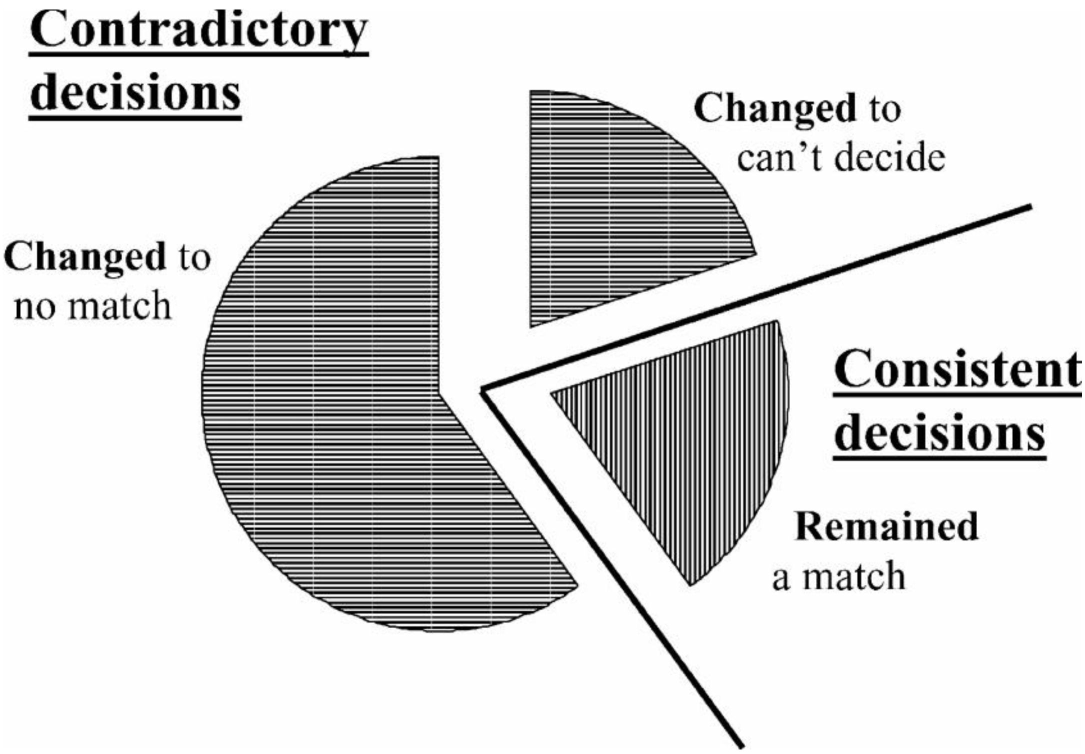


Figure 34 The covert empirical data showing that most of the expert LPE changed their decisions when the same pair of fingerprints was presented in a different context.

Discussion

This experiment shows that fingerprint identification decisions of experts are as vulnerable to irrelevant and misleading contextual influences as were the students tested in the previous chapter. This experiment specifically demonstrated that the extraneous context in which fingerprint examinations occur can determine the identification decision. When presented within a different context four out of five of the experts made different identification decisions. One of the four decided that there was insufficient information available in the latent print to make either a 'match' or 'non-match' decision, whereas, the other three fingerprint experts decided that the fingerprints were a definite 'non-match'. This is striking given that all five experts had seen the identical fingerprints previously and all had decided that the prints were a sound and definite match.

This was the first research to experimentally examine the possible impacts of extraneous context in the real world of biometric and forensic science. One reason for the lack of research in this area is the difficulty in conducting proper scientific research with experts without their knowledge and in their real working environment, while obtaining their consent. Only experts were used for whom it was possible to covertly access and obtain archival files of their own past judgments and who were not familiar with the Mayfield fingerprint. This stipulation further decreased the availability of suitable participants, but had the added advantage of providing a unique opportunity to conduct a within-subject study. The magnitude of the contextual effect and the fact that the experts had judged the same fingerprints in the past enabled the sample to provide clear findings with a high level of confidence. Furthermore, given that the experiment was conducted in real world conditions of the criminal justice system, even if only one expert out of five was susceptible to such effects that would by definition have serious implications.

Even if it were possible to increase the sample of expert participants 10-fold (which is unrealistic, given all the constraints detailed above) and assuming that none of the additional participants would have been vulnerable to the manipulation and changed their judgments (which is statistically highly unlikely), the data would still demonstrate that approximately 10% of the experts were susceptible to misleading extraneous contextual information. Thus, these results are striking even though five expert participants were used and a strong extraneous context employed.

The critical question is what do these results reflect and what do they imply. Are the inconsistent fingerprint identification decisions a reflection of practitioners' errors? Do they reveal deeper methodological and procedural problems in the way that fingerprint experts are trained and identifications are conducted? Or do the results point out basic flaws in the scientific basis and assumptions underlying fingerprint identification altogether?

The data presented in this experiment, along with some of the rare examples where erroneous identifications are publicly revealed and acknowledged, do not necessarily indicate basic flaws in the scientific underpinning of fingerprint identification. The fundamental question as to whether fingerprint identification is a science is not addressed in this chapter or the overall thesis, since that raises a different set of issues that pertain to a variety of ‘‘sciences’’ (Dror et al 2005). These results also do not reflect or reveal practitioners’ errors whereby experts’ negligence, carelessness, and personal fault (intentional or not) produce erroneous identifications. Such causes are often used to deflect deeper scrutiny and discussion. Rather, it seems that these findings of inconsistent identification decisions may reflect cognitive flaws and limitations in conducting objective and independent processing and evaluation of the information. It is important to note that such problems arise mainly in the more difficult and challenging cases, such as with latent fingerprints collected at crime scenes that are distorted, partially missing, and contaminated. In such cases subjectivity is more pronounced (Dror et al 2005).

As extraneous contextual effects are more pronounced, greater distortions can arise. The sources of such distortions are many and varied, including emotional context, pressure, contextual information, group think, biases, hopes and expectations, self fulfilling prophecies, and peer pressure. In this study, a strong misleading extraneous contextual influence was used, but such influences do occur.

It is important first to establish empirically that experts can be influenced by extraneous contexts. Now that it has been established that such an effect is real, further research can and should use different and more subtle manipulations to examine in greater depth when such factors affect performance and render the experts vulnerable to misjudgments, and when such factors are unlikely to affect performance. When vulnerable, these effects can cause a variety of distortions that arise from ignoring parts of the evident information, over-emphasising and over-evaluating other parts of the information, and changing decision criteria, to name but a few.

Vulnerabilities in fingerprint identification may be minimized by better initial selection and screening of fingerprint experts; appropriate training and professional development, and the adoption of methodological procedures that adequately address potential pitfalls. These results show that even in the face of strong extraneous contextual information one expert nevertheless did maintain their original judgment. That expert was able to focus objectively and consistently on the data, ignoring the extraneous misleading contextual information. This demonstrates that it is possible to be much more objective, and that some experts may not be optimizing objectivity.

The reliability and validity of a scientific method such as fingerprint identification is maintained only when analysis is relatively objective, and hence consistent, across individuals, times, and extraneous contexts. For fingerprint examination to remain a credible forensic science, it must achieve this level of objectivity of analysis. This experiment shows that it is possible to alter identification decisions on the same fingerprint, solely by presenting it in a different context. This does not imply that fingerprint and other forensic identifications are not a science, but it does highlight problems of subjectivity, interpretation, and other psychological and cognitive elements that interact and may distort any scientific inquiries (Dror 2010).

One of the main sources of weaknesses in biometric and other forensic sciences is the lack of research, attention, and application of psychological elements that play a key role in the identification processes. These range from the ways in which perceptual factors (such as similarity and orientation) affect the process of pattern recognition (Ashworth et al 2000) to how we consider decision alternatives and shift response criteria (Dror et al 1999). With new and future statistical tools and technologies the face of fingerprint and biometric identification is changing; however, psychology and cognitive elements continue to play a critical role in their implementation and success (Dror 2005). To highlight and address such potential pitfalls, cognitive research needs to be applied systematically to the world of biometrics and forensics. This is all the more necessary in view of the findings that extraneous contextual information is able to determine experts' evaluation of fingerprints. Given that fingerprint examination is a well-established and relatively objective forensic discipline, then distorting effects are undoubtedly as prevalent, if not more so, in other biometrics and forensic disciplines (Risinger et al 2002).

Chapter 5:

Why Experts Make Errors:

Expert performance and accuracy is an important issue in almost all specialised domains. In contrast to novices, experts possess abilities and skills that enable them to perform certain tasks, such as medical procedures or flying aircraft (Dror et al 1993). An expert needs not only knowledge, but also skill, judgment, and experience to evaluate and interpret information correctly to make correct decisions. However, being an expert does not necessarily mean error-free performance; in fact, almost every specialist domain is subject to error. The pivotal question is the source of the error. Errors, broadly speaking, fall into three categories.

The first category relates to human error. Human errors can be intentional errors (whereby experts are involved in fraudulent behavior), negligent errors (whereby experts do not pay attention, do not follow procedures, etc.), and competency errors (whereby experts are unable to make correct judgments because of a lack of appropriate skill sets; this can be due to declining eye sight, faulty initial selection that results in recruiting people who do not possess the proper cognitive abilities that are needed to underpin the expertise, inadequate training, and so forth).

The second category relates to instrumentation and technological errors. In this category, errors derive from failure and breakdown of instruments and technology. These types of errors are rare in the fingerprint domain. Technical malfunction certainly accounts for errors in other domains, such as breakdown of equipment on aircraft. The first and second categories both relate to chance malfunctions and breakdown, either human or machine.

The third category of error relates to more fundamental methodological factors that are inherent to the field in question. These may include errors associated with the technology in question, instrumentation, and measurements. In the fingerprint domain, for example, a failure of the Automated Fingerprint Identification System algorithms to provide a matching file from the database (assuming such a file is present and enough information exists to make the match) can lead to an error.

This, in contrast to category two, reflects not a malfunction or a bug in the software, but the inherent inaccuracy of the algorithm. Technologies and instrumentations have their limits, range of accuracy, levels of precision, variations, and so forth, which are not due to their breakdown and malfunction but to their very nature.

An error is sometimes defined only as an incorrect individualization (i.e., a false positive, what is referred to in signal detection theory as a false alarm). However, such a definition is limited because it does not include cases where there is sufficient information to make a positive individualization, but, because of the error, no such individualization is made (i.e., a false negative, what is referred to in signal detection theory as a miss). These two types of errors are different, but nevertheless, both can be regarded as errors.

However, such errors are not limited to technology or instrumentation, especially in a domain like fingerprints where much of the individualization process falls on human experts and their interaction with technology (Rudin et al 2005, Dror et al 2005). Here, the error is due not to the nature of the technology and instrumentation but to the nature and mechanisms of the human mind and cognition. This is particularly noticeable when dealing with latent prints that are collected from crime scenes and are thus degraded, contaminated, partially missing, and distorted.

This third category thus includes errors that are not simple practitioner error that can be attributed to the specific expert involved (as specified and belonging to category one). When practitioners are competent, well trained, and following procedures, when instrumentation and technology operate properly, and errors happen nevertheless, then these errors belong to category three.

This chapter examines the possible role that psychological and cognitive factors may have in causing these types of errors. When expert practitioners perform well and technology is effective, can errors still occur? And if so why and how? Some people attempt to dismiss this possibillity a priori, claiming that an error results either from a practitioner's error (such as those specified in category one) or from the lack of scientific basis for the domain (such as the uniqueness of fingerprints).

This dichotomized attribution of an error as either reflecting a basic scientific flaw in the domain or a specific practitioner's error fails to consider a third alternative: errors that derive from psychological and cognitive elements involved in fingerprint individualization. This process falls on human experts, allowing the possibility that errors may result from the way the brain processes information and makes decisions.

If the nature of the mind and cognitive processing can give rise to error in fingerprint individualization, then these errors are inherent to the domain. Nevertheless, they do not reflect a basic ontological scientific flaw in the domain nor are they the fault of a specific practitioner. They are, in essence, epistemological problems that derive from the mechanisms of human cognition and the workings of the mind.

As with technological and instrumentation advances that improve their limits, accuracy, and levels of precision, so can human performance be improved with correct selection, training, and procedures. However, such endeavors need to be based on systematic and scientific research, and even then they will not totally eliminate human error of category three. Nevertheless, with such research, these errors can be drastically minimized, so minimized that although they are theoretically possible, they are in fact so very rare that de facto they do not exist.

The above division in to three categories of error is a simplistic model for methodological reasons; reality can be more complex. For example, because practitioners and methodology are so intertwined, it is difficult (perhaps sometimes impossible) to separate and distinguish between the two (Rudin et al 2005), thus making it problematic to attribute an error solely to either category one or three. For instance, if errors occur because a practitioner is incompetent, but his or her incompetence is due to basic flaws in the domain, then the errors are not purely in the practitioner error of category one. Practitioners' incompetence may arise from the lack of appropriate scientifically based screening tests for recruitment and certification of fingerprint examiners and thus may reflect deeper flaws in the domain. To date, there is no systematic scientific research into the psychological and cognitive skills that underlie fingerprint expertise. Research is needed to construct appropriate tests for recruitment screening and selection of fingerprint examiners. Therefore, in some cases, an expert's incompetence may also be attributed to more basic flaws in the domain rather than purely to the individual practitioner.

Understanding the source of errors and their assignment to one of the categories can be highly insightful and have important implications. For example, the Mayfield erroneous individualization (Stacey 2004) raises interesting issues. Would such a mistake be totally attributable to practitioner error? Or should the error be attributed also, at least in part, to the lack of appropriate procedures, training, and quality assurance to address and deal with the causes of such errors? These types of questions are critically important to investigate to allow advances in this and other domains of forensic science.

However, researching and discussing errors is a problematic and challenging endeavor. First, by its very nature, "error" is a sensitive issue that often meets defensive responses. This is especially true in the criminal justice system which deals with incarcerating and even executing people. Second, the framework of the criminal justice system does not enable the "ground truth" to be positively known. Thus, individualization cases can always be open to suspicions as being erroneous. Simon Cole (Cole 2005) tries to catalogue such suspected erroneous individualizations. Third, even when errors are detected and acknowledged, their classification and examination are constructed post hoc in a highly political and personal environment.

Information presented in this chapter examines whether inherent psychological and cognitive mechanisms predispose fingerprint and other forensic identification experts to commit category three errors (Risinger et al 2002, Dror 2010). This important area of research has been highly neglected. Some of the psychological phenomena and mechanisms that may underlie such errors are derived from the nature and architecture of the human mind (Dror 2005). It is imperative to conduct empirical experimental research to examine whether such errors actually do exist and, if so, to find ways to minimize these errors.

To examine such issues, it is important to conduct scientific studies within a real world setting. This is challenging because when people know they are being studied, their behavior and performance change, and thus puts into question the applicability and ecological validity of the findings. One must try to observe and examine performance as well as collect data in the normal routine setting with minimal (or no) knowledge of the people involved. This of course is very difficult to achieve and necessarily results in small data sets. However, these data sets are statistically very powerful, meaningful, and more interpretable.

Previous Research

In previous empirical research in this thesis it has been possible to start to address issues relating to errors that derive from psychological and cognitive influences. Emotional states (See figure 35) have been shown to cause non experts to be more likely to match ambiguous pairs of fingerprints when they performed the comparison in a highly emotional context (Dror et al 2005).

Figure 35 An example of an emotional context used in the previous (non-expert) study. (Collection of Dr Dror)

In a previous experiment five experts were shown fingerprints and were told that the prints were from a highly publicised erroneous identification suggesting that the fingerprints in front of them were an exclusion. However, rather than giving them prints that were an exclusion, they were presented with fingerprints that had been compared and individualized. The fingerprints were not only individualizations, but they had also been previously individualized by the same experts now being tested. Although the experts were instructed to ignore all the contextual information and to focus solely on the actual prints, most of the experts (four of the five) were affected by the context and made inconsistent decisions. In previous research, it was possible to demonstrate the vulnerability of experts to extraneous contextual information, but only when subjected to the relatively extreme context and when presented with difficult matches.

The experiment reported in this chapter follows up on previous investigations, further examining psychological and cognitive influences that may play a role in the work of fingerprint experts. The specific purpose of this experiment was:

- To determine whether the original findings would replicate to another and larger set of data.
- To use less extreme and more routine day- to - day contextual influences.
- To use pairs of prints of varying levels of difficulty.
- To examine the possible influence of contextual information on different decision types (thus, not only whether it can change a past individualization decision to an exclusion, but also to examine whether it can change a past exclusion decision to an individualization).
- To examine the basic consistency of decisions by representing to experts the same fingerprints they judged in the past but without introducing any contextual information manipulation, thus examining the reliability of experts.

Method

Participants

Six fingerprint experts, representing more than 35 years of experience in examining fingerprints (each with a minimum of 5 years' experience in latent prints), participated. The experts were experienced and specialized in latent print examination and were not field operatives; hence, their experience was full - time in latent print comparisons. Each of the expert participants was not only highly experienced, but was highly trained, certified by a nationally recognized independent authority, and had successfully completed proficiency testing. None of them had been the subject of a poor competency review, and they were all considered by their respective laboratory directors or bureau chiefs to be effective and competent.

The participants were taken from an international fingerprint expert pool of volunteers. This pool includes fingerprint experts from across the world (including the USA, UK, Israel, the Netherlands and Australia).

The only experts used were those where past work could be covertly accessed. Because the collation of individual casework is so difficult in such circumstances, and because of the difficulties in covert testing, this effectively limited the number of participants within this study. This permitted, however, the possibility of collecting more meaningful and powerful data. This derives not only from the covert nature of the study but also from the use of a within- subject experimental design in which participants are used as their own control through repeated measures (Rosenthal et al 1991). Thus, variations are more easily attributable to the experimental manipulations and conditions rather than individual differences between people. This provides clearer and more interpretable results, as well as more statistical power for each data point (Rosenthal et al 1991).

Materials

A different and unique set of eight pairs of fingerprints was prepared and tailored for each of the participants. Each set included four pairs of prints that the specific fingerprint expert had in the past judged as individualizations and four pairs of prints that he or she had judged in the past as exclusions. All of the eight pairs of fingerprints had been deemed in the past by the specific participant to have sufficient information to make definite judgments. Within each of the four past individualizations and the four exclusions, two pairs of prints were relatively difficult to judge, and the remaining two pairs were relatively not difficult to judge.

The latent fingerprints had all been obtained from real crime scenes and were all presented to the participants in their original format for comparison against suspect tenprint exemplars. It should be stressed that all the pairs of prints that were used were fingerprint comparisons that were obtained from the archives and had been evaluated some years before by the very same experts. The within- subject design of the experiment was deemed vital to the overall robustness and credibility of the findings. Two additional experienced fingerprint experts who did not take part in the study (each had more than 20 years of experience) independently confirmed and verified that all the pairs of fingerprints were indeed either correct individualizations or correct exclusions. They also characterized the pairs of prints as either relatively difficult to evaluate or as relatively not difficult to evaluate.

An instruction and response sheet was prepared for each of the eight pairs of fingerprints. Four pairs of prints were used as controls and provided no contextual information. The second four pairs of control prints included two pairs that had been judged in the past as individualizations and two pairs that had been judged in the past as exclusions. Of each of the set two pairs, one was relatively difficult to judge and one was relatively not difficult to judge. The instruction sheet that was given with the control pairs included minimal instructions that told the experts to evaluate the prints. A response sheet was provided for participants to write their conclusions. In the response sheet, the experts were first asked whether there was "sufficient information within the prints to either identify or exclude". If the answer was, "no", then they finished with this pair of prints and moved onto the next pair. If the answer was, "yes", then the experts continued, stated what their decisions were, and then moved on to the next pair of prints.

Of the remaining four pairs of prints, two pairs had been judged in the past as individualizations and two had been judged in the past as exclusions. These four pairs of prints were presented within contextually biasing information that was hypothesized to influence the conclusions reached by the experts. The two pairs of prints that had been judged in the past as individualizations were presented in a context that suggested that they were exclusions (one of the pairs was relatively difficult to individualize, and the other was not). Similarly, the two pairs of prints that had been judged in the past as exclusions were presented in a context that suggested that they were individualizations. (Again, one of the pairs was relatively difficult to individualize, and the other was not.)

In contrast to previous experiments, a more subtle routine was now used, using more day-to-day, contextually biasing information. In the instructions the participants were told, for example, that the "suspect confessed to the crime" (for contextual information that the prints were an individualization, when in fact they were not and had been judged by the same expert as an exclusion in the past) or that the "suspect was in police custody at the time of the crime" (for contextual information that the prints were not an individualization, when in fact they were and had been judged by the same expert as an individualization in the past).

Thus, a total 48 unique experimental trials were prepared. Each one consisted of a latent print from a crime scene and a suspect tenprint exemplar, accompanied with the proper instructions and response sheets. For each expert, a customized folder was prepared containing the eight pairs of fingerprints that they themselves had judged in the past. The eight pairs of fingerprints in each folder were counterbalanced. Counterbalance presentations assure that results are not due to order affects and cross-contamination between the different conditions, because the order of presentation is systematically permutated across participants (Rosenthal et al 1991).

Procedure

Participants were approached by the director or head of the laboratory or bureau and were asked to provide opinions on a variety of latent prints and their comparisons to ten-print exemplars. They were told that the conclusions they reached after the examination would be used for an assessment project. They were further told that the project was intended to look at problematic prints and assessments.

Then, the first assessment, along with the instruction and response sheet, was given to the participants. After they finished the comparison and documented their conclusions on the response sheet, the materials were put back into the original folder, and the second assessment was given to them. This continued until all eight assessments were completed.

During the comparison process, all of the participants were allowed to evaluate the prints as they would do routinely (handling the prints, using magnifying and lighting equipment, and so forth). The participants were allowed an unlimited amount of time and all normal resources (e.g., comparators) to make their evaluations.

Results

Overall, from 48 experimental trials, the fingerprint experts changed their past decisions on six pairs of fingerprints (see table 12). The six inconsistent decisions (12%) included the 24 control trials that did not have any contextual manipulation. From the 24 experimental trials that included the contextual manipulation, the fingerprint experts changed four of their past decisions, thus making 16.6% inconsistent decisions that were due to biasing context. The inconsistent decisions were spread between the participants. (The inconsistent decisions were by four of the six experts, but one expert made three inconsistent decisions while each of the other three made only one inconsistent decision.) Only one-third of the participants (two out of six) remained entirely consistent across the eight experimental trials.

Table 12 Results of fingerprint examinations.

	1	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Past Decision	Identified	Identified	Identified	Identified	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 Identified	None	Suggest Identified
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 Identified	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

A further examination of the inconsistent decisions revealed that they were most common in the more difficult decisions. In most cases (five out of six), the inconsistent decisions were made in the difficult comparisons. Nevertheless, inconsistent decisions were not totally limited to the difficult comparisons; one of the six was in the relatively not difficult comparisons.

The inconsistent decisions were most prevalent in past individualization decisions (five out of the six inconsistent decisions; they were changed to either exclusion, in four cases, or to cannot decide, in one case). However, a conflicting judgment also occurred in a past exclusion decision (now changed to an individualization).

Finally, two inconsistent decisions were present in the control condition that had no biasing contextual information. These decisions constituted 33.3% of the conflicting data (two out of six) and 8.3% of the control data (two out of 24).

Discussion

Experts, just like all humans, are bound by the way the mind and the brain work. One of the most notable characteristics of humans is the active and dynamic nature in which we process information (Dror 2005). This fundamental cognitive architecture enables us to deal with vast amounts of information and is the basis of human intelligence. However, it can also affect and distort what we see (or why and what we do not see), how we assess and evaluate visual information, and our decision making (Dror 2010, Dror et al 1999).

Fingerprint and other forensic experts are not immune to such psychological and cognitive factors. Researching and understanding them better is a necessary step toward dealing with and minimizing such influences.

Previous research has already demonstrated that under conditions of relatively extreme and rare extraneous contextual information, fingerprint experts may change the way they compare and judge fingerprints (Dror et al 2006). The extraneous contextual information effect caused some experts to reach different and inconsistent decisions to those they had made in the past on the very same pair of fingerprints.

The research reported in this chapter replicated and expanded on previous findings and found that across eight comparisons made by each of the six participants, two thirds of the fingerprint experts made inconsistent decisions to those they had made in the past on the same pairs of prints. The findings of this experiment not only further substantiate the vulnerability of experts to contextual effects within a larger data set, but they further contribute to understanding of this phenomenon.

The data demonstrate that fingerprint experts were vulnerable to biasing information when they were presented within relatively routine day-to-day contexts, such as corroborative (or conflicting) evidence of confession to the crime. Thus, contextual information does not need to be extreme and unique to influence experts in their fingerprint examination and judgment.

Varying the levels of difficulty in comparing the prints demonstrates that psychological and cognitive vulnerabilities are most pronounced in the difficult cases. However, the data also show that such vulnerabilities can also occur and cloud judgment in non difficult cases, because the contradictory findings were not limited to only difficult comparisons.

Previous research has only examined whether past decisions are susceptible to change when the past decisions were individualizations (Dror et al 2006). It seems that the threshold to make a decision of exclusion is lower than that to make a decision of individualization. Indeed the data support this claim, as reflected by the fact that most of the conflicting decisions were past individualizations. It was observed, however, a case in which an exclusion decision was now judged to be individualization. This relates to the decision-making model used by experts in the fingerprint domain. Changes in decisions may reflect changes in decision thresholds or changes of the decision strategy itself. The former reflects changes in decision criteria within a single strategy whereas the latter reflects modifying (or totally abandoning) the decision strategy and replacing it with a different strategy. Examining the decisions and how they change can reveal which occurred; however, this needs to be done carefully, because changes in thresholds and decision strategies can yield similar (and even identical) decision changes (Dror et al 1999).

Finally, quite surprising and alarming are the data points of inconsistent decisions made in the control “context-free” condition. This may reflect that expert decisions are inconsistent across time regardless of context. If this is indeed the case, then further research needs to examine the root of this observation. Is this due to the accumulation of expertise and experience gained between the first and second exposures (and this led to change in decision threshold or in decision strategy itself), or is some other explanation the cause? Such inconsistency does not suggest that contextual information does not affect the judgment of fingerprint experts; it only suggests that contextual information is not the only factor that may affect fingerprint experts. The control condition needs to be evaluated with a number of reservations. First, the control condition was not context free. It did not include any extraneous contextually biasing information; however, that does not make it context free. Second, even if it were possible to achieve this ideal notion of a context-free environment, this context would still be different than that which was present during the first exposure and judgment made years ago.

These findings are especially robust because a within-subject experimental design was employed. Thus these findings do not reflect individual differences among experts. These findings encourage particular confidence because these findings support the findings of previous investigations and thus the extra data further substantiate and validate conclusions. This study was conducted covertly, which is critical for the correct measure of performance. When participants know they are participating in a study, their behavior changes (Dror 2010). These experimental design criteria make it hard to collect data and enable only small data sets, but the data are meaningful and statistically powerful (Rosenthal et al 1991 and Dror et al 2006).

This entire area of research is new in the forensic sciences and has rarely been considered before. Therefore, such studies constitute the initiation of a research program that is aimed at examining the psychological and cognitive elements that are involved in fingerprint and other forensic identifications. It is hoped that these findings will contribute to better selection, better training, and better procedures for work in this domain. However, such an endeavor to deal with and minimize these vulnerabilities is dependent on the cooperation of fingerprint experts worldwide.

Further research is needed in a wide range of issues pertaining to individualization. Within the issues raised in this chapter, additional data may shed light on the characteristics of the experts who were immune to contextual manipulations, issues pertaining to the circumstances in which the manipulations were more (or less) effective, as well as additional issues.

Research needs to examine psychological and cognitive influences in all the stages that lead to decisions in comparing fingerprints, from feature perception and selection, to evaluation and comparison, to the final interpretation and weighting of alternative choices that determine the decision outcome. Additional perspectives for future research in the fingerprint and other forensic domains relate to verification, selection and training of experts, and integration of technology.

There is no possibility of 100% objectivity (Dror 2010, Dror 2005 and Hofstadter et al 1979), but there is potential for very high levels of objectivity. How rare and under what conditions errors occur at a practical level is still unclear at this stage. Experts, as humans, are prone to errors; however, with proper research and its systematic application, these errors can be reduced and minimized.

Chapter 6:

Cognitive issues in fingerprint analysis: Inter- and intra-expert consistency and the effect of a 'target' comparison

Cognitive processes underpin much of the work carried out in many forensic disciplines which require examination of visual images. Fingerprints, in addition to bite and shoe impressions, tire tracks, firearms, hair, hand writing and other forensic domains all hinge on comparative examination involving visual recognition of visual cues, patterns or striations. Although human experts are the 'instrument' in judging whether two patterns originate from the same source, understanding the factors that shape such judgments in forensic science has been relatively neglected. Often it has been misconceived that fingerprint identification is an exact science (Evett & Williams, 1996); and this perception goes across all forensic disciplines, (Evett 1996). The recent National Academy of Sciences report further highlights that the findings of cognitive psychology and the extent to which practitioners in a particular forensic discipline rely on human interpretation are significant (NAS, 2009), and that at least to date, there is no evidence to indicate that the forensic science community has made a sufficient effort to address the bias issue.

The task demands imposed on the examiners require them to search through a rich stimulus, filter out noise, and determine characteristics and 'signals' for comparison (see Phillips, Saks, & Peterson, 2001, and Vokey, Tangen, & Cole, 2009, for discussion of signal detection theory (SDT) applied to fingerprint evidence). This initial analysis and determination of 'signals' usually takes place before the actual comparison between stimuli (e.g., the latent fingerprint mark left at a crime scene and the comparison print of a suspect). Scientists have long accepted that observations, including those in their own scientific research, encompass errors.

A study examining 140,000 scientific observations reported in published research not only revealed that erroneous observations were made, but that those were systematically biased in favour of the hypothesis being researched (Rosenthal, 1978). For many years, laboratory technicians who counted blood cells visually were taught that correct counting would keep blood cell counts within a certain range of variation. Using a more accurate photographic method to count blood cells, researchers discovered that for years technicians had been reporting blood cell counts that were within an impossibly narrow band of variability. The technicians made observations consistent with the expectations they held, but inconsistent with reality. Many different forms of contextual and cognitive influences affect our perception and bias it in a variety of ways (Nickerson, 1998).

Previous research specifically examined the potential cognitive contextual influences impacting upon the reliability of comparing fingerprints and decision making when related to whether or not they originated from the same source (Dror & Charlton, 2006; Dror, Charlton & Péron, 2006; Dror & Mnookin, 2010; Dror, Péron, Hind & Charlton, 2005; Dror & Rosenthal, 2008; Hall & Player, 2008; Langenburg, 2009; Langenburg, Champod & Wertheim, 2009; Risinger, Saks, Thompson, & Rosenthal, 2002; Saks, 2009; Stacey, 2004; Wertheim, Langenburg, & Moenssens, 2006; Evett and Williams, 1996). For a review of the literature on bias in forensic decision making, see Dror and Cole (2010).

When specifically focused on the initial analysis phase of the latent finger mark, before comparison to a ten print exemplar is initiated, Langenburg (Langenburg 2004) found that examiners generally reported observing more minutiae than novice controls. Furthermore, although the examiners varied in how many minutiae they observed, they were more consistent than the novice control group (in 8 out of the 10 latent marks used in the Langenburg experiment). These results were consistent with those of Evett and Williams (Evett et al 1996). Following Langenburg's 2004 experiment, Schiffer and Champod (Schiffer et al 2007) found that training and experience increased the number of characteristics reported, and at the same time reduced the variability among observers. Schiffer and Champod (2007) also reported that the number of characteristics observed during the analysis phase was not affected by contextual information about the case or by the presence of a comparison ten-print. Consequently, they concluded that the initial analysis stage (pre-comparison) is relatively robust and relatively free from the risk of contamination through contextualisation of the process.

Although Langenburg (2004) Schiffer and Champod (2007) show that these inconsistencies decrease with training and experience, they also make the point that there are important variations between examiners (Schiffer and Champod, 2007). These investigations consistently show that there is variability in the number of minutiae observed in the analysis stage, but that these inconsistencies are attenuated but not eliminated as a result of training and experience gained operationally in fingerprint examination.

As reported by Schiffer and Champod (2007), even in the relatively robust stage of analysis a clear subjective element persists. A further study (Schiffer 2009) suggests that the combined presence of contextual pressure and availability of the target comparison print influences the evaluation stage (following the analysis and comparison), but this affect varies among different marks (the reported features in correspondence obtained following the ACE process). It is suggested that as finger marks are more difficult (bottom-up), the more influence external factors (top-down) have on the observations (Dror et al., 2005). Bottom-up refers to the incoming data, where as top-down relies on pre-existing knowledge (Busey & Dror 2010). Top-down has many forms and manifestations, which include the context in which the data is presented, past experiences and knowledge, expectations, motivations and so forth. Expertise is a manifestation of top-down information processing, in that expert knowledge makes certain assumptions based on experience of ridge flow in a fingerprint, the configuration and reliability of certain feature types. Expertise is a form of cognitive shorthand that allows the mind to process huge amounts of information efficiently and effectively and as such experts rely more on top-down information to facilitate faster information processing based on experiential assumptions. This efficient and effective processing of the bottom-up data can distort and bias how the data is processed (Dror, 2010). However, Schiffer (2009) also showed that the availability of the comparison print does not influence the observation of minutiae in the analysis stage. The analysis stage seems to be relatively robust against "circular reasoning" if the ACE process is applied in sequence and each step is properly documented.

Furthermore, variations in observation amongst different observers ("inter-observer" differences) and variations in observation for the same observer for the same task, taken at different time ("intra-observer" differences) are a well-known phenomenon in other fields involving expert decisions, such as radiologists or other medical technicians (Potchen, 2000; Bektas, 2009). A general finding in these medical studies is that intra-observer variability tends to be lower than inter-observer variability.

In the research reported in this chapter three main issues were investigated:

- The potential effect that a 'target' comparison fingerprint may have on the analysis of the latent mark.
- The consistency in analysis among different examiners.
- The consistency in analysis within the same examiner.

This chapter further investigates and contributes to the studies on the analysis of fingerprints in the following ways:

- Using actual latent fingerprint examiners, rather than forensic science or psychology students (such as in Schiffer, 2009; Dror et al., 2005).

- Applying a within-subject (intra-observer) experimental design. This allows us to measure consistency in analysis, as examiners are compared to themselves, in addition to comparing between examiners. Such intra-observer measurements are useful because they are not only statistically more powerful then inter-observer measures, but they allow us to confidently draw conclusions because the data cannot be attributed to individual differences, such as visual acuity, experience, strategy, cognitive style, and training.

- Subjecting the experimental data to statistical procedures and standards (retest reliability) that quantify the consistency of latent fingerprint examiners in analysis of latent marks.

- Statistically differentiating between factors that contribute to inconsistencies in latent mark analysis; thus determining what portion of the variance is attributed to the examiners' performance and what portion is attributed to the latent marks themselves (using statistical effect sizes).

- Suggesting a number of recommendations for handling and dealing with issues surrounding latent mark analysis.

Effects of a 'Target' Comparison on Examiner Visual Perception

The human cognitive system is limited in its capacity to process information. The information available far exceeds available brain power and cognitive resource, and therefore we can only process a fraction of the information presented to us. This mismatch between computational demands and available cognitive resources provided the catalyst for the development of cognitive mechanisms that underpin human intelligence. For example, we do not randomly process information, but rather prioritize it according to our expectations (Summerfield & Egner, 2009). Expectations are derived from experience, motivation, context, and other top-down cognitive processes that guide visual search, allocation of attention, filtering of information, and what (and how) information is processed. These mechanisms are vital for cognitive processes to be successful. Expertise is characterised by further development and enhancement of such mechanisms (Busey & Dror 2010; Dror 2010; Ericsson, et. al, 2009; Ericsson, 2006).

There is good scientific data showing that the presence of any contextual information may affect cognitive information processing. Various factors and specific parameters define the context, who it may affect, how, and to what extent such context influences the processing of information. Understanding these factors and the parameters of their sphere of influence will help develop science-based training and procedures that will enhance objectivity in fingerprint and other friction ridge skin analyses, as well as other forensic comparative examination involving visual recognition.

In the first experiment reported in this chapter 20 experienced latent fingerprint examiners were used to examine whether the presence of a comparison 'target' print would affect the characteristics they observe in the latent mark. Each of the 20 experts received ten stimuli: five latent marks by themselves (solo condition) and five latent marks with the matching target print (pair condition). All the participants were instructed identically, requiring them to examine the latent marks and to count all the minutiae present in the image. The experimental conditions were counterbalanced across participants using a Latin Square design to minimize any affects due to the order of presenting the experimental trials (Rosenthal & Rosnow, 2007).

Results and Discussion

This experiment found that the presence of the accompanying comparison print affected how many minutiae were perceived by the expert latent print examiners. These differences were statistically significant ($t_{(9)} = 2.38$, $p = .021$; with an effect size, $r = .62$). As evident in Table 13, the presence of the accompanying matching comparison print mainly reduced the number of minutiae perceived. This is consistent with attention guided visual search, whereby our cognitive system operates within the contextual expectation. It is important to note that the reduced number of minutiae was perhaps due to the comparison print being from the same source (a match); if it had been a non-match, then it may have directed the perceptual cognitive system differently, possibly providing different results. The importance of the finding is not whether the presence of the comparison print reduced or increased the number of minutiae perceived in the latent mark, but that the presence of a target comparison print had an effect on the perception and judgment of the latent mark.

Table 13 The mean number of minutiae observed when the latent mark was presented within the context of comparison print ('pair'), by itself ('solo'), and the differences between these two conditions.

CASE	PAIR	SOLO	DIFFERENCE
A	14.1	20.6	-6.5
B	9.9	13.4	-3.5
C	10.8	20.1	-9.3
D	9.7	9.8	-0.1
E	11.1	10.7	0.4
F	8.8	8.4	0.4
G	10.7	12.1	-1.4
H	10.5	15.6	-5.1
I	8.5	7.1	1.4
J	6.6	9.1	-2.5
MEAN	10.1	12.7	-2.6
SD	2.0	4.7	3.5

This emphasises the importance of examining the latent mark in isolation (see table 13), prior to being exposed to any potential comparison print. This is to maximize the 'clean' bottom-up and more objective analysis, driven by the actual latent mark, and to minimize any external influences that may bias the process of analysing the latent mark itself. This is especially important when the latent print is of low quality. Such recommendations are also appropriate for other forensic domains (e.g., DNA, see sequential unmasking, Krane et al., 2008), as well as for scientific research in general: "Keep the processes of data collection and analysis as blind as possible for as long as possible" (Rosenthal, 1978, p. 1007).

However, the comparison print can play an important role in helping examiners optimize their analysis by correctly guiding their cognitive resources and interpretation (Dror, 2009). Therefore it seems reasonable to balance the vulnerabilities and clues presented by making the comparison print available to the examiner by instigating the introduction of checks and balances to mitigate against the introduction of such contextual information. A reasonable solution may be to first examine the latent mark in isolation, clearly documenting this more objective and uninfluenced analysis, but at the same time also allowing further analysis to be conducted later after exposure to the context of the target comparison print. Hence, the ACE (Assess, Compare and Evaluate) approach needs to be applied linearly initially, making sure that the initial Analysis of the latent mark is done in isolation, prior to moving to Comparison and Evaluation; yet still allowing flexibility, with well documented transparency of when and why this took place.

This methodology should be part of a wider strategy that controls and limits the circumstances and extent for such retroactive changes so as to maximize performance but avoid (or at least minimize) circular reasoning and bias (for details, see Dror, 2009). An attempt to provide such flexible and transparent practices have been captured by the PiAnoS (Picture Annotation System) software developed by Champod, et al. at the University of Lausanne and by Langenburg's GYRO system of annotation for fingerprint comparisons (Langenburg, 2009).

It is important to note that some latent marks were more susceptible to this effect than others. For example, Table 13 shows that latent D was unaffected by the presence of the comparison print, whereas latent B was quite dramatically affected (see Figure 36, below, for the actual latent marks). It is clear from all the studies on latent mark analysis that findings are highly dependent on the specific fingerprints used. This suggests that we can (and probably should) tailor procedures and best practices to specific types of prints, rather than inflexibly applying identical procedures prescribed to all prints (Dror, 2009). Such knowledge-based procedures will allow for higher quality work without requiring more resources, because it wisely and appropriately allocates resources to where they are needed.

Figure 36 Some latent marks were more affected by the presence of a target comparison print than other latent marks. For example, latent mark B (left panel) was more affected then latent mark D (right panel) (personal collection).

The large variability in the effects of the presence of the comparison print (see table 13) on the latent mark analysis may explain why Schiffer and Champod (2007) did not find such an effect: The latent marks they used may have been prints that are less (or not at all) affected by the presence of the comparison print, such as mark D in this study. An alternative (not mutually exclusive) explanation of why Schiffer and Champod (2007) did not find this effect is that these effects may occur as examiners are more experienced and knowledgeable, and hence have expertise in how to utilise the information from the comparison print more effectively. The study reported here used experienced experts in latent print examination whereas Schiffer and Champod (2007) used students.

It is also noteworthy that the largest differences were observed with the latent marks that had the highest number of minutiae observed in the solo condition (see e.g., A, C, & H in table 13). Overall, the correlation between the number of minutiae observed in the solo condition and the difference (absolute value) from those observed when shown in the pair condition was 0.9 (see Figure 37). This may be due just to a ceiling effect, i.e., an artifact reflecting that as there are more minutiae marked in the solo condition, then there is more scope to reduce this number in the pair condition, and as the number of minutiae are lower in the solo condition, there is much less scope for a drop when they are presented in the pair condition.

An alternative, not mutually exclusive, explanation is the affect of motivational factors. In the solo condition, examiners may be motivated to mark as many minutiae as they can, as they are not sure which ones may be useful and informative when they have a target print at the comparison stage. However, when the latent is analysed while the target comparison print is available (as in the pair condition), examiners' motivation may drop when they get to a critical mass of minutiae they need for comparison purposes. Once they get to that threshold, they may be less likely to detect more minutiae. This effect further strengthens the suggestion that the initial analysis of latent marks should be done in isolation of a comparison exemplar print (especially when the latent mark is judged to be low quality, distorted, or has limited information available).

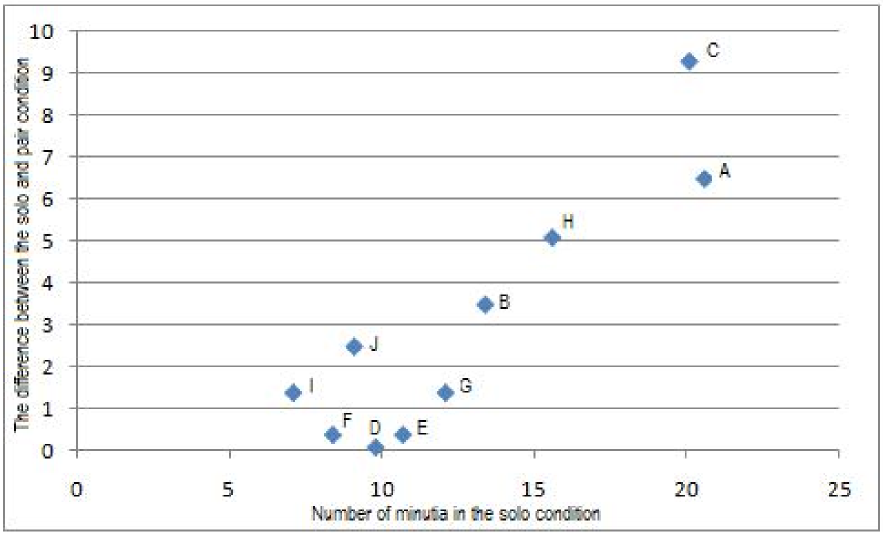


Figure 37 The high correlation (.9) between the mean number of minutiae observed when analysis was conducted when latent mark was presented by itself ('solo' condition) and the differences in analysis between the solo and pair conditions.

Inter-Observer Consistency

The presence of a 'target' comparison print can affect the perception and judgment of the latent mark in a number of ways. The next issue investigated was the consistency in the perception and judgment of minutiae in a latent mark across participants, even without the presence of a target comparison. The 'solo' condition data contains the answer to this question; it allows us to examine and compare the minutiae observed by different experts, and hence to report the variability in how latent print examiners may perceive and judge minutiae. Table 14 presents the relevant data, with the range of values for each mark (bottom row). The lack of consistency may reflect the absence of objective and quantifiable measures as to what constitutes a ridge feature, especially with latent marks that are of varying quality. However, these differences may also reflect individual differences between the examiners (arising from variations in eyesight, training, feature selection strategy, cognitive style, threshold criteria, etc.; these issues are currently been research by Langenburg and Champod (forthcoming).

Evett and Williams (1996), Langenburg (2004), Schiffer and Champod (2007), and Langenburg (2009) all found inconsistencies among examiners regarding the number of minutiae observed. Evett and Williams (1996) suggest that this confirms the subjective nature of points of comparison, and Langenburg (2004) and Schiffer and Champod (2007) report that these variations are larger with novices.

Table 14 Target Comparison Print Data. The number of minutiae observed by each examiner for each fingerprint, the minimum number per fingerprint ('Min'), the maximum number per fingerprint ('Max'), the standard deviation ('SD') and the relative standard deviation (RSD) for each fingerprint, and the range of minutiae observed for each fingerprint (presented on the bottom row).

	LATENT FINGERPRINT									
	A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**
Observed Minutiae	22	9	15	8	9	3	8	11	7	10
	21	11	25	7	10	9	9	10	6	5
	19	9	18	10	7	9	15	19	6	6
	21	21	29	14	12	9	8	9	4	8
	17	16	15	11	16	9	7	12	5	5
	20	14	22	9	10	7	13	18	7	9
	22	17	15	10	10	8	11	24	8	11
	9	9	19	6	9	8	18	16	9	10
	30	15	25	10	12	12	19	22	12	17
	25	13	18	13	12	10	13	15	7	10
Min	9	9	15	6	7	3	7	9	4	5
Max	30	21	29	14	16	12	19	24	12	17
Mean	20.1	13.4	20.1	9.8	10.7	8.4	12.1	15.6	7.1	9.1
SD	5.49	4.01	4.93	2.49	2.45	2.32	4.25	5.15	2.23	3.54
Range	**21**	**12**	**14**	**8**	**9**	**9**	**12**	**15**	**8**	**12**

As fingerprint examination advances, more objective measures and standards will ensure greater consistency among examiners (Dror, 2009). The potential influence introduced by a 'target' comparison print was addressed earlier. Another issue is the calibration of the threshold for determining a minutia is a 'signal.' Different examiners may be using different threshold criteria, and hence the large variance in how many minutiae different latent fingerprint examiners report observing on the same latent mark (similar problems occur in other forensic domains; see, for example, the lack of agreement on colour description used to determine the age of a bruise, see Dror, 2009).

A simple training tool can help deal with this problem. A set of latent marks can be made available for examiners to analyse. After analysis, personal feedback will be provided to the examiner as to how consistent they are with other examiners. For example, the feedback supplied may state that ‘your analysis resulted in similar minutiae as most examiners (and hence no need to calibrate thresholds), or it may state that ‘your analysis resulted in a larger (or much larger, or smaller, as the case may be) number of minutiae relative to most examiners (and hence the examiner may consider changing their thresholds). The idea is that this would be a private measure, with results and feedback confidentially available only to the individual examiner. The full technical details of such a training calibration tool and its implementation are beyond the scope of this chapter, but they are straightforward. For example, the previously mentioned PiAnoS software allows for an examiner to select minutiae, and can display the annotations of other examiners. Some more conceptual issues that need to be addressed are which latents should be used for this purpose, and how to make sure the feedback is taken on board when examiners do indeed need to re-calibrate their judgments. These must be scientifically based decisions. Furthermore, a fundamental issue that needs to be addressed is that the calibration is done to the ‘correct’ level, because ensuring different examiners use the same criterion, does not mean they are using the ‘correct’ one.

Intra-Observer Consistency

Judgment and subjectivity affect the number of minutiae characteristics reported, resulting in inconsistency among experts on how many minutiae are present within a specific latent mark. The experiment within this chapter and others (Evett & Williams, 1996; Langenburg, 2004; and Schiffer & Champod, 2007) all consistently show that these variations are further dependent on the actual mark and prints in questions (i.e., some produce higher inconsistency than others). Furthermore, as already reported in this thesis, some examiners are more affected by context than others. To ascertain the role of individual differences (such as experience, motivation, training, feature selection strategy, thresholds, cognitive style, personality, etc.) versus the contribution of lack of objective quantifiable measures for determining characteristics in analysis of latent marks, an intra-observer (within-expert subject) experimental design was introduced.

Within-expert experimental design examines intra-observer effects, comparing an examiner to his or herself, thus controlling for individual differences (see Dror & Charlton 2006 and 2006). Another experiment within this chapter also examined the consistency in analysis of latent marks by the same expert examiner at different times. A new set of expert examiners was used. They were asked to report the minutiae present on ten latent marks. A few months later, they were asked to do the same exercise, thus receiving the same identical instructions at time 1 and at time 2. The experts overall analyzed 200 latent marks, 100 prints twice. Table 15 presents the actual data: 10 latent print examiners, each making 20 analyses in total, analysing 10 latent marks (A to J), at $Time_1$ and at $Time_2$. In contrast to Table 14 where we examined the overall range and consistency obtained across examiners, here the focus was on comparing each examiner to his or herself, specifically looking at the degree to which the experts were consistent with themselves.

Analysis of variance (ANOVA) of the data from Table 15 showed that examiners differed significantly from each other in the number of minutiae reported: $F_{(9,81)} = 8.28$, $p = < 0.001$, effect size correlation eta = 0.69. This analysis also showed that the number of minutiae observed differed significantly from each other depending on the finger mark: $F_{(9,81)} = 57.30$, $p = < 0.001$, effect size correlation eta = .93. Note the larger effect size for the contribution of the latent marks compared to the effect size for the contribution of the examiners. Most important is the Retest Reliability reported in Table 15 (right column) which is a statistical measure for quantifying consistency; see also the Stem-and-Leaf Plot and the Five Point Summaries of Retest Reliabilities in Figure 38.

Table 15 The number of minutiae observed by each examiner (1-10), for each mark (A-J), at Time 1 and at Time 2. The last column shows the retest reliability statistic for each of the 10 examiners.

EXAMINER		Finger Mark										Retest Reliability (r_{12})
		A	B	C	D	E	F	G	H	I	J	
1	Time 1	27	15	17	9	9	7	16	13	7	13	.95
	Time 2	26	14	21	10	8	5	13	15	7	12	
2	Time 1	31	16	14	9	10	7	12	13	6	9	.85
	Time 2	23	13	19	10	9	9	10	8	8	11	
3	Time 1	19	11	13	5	9	5	8	12	6	10	.65
	Time 2	18	8	16	8	15	9	17	21	7	12	
4	Time 1	20	12	17	6	10	8	7	8	6	7	.92
	Time 2	22	9	19	11	10	9	8	8	6	8	
5	Time 1	19	11	19	6	10	13	9	14	8	12	.84
	Time 2	25	13	21	9	14	12	12	11	8	9	
6	Time 1	34	16	21	12	13	13	12	11	8	12	.80
	Time 2	25	12	23	11	17	7	12	16	9	13	
7	Time 1	21	9	19	9	12	9	10	18	6	10	.80
	Time 2	21	13	14	7	8	6	7	11	6	10	
8	Time 1	19	14	14	10	9	6	12	13	7	11	.87
	Time 2	22	13	18	10	15	8	13	17	5	11	
9	Time 1	19	11	11	7	9	4	8	15	5	2	.88
	Time 2	23	14	20	7	13	8	11	14	4	5	
10	Time 1	19	10	9	8	4	2	10	8	6	5	.91
	Time 2	20	10	9	7	8	3	6	7	6	5	

Inconsistencies were observed to see if they occurred over the typical range of thresholds for potential decisions (e.g., 8 vs. 17, see Examiner 3, latent finger mark G), or in ranges that do not typically matter for identification (Examiner 6, latent finger mark A, 25 vs. 34). Both cases had a difference of 9 minutiae, but the former variability is more likely to cross a decision threshold for identification, while the latter's range of values are more likely to all be over decision threshold (of course, this cannot be determined from the data in the present study with, as this analysis is on the latent finger mark alone, prior to comparison to a print). Do examiners even consider identification thresholds when conducting the initial analysis? Evett & Williams (1996) reported that the number of minutiae participants observed was influenced by decision thresholds, e.g., “participants tended to avoid returning 15 points” (p. 7). Categorical perception makes people perceive information according to psychological categories rather than by their actual physical appearance (see Harnad, 1987).

	Examiners		Finger Marks
	1, 2, 5	.9	
	0, 0, 4, 5, 7, 8	.8	7
		.7	
	5	.6	8
		.5	4, 9
		.4	9
		.3	0, 5, 6
		.2	4
		.1	6
Mean	0.85		0.46
SD	0.085		0.22
Min	0.65		0.16
Median	0.86		0.43
Max	0.95		0.87

	Examiners	Finger Marks
Maximum	.95	.87
75th percentile	.912	.612
Median	.86	.425
25th percentile	.80	.285
Minimum	.65	.16
Mean	.85	.46
SD	.085	.218
S^2	.007	.048

Figure 38 Stem-and-Leaf Plot of Retest Reliabilities of 10 Fingerprint Experts and 10 Finger marks (top panel) and Summaries Statistics of Retest Reliabilities of 10 Finger marks Experts and 10 Fingerprints.

To further investigate and understand the inconsistency the absolute differences in the analysis between Time 1 and Time 2, for each examiner (1-10) were calculated for each finger mark (A-J), see Table 16. A score of ‘0’ reflects a consistent analysis. As evident in Table 16, there were only 16% such consistent analyses. If we ‘relax’ the criteria for consistency, and characterize consistency as a difference of 0 or 1, then there are 40% consistent analyses; if we further relax our criteria for consistency to include a difference of 2, then there are 55% consistent analyses (or, stated differently, 45% of the analyses differed in at least more than two minutiae between the two analyses conducted by the same examiner. These data raise questions about objective assessment even at the analysis stage (which seems to be more robust to influences and context than the other stages of the ACE). The data reported here are conservative, as the variability may be much higher.

Table 16 The differences in number of minutiae observed by the same examiner at different times. The bottom row is the mean difference per latent mark (A-J), and the right most column is the mean difference per examiner (1-10).

	LATENT FINGER MARK										MEAN
PARTICIPANT	A	B	C	D	E	F	G	H	I	J	
1	1	1	4	1	1	2	3	2	0	1	1.6
2	8	3	5	1	1	2	2	5	2	2	3.1
3	1	3	3	3	6	4	9	9	1	2	4.1
4	2	3	2	5	0	1	1	0	0	1	1.5
5	6	2	2	3	4	1	3	3	0	3	2.7
6	9	4	2	1	4	6	0	5	1	1	3.3
7	0	4	5	2	4	3	3	7	0	0	2.8
8	3	1	4	0	6	2	1	4	2	0	2.3
9	4	3	9	0	4	4	3	1	1	3	3.2
10	1	0	0	1	4	1	4	1	0	0	1.2
MEAN	3.5	2.4	3.6	1.7	3.4	2.6	2.9	3.7	0.7	1.3	2.58

Analyses of variance (ANOVA) of the data of Table 15 showed that although examiners differed significantly from each other in their degree of consistency in judging fingerprints (eta = .44), and in the number of minutiae observed (eta = .69), they still showed a high degree of inter-observer reliability with each other ($r_{intraclass}$ = .85) and with themselves (retest reliability r = .86). The examiners who showed the highest retest reliabilities also tended to show the smallest discrepancy between their two evaluations of the same fingerprints at time and at time, r = .84.

The differences between Time 1 and Time 2 (see table 15) show that some examiners are more consistent than others (see, e.g., examiner 10, who is relatively highly consistent vs. examiner 3). Indeed, analysis of variance (ANOVA) of the different scores in Table 15 showed that examiners differed significantly from each other in the consistency with which they judged the 10 fingerprints: F= 2.17, p = .032, effect size correlation eta = .44. Are the more consistent examiners characterized by personality type and cognitive aptitudes? If so, we need to know how to select candidates with such cognitive profiles during recruitment. Or perhaps these examiners receive a certain type of training, or maybe they adopted more objective definitions? All these are important questions that may help pave the way to understanding how such variations can be minimized.

However, the inconsistencies did not only vary between examiners, they were also dependant on the finger mark itself. The analysis of this variance also showed that finger marks differed significantly from each other in the consistency with which they were judged: F= 2.82, p = .006, effect size correlation eta = .49. This means that some marks are just more susceptible to issues of consistency than others. However, understanding and characterizing what constitutes such marks is not a simple matter, and we must be careful and not be hasty in determining how to a priori know which prints are susceptible to inconsistent analysis. With careful further research and converging investigations, it should be possible to learn and predict which marks are likely to be problematic. This is an important step to remedy the problem. Once we know which marks are likely to cause consistency issues, we can recommend appropriate scientifically based procedures that attenuate the problem. For example, in marks of low quality, instructing a number of examiners to only mark minutiae that they have high confidence in. And then allow only use of those minutiae that have been marked across different examiners, thus using consensus to determine the reliable features to use in such marks. Another approach is for mapping quality and clarity across a latent finger mark, so as to map high, medium, and low quality regions. Variability of feature selection may be lower if examiners are required to select only from the higher quality regions, but that may entail losing out on information. In this chapter it has been possible to identify a common phenomenon found in many expert domains, invite debate on the topic and its significance, and have suggested recommendations to deal with it.

Experts		Finger Marks
1	4.	
1, 2, 3	3.	4, 5, 6, 7
3, 7, 8	2.	4, 6, 9
2, 5, 6	1.	3, 7
	0.	7

	Experts	Finger Marks
Maximum	4.1	3.7
75th percentile	3.22	3.52
Median	2.75	2.75
25th percentile	1.58	1.6
Minimum	1.2	0.7
Mean	2.58	2.58
S	.922	1.049
S^2	.851	1.100

Figure 39 Stem-and-Leaf Plot of Absolute Difference Scores of 10 Fingerprint Experts and 10 Fingerprints (top panel) and Summary Statistics of Absolute Difference Scores of 10 Fingerprint Experts and 10 Fingerprints.

The experiments reported in this chapter have identified for some finger marks or for some experts, significant inter and intra-observer variations during minutiae selection and it is relevant to ask: What impact can this have on the overall comparison decision making outcome? Is the lack of consistency a practical concern or an academic issue? The answer to these questions appears to be complex and depends on a number of factors. For example, in Evett and Williams (1996) the variations in reported minutia did not totally predict the variations in overall decision outcome. In their study, Trials B, E, and F (which varied a lot in minutiae reported by some examiners), had 99%, 92%, and 100% consensus (N = 130) that the mark and the print originated from the same source. In other words, the variations (e.g., Trial F varied up to 42 minutiae), did not necessarily prevent experts reaching the same final conclusion. In contrast, other trials (such as Trial H) which had smaller variations, had less consensus on the final overall decision (in Trail H, e.g., 54% concluded they are likely from the same source, 38% reported insufficient detail to make a decision, and 8% reported they are not from the same source). Here the variation in feature selection appeared to be critical. In Langenburg, Champod, and Wertheim (2009) a similar trend was observed. In their study, participants reported ranges (maximum differences) of 21, 17, and 12 minutiae respectively for Q1, Q4, and Q5 trials. However, trials Q1 and Q5 resulted in 100% consensus (N = 43) for the reported decision. Q4 on the other hand resulted in three errors, and the remaining participants nearly split on reporting "identification" or "inconclusive".

Those that reported "identification" had a statistically significant higher likelihood of also reporting more minutiae. In this trial, it appeared that the number of minutiae observed directly correlated to the decision reported and was a critical function of the decision making process. Therefore, although it is clearly a critical issue, variation needs to be researched and understood better.

It appears as a general trend in the previous investigations, that the reduction of available minutiae in a finger mark, especially to a point where the amounts may hover around categorical decision thresholds (i.e. “identification” versus “inconclusive”), can lead to different reported decisions. Therefore, a possible best practice would be to identify a priori which marks are likely to produce such decision variations and apply special procedure, such as previously discussed (use of consensus minutiae, quality mapping, conservative selection procedures, etc.). Further research is recommended here, particularly to determine which suggested variation reduction technique is appropriate and effective.

Summary and Conclusions

Feature selection during the analysis stage of a latent mark is important because it sets the stage and the parameters for comparisons and decision making. Although this stage is relatively robust, it is still susceptible to observer effects. In this chapter it was established that the presence of a comparison ‘target’ print may affect the analysis stage. Furthermore, there is lack of consistency in the analysis not only among different examiners (e.g., reliability among examiners r = .85), but also within the same examiners analysing identical latent marks at different times (retest reliability r= .86). The characterization of experts’ consistency depends on the standard applied. If we examine the purest test of consistency, i.e., how consistent examiners are with themselves, then the retest reliability of r= .86, though far from perfect is respectably high; but using another standard, we find that at best only 16% of experts observed the exact same number of minutiae when analysing the same latent mark (40% of the experts were within one minutia difference, and 55% were within a difference of two minutiae).

This experiment goes beyond establishing that analysis of latent marks by experienced latent print examiners is inconsistent. First, it demonstrates that the presence of a comparison print can affect the analysis of the latent mark. Second, it shows that examiners are inconsistent among themselves; i.e., different examiners vary in their analysis. Third, it reveals that the consistency of examiners with themselves varies; some examiners are relatively consistent with themselves and others are not. Fourth, the lack of consistency does not only depend on the examiner in question, but it also highly depends on the nature of the latent mark itself.

For each of these findings there are potential recommendations to mitigate the problems. First, given the effects of the comparison print, it is suggested that initially the analysis of a latent mark should be done in isolation from the comparison print. Furthermore, the examiner should not rule out reconsideration of the analysis after exposure to the comparison print, but stipulate that this process, should it occur, must be clearly and transparently documented, and justified. Further research needs to consider other ways to deal with variation in the analysis stage. One suggestion may be that examiners should mark confidence levels in minutia detection; thereafter they can only reconsider low confidence judgments but cannot change those that were analyzed initially with high confidence (see Dror, 2009, for details).

Second, given that examiners vary among themselves in their analysis, it is recommended that there should be the development of a simple calibration tool that enables examiners to adjust their threshold so as to meet the standards in the field.

Third, given that some examiners are more consistent than others, I am confident that with proper selection of examiners with the right cognitive profiles specifying the exact skills needed for latent fingerprint examination, and with proper training, can reduce the examiners' contribution to inconsistencies observed in finger mark analysis.

Fourth, given that the latent marks themselves play a major contributing role to the inconsistencies, and that these contributions vary with different marks, it is suggested that such marks be subject to a different analysis procedure. Namely this would require using only higher confidence consensus minutiae that a number of independent examiners agree on.

Determining characteristics in finger mark analysis is critical and measures must be taken to minimize inconsistency and increase objectivity. These issues are not limited to fingerprint examination; there are similar issues across the forensic disciplines, including DNA. It is noted that the potential problems with inconsistent analysis may be acute only when the comparison and latent print are near the threshold for identification (and thus one analysis may result in identification whereas another analysis does not; problems may also arise around judgments of 'inconclusive' when another analysis may be sufficient for identification). When the decision is considerably beyond the threshold of determination, then these issues may not have important practical implication (as both analyses, although inconsistent, still will result in the same overall decision).

Understanding the cognitive issues involved in pattern matching and decision making, and researching them within the realm of fingerprinting is a promising way to decrease expert variation, improve the reliability of fingerprinting, and to gain insights into the human mind and cognitive processes.

Chapter 7:

Cognitive Technology in the Fingerprint Domain: A Theory of Phased Cognitive Engagement

> *'The power of the AFIS match, coupled with the inherent pressure of working an extremely high-profile case, was thought to have influenced the initial examiner's judgment and subsequent examination. This influence was recognised as confirmation bias (or context effect) and describes the mind-set in which the expectations with which people approach a task of observation will affect their perceptions and interpretations of what they observe'.*

Robert B. Stacey, Quality Assurance and Training Unit, Federal Bureau of Investigation (2005) (Referring to the circumstances around the erroneous identification of Brandon Mayfield by the FBI as part of the Madrid bombing investigation).

Technology in the realm of fingerprint examination has been a part of the discipline for many years. Indeed, the very first automated fingerprint search tools were being developed as long ago as the 1930's.

In 1963 one of the most infamous crimes of the 20th century was committed in the UK, The Great Train Robbery. The Scotland Yard Fingerprint Branch was instrumental in identifying many of the perpetrators (including Ronnie Biggs) after an exhaustive forensic examination (fingerprints) of the hideout at Leatherslade Farm. The fingerprint examiners at the time would have conducted exhaustive manual comparisons of the crime scene marks against nominated suspect ten-print cards, or, they would have employed a system of coding latent fingermarks found at the crime scene to facilitate a manual search through the National Ten-Print Collection using nothing but their eyes and a thumb stool with which to painstakingly manually compare the myriad fingerprint records (see figure 40).

Figure 40 Typical fingerprint comparison benches circa 1932 (Internet images)

Figure 41 Margaret McCarthy demonstrates Automated Fingerprint searching on the new IBM Card Sorter in 1937 (Internet images).

In the same year (1963), New Scotland Yard began investigating more seriously the potential use of fingerprint computer search systems. The rationale behind the drive to automate processes was as obvious decades ago as they are now. It was stated by Petersilia that no matter how competent the evidence technician is at performing his or her job, the gathering of physical evidence at a crime scene will be futile unless such evidence can be properly processed and analyzed. He went on to state that since fingerprints are by far the most frequently retrieved physical evidence, making the system of analyzing such prints effective will contribute the most toward greater success in identifying criminal offenders through the use of physical evidence. (Petersilia, 1975)

Some attempts at automated fingerprint identification were made as long ago as the 1930's (see figure 41).

In 1937, the first automated fingerprint searching machines were installed by IBM at the New York State Bureau of Identification (Harling 1996). These horizontal sorters, using punched cards containing coded fingerprint classifications, could search suspects at a rate of 420 comparisons per minute. The bureau also established a mechanical Personal Appearance File (PAF) which searched physical descriptions at the rate of 420 per minute. However, it was the introduction of the system for searching latent prints which the International Association for Identification (IAI) called "the most notable single contribution in the field of dactyloscopic work in many years".

By 1966, the back record conversion (BRC) of the National Fingerprint collection at New Scotland Yard had begun. The back record conversion involved the digitization of the entire National Fingerprint Ten-Print Collection. This was in preparation for the loading of digital coding information on to the planned computer systems to follow.

In conjunction with this work, the United Kingdom's Home Office began conducting research into AFR (automatic fingerprint recognition) (Millard 1975 and 1983). The UK Home Office developed a reader to detect fingerprint minutiae that could record position and orientation, and determine ridge counts to the five nearest neighbours to the right of each minutia detected. This was the first use of ridge count information (Moore, 1991). Over the next few years, further work to prepare fingerprint identification for future digitization occurred as follows:

1977 - The Videofile imaging system came on line at New Scotland Yard.

1984 - The first AFR (automated fingerprint recognition) system was installed at New Scotland Yard.

By 1993 Over 10,000 scenes of crime identifications were made annually at Scotland Yard as the National Fingerprint Collection grew to over 4,500,000 sets of fingerprints. The National Automated Fingerprint Identification System (NAFIS) was the first nationally integrated fingerprint matching tool the World had ever seen. It went live in 1999. By 2004 100,000 records were processed every month and 80,000 crime scene marks were searched against the National Fingerprint Collection. What used to take fingerprint examiners weeks and months could now be completed in a fraction of the time.

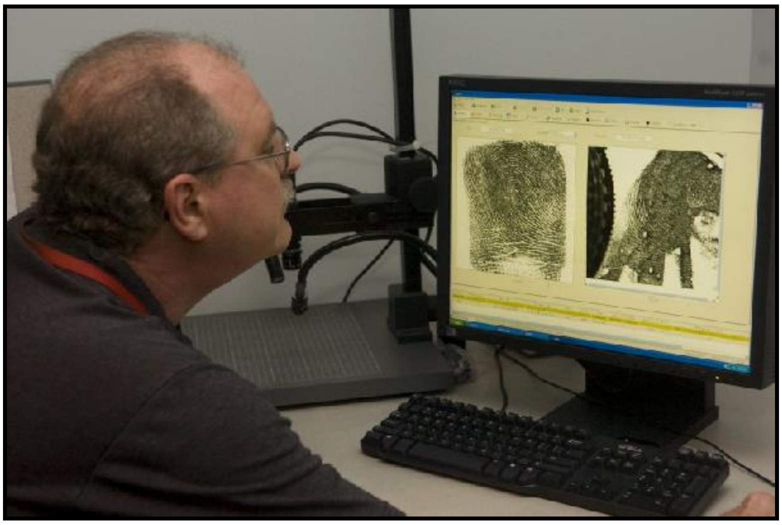


Figure 42 Modern automated fingerprint matchers (Fingerprint Society archives)

What might have taken one fingerprint examiner an entire career (several careers) to perform (for example to conduct a manual search of one crime scene mark from the Great Train Robbery against the entire National Fingerprint Collection comprising 7 million tenprint cards = 70 million individual comparisons) can now be performed by modern AFIS technologies in less than five minutes (see figure 42).

Technology is now available that facilitates the rapid remote transmission of crime scene material from the crime scene to the fingerprint bureau (FISH, Forensic Information Scanning Hub). Such technologies (see figure 43), in addition to AFIS fingerprint matchers are changing the nature of the human contribution to the assessment of fingerprint evidence. Decisions are now made more quickly as to what crime scene marks may be suitable to search on AFIS, and thus, which marks require either re-scanning on another digital system, or maybe require transmission through a GMCI (generic mark camera interface).

Figure 43 FISH (Forensic Information Scanning Hub) (Courtesy Sussex Police)

The speed with which evidence can now be moved from one location to another as well as the speed in which forensic evidence can now be processed, means that cognitive processing of information to understand the evidence being observed, as well as the interpretation of that evidence and the intelligence it produces, is now subject to a time driven need for results that in itself may be detrimental to the ability of both forensic scientists and police investigators to effectively and efficiently rationalise the data put before them, leading potentially to inappropriate strategies for both the interpretation of evidence and the inferences to be drawn from the results of such examinations.

Aligned with the introduction and use of technology in the fingerprint domain are adapted operational processes for fingerprint examiners that have arisen from a need to use such technology as efficiently as possible to maximise the potential to find evidential material that will aid an investigation. The UK National Fingerprint Manual (ACPO issue 1 2006), for example, stipulates for latent print searching on AFIS that finger marks recovered from crime scenes should be assessed before being put on the system. In addition, search parameters should be established (local, regional, national) and case records will be updated in line with national requirements.

The boundaries between human and technology in the fingerprint domain and where cognition is shared or distributed between the two (Hollan et al 2000, Hutchins et al 2000, Rogers et al 1994) is becoming less well defined. As a consequence it is important to understand how the 'blurring' of the cognitive boundaries between technology and the practitioner impacts on human performance when examining fingerprints within modern fingerprint Bureaux.

New technologies continue to impact and encroach upon the traditional human cognitive process, which includes the intensity of the cognitive links between human and technology. Technologies that cognitize with us as humans have strong influences over our own cognitive functions, sometimes taking over such functions altogether. Automated fingerprint identification systems (AFIS) are examples of cognitive technology (Dascal and Dror 2005). Human cognition is dynamic and adaptive so when such dynamic systems interact with cognitive technologies then it can be assumed that the interaction between the human and technology will form and shape new cognitive thinking, shaping the very decision making processes and other cognitive functions.

Cognitive technology and how the cognitive burden is shared between the technology and the human is a very important area for consideration in the forensic domain and fingerprint examination specifically. Cognitive technologies offer great opportunities and benefits but can also be dangerous and counter productive. Anecdotal comments made by fingerprint examiners indicate that the balance of cognitive function between human and technology may not be distributed appropriately within the fingerprint domain and that it can be hypothesized is causing a detrimental imbalance between the needs for procedural and technological efficiency as well as human efficiency. This, in turn, may possibly explain some of the demotivating language used by examiners when discussing technology which will be discussed in more detail later on in this chapter.

The deployment and use of cognitive technology is not a simple matter to get right (Dror and Mnookin 2009). Indeed, if the interaction between the human and technology is to be appropriately balanced to achieve a mutually beneficial state that achieves maximum efficiency, then more research is required to better understand the relationship between the human fingerprint examiner and the technology currently designed and deployed in this domain.

While Dror and Mnookin centre their arguments on response similarities associated with AFIS fingerprint searches, whereby the sheer weight of database size will result in fingerprint response similarities that may make decision making harder, other risks present themselves that result from a potential change of cognitive state on the part of the human as a result of the interaction with technology. For example, the very design of AFIS systems precludes the use of level 3 information when assessing a crime scene mark and it's suitability for search. The generic AFIS system can only utilize so called level 2 Galton details for search purposes as the mathematical algorithms deployed are only designed to operate within such design scope (Maltoni et al 2003). Therefore it can be hypothesised that AFIS systems exert an influence on human thought processes to employ contrary, or at least truncated assessment strategy when considering technology than they might otherwise have employed in a manual search and comparison against suspects, where more holistic strategies using accepted processes such as ACE-V (Ashbaugh 1985, 1991, 1995) may be employed by the examiner. In short, whatever might be taught or indoctrinated into fingerprint examiners through experience and learning might be modified by the interaction with technology in the work place. For example, most AFIS systems, utilizing level 2 Galton detail require a minimum number of minutiae (level 2 features) to facilitate a successful search.

The number usually suggested ranges from 8-10 Galton points (ridge endings and bifurcations etc) depending on the AFIS system (Maltoni et al 2003), and any fewer Galton points severely impacts on both the respondent list length and the quality of respondents for comparison (ambiguous and poor quality) as well as impact on AFIS accuracy thresholds. For this reason fingerprint examiners may be discouraged from attempting to search relatively poor quality fingerprint material on AFIS systems not just on a personal level, but perhaps even by management intervention too.

Such management intervention might result in assessment of examiner performance results that might prove unfavourable to the examiner should the search ratio of input to identification achieved be less fruitful to the examiner who searches all useful material on AFIS systems, rather than someone who chooses the most likely marks likely to yield positive outcomes.

It can be argued that AFIS technology dictates when and how cognition is shared and that the human is to some point required to act as the expert 'fail safe', ensuring that what technology is incapable of performing, the human can augment in overall performance. In contrast, examiners are asked to employ a totally different cognitive strategy when using manual methods in that now the total cognitive burden is now with the human. In this scenario, fingerprint examiners are now expected to consider all ridge detail holistically and to consider that there is no minimum level of ridge detail that cannot be considered for comparison against a suspect. This difference in strategic approach as to how assessment of fingerprints (The 'A' in ACE-V) may be affected by cognitive technology is worthy of further investigation.

Human interaction with technology is discussed by Dror et al (Dror 2012) as well as Dror and Mnookin (Dror and Mnookin 2010) who describe three categories of cognitive distribution. They describe low level distribution which has the impact of improving efficiency, such as storing data on a computer, a higher level of cognitive distribution that represents more of a share of tasking responsibilities, for example, the clinical discussion arising from the production of an X-ray, and finally, the highest form of distributed cognition where the technology effectively takes over much of the function of the human actor. It can be argued that automated fingerprint matchers that are approaching a lights out capability (where there is little or no human intervention) are an example of the highest form of cognitive distribution between the technology and the human in the fingerprint domain.

The use of AFIS systems by fingerprint examiners may also have an impact on decision thresholds in establishing whether there is enough detail within the fingerprint to search on AFIS systems (utility). Traditional concepts and values associated with the assessment of utility in the fingerprint domain may change as technology becomes more powerful and dominant. Fingerprint examiners' assessment of utility is based on the relationship between the clarity of the print in question in relation to the quantity of features observed. So, high clarity marks showing all three levels of detail (pattern, Galton details and sweat pores and ridge detail) but with few actual details available may be better for comparison purposes than finger marks with high numbers of Galton detail features where there is a lack of clarity and where there is ambiguity.

It is possible that AFIS technology, that is not designed to think in the same way as the human in this regard, may induce different strategies to the assessment of utility that are not part of traditional fingerprint teachings and that are subconsciously adopted by the expert examiner without him or her knowing it. In essence exerting top down contextual influence on the examiners, not through the context of the crime or the case details (Dror and Charlton et al 2005, 2006), but rather through previous experiential knowledge of how AFIS works and preconceived ideas of what such systems are capable of which will in turn exact a motivational influence on decisions made around the use of such technology.

Experts are the 'instrument' of interpretation and forensic decision making, even when cognition is distributed between human and machine. In a recent paper by Dror et al (Dror et al 2012) the impact of using Automated Fingerprint Identification Systems (AFIS) on human decision makers was investigated. They provided 3,680 AFIS lists (a total of 55,200 comparisons) to 23 latent fingerprint examiners as part of their normal casework. They manipulated the position of the matching print in the AFIS list. The data showed that fingerprint examiners were affected by the position of the matching print within the respondent list in terms of decision making around false exclusions and false inconclusives. The data also showed that false identification errors were more likely at the top of the respondent list, and that such errors occurred even when the correct match was present further down the respondent list.

So does technology impact upon human decision making in the forensic domain? In a review of the use of automated fingerprint recognition technology in the UK authored by the Home Office Police Standards Unit in 2004 (Gold 2004), several pertinent observations were made on the use of automation in the fingerprint domain. Amongst the most striking observations made included a substantial variance between different fingerprint bureau in the overall use of the technology, considerable underuse of the capabilities of the system, and a disregard for the guidance issued on best practice use of the deployed automated system. Indeed, recommendation 1 of this review stated that 'fingerprint Bureaux should ensure that all searchable marks should be searched on automated matcher systems'. However, the review went on to add that 'some fingerprint Bureaux continued to perform manual search and comparison of crime scene marks despite the empirical evidence suggesting that automated matchers are robust and reliable'. The observations Gold made seem to suggest that the AFIS technology was underused, even though fingerprint examiners are trained and instructed to use AFIS technology as often as possible. Further investigation is required to understand not only when there is conflict in the interface between human and AFIS technology, but also whether this has an impact on overall cognitive efficiency given the task involved, namely to search and find fingerprint matches as efficiently and accurately as possible.

While it has been shown that AFIS and its influence has an impact on management and performance of the organization (e.g., Klug et al 1992, Petersen et al 1996), there has to date been very little discussion of the effects of AFIS systems on the operational and theoretical decision processes of human examiners. To date there is little empirical data to assess the nature of the imbalance between technology and human in the fingerprint domain, or in just how demotivating technology is on the examiner given the broad variance of not just fingerprint pattern types but also the variances in fingerprint quality too.

In summary, one of the key attributes of expertise is the ability to work efficiently and effectively, filtering out extraneous information that will impact on cognitive processing. However, such attributes also bring with them less advantageous traits that might include a greater influence of top down cognitive processing and selective attention to detail (Dror 2012 and Busey et al 2010). An expert in fingerprint comparison who may also be expert in using both AFIS and manual techniques to search fingerprints, may use parallel and diverse strategies for examining fingerprints that may be directed by such top down influences as context, which could include the perceived technical requirements, benefits and disadvantages of using AFIS technology, as well as the crime type from which the fingerprints came, whether there are suspects available for comparison, as well as other experiential knowledge, such as the difficulty of comparing certain types of fingerprint patterns on screen (as opposed to manual comparison techniques) and their relative ridge flow configurations.

In order to better understand the relationship between human and technology in the forensic domain a comprehensive four phased investigation was undertaken to assess how fingerprint examiners perceive different fingerprint patterns, the level to which usable fingerprint evidence is searched on AFIS systems as well as to understand under what circumstances certain decisions around AFIS search strategy is made, and whether there is any change to ACE-V process as a result of the presence of cognitive technology.

Interviews with Fingerprint Examiners to Determine the Difficulties of Latent Print Analysis

Between the 9th January 2010 and the 8th February 2010 a series of interviews were conducted with 30 individual fingerprint practitioners from a variety of law enforcement and consultative agencies in the UK and the US. The interviews were constructed in such a way as to encourage latent print examiners to volunteer information pertaining to their daily workload. Each interview was digitally recorded by the interviewer to enable accurate transcription of the data produced; however, each recording was destroyed as soon as was practicable after the completion of the transcription task as part of the agreement to conduct these interviews. No audible records now exist of these interviews, although the transcripts are preserved. Each interview lasted approximately 30 to 45 minutes and was held in a variety of settings, which provided seclusion, privacy and quiet that promoted free discussion. Each participant was given a coded identifier to anonymise their contribution. Full consent to participate was sought and agreed to.

Specific themes were assimilated, whereby individual thematic categories were generated from the interview text itself rather than from specific theory (Joffe et al 2004). Themes were broken down into separate subjects; for example, participants were asked to describe what types of fingerprint examinations are most problematic to analyse and each interviewee was asked to describe their thoughts on AFIS systems and their use, as well as other areas of discussion around distortion of prints and how this impacts upon examinations. It was important to be cautious not to generate too many categories. Consequently, a few, broad, general themes were chosen as this allowed much greater generalization.

Analysis

Using thematic analysis protocol, codes were assigned to various segments of the text. It was not clear what the findings would be and there was little in the way of guidance from past research. As a result, specifics were inductively coded, whereby individual categories were generated from the interview text itself rather than from specific theory (Joffe et al 2004). Initially, the coding was very broad to encompass anything which had content aligned to discussions on fingerprint complexity. It was important to be cautious not to generate too many categories. Consequently, a few, broad, general themes were chosen as this allowed much greater generalization. It was also important to this particular investigation that both latent and manifest content was coded. Although this involves a certain amount of interpretation by the researcher, it was hoped that any clarifications made during the interview and the concurrent notes made would avoid inaccurate interpretation.

Reliability is vital in any qualitative study. It was important that the coding was both stable and consistent and that it had good reproducibility (Krippendorff 1980). There is another reliability measure, accuracy, which refers to the extent that the coding corresponds to a previously generated standard or norm, which provides the strongest form of reliability (Weber 1985). However, as this is a new area of study there are no standards to compare it against. As a result it was not possible to measure accuracy reliability. However, it is hoped that the themes from this study might be used to gauge further qualitative studies investigating the technical experiences of forensic or criminal investigative personnel.

Results:

Experience, length of service and what is easier now than early on in careers:

Some participants felt that it was harder to differentiate between individual Galton details when first starting out in their career, and that it was subsequently possible, with experience, to focus in on detail even in light (faint) prints, establishing whether a feature was a ridge ending or bifurcation. Vague detail was seen as problematic at the start of a fingerprint examiner's career. Many examiners alluded to seeing more detail today in difficult prints, than they could earlier in their careers, with examiners having more confidence to call matches in difficult prints as their experience grew. With experience, examiners apparently find it much easier to assess distortion and to determine substrate and matrix material. In contrast, some participants suggested that those just starting out in their careers often find it difficult to analyze the arrangement of ridge flow and features.

Why everything seems easier as experience is gained was explained by one participant who discussed the examination of fingerprints in the following way, 'the eye is used to seeing ridges, bifurcations and dots'. In other words, it appears that there is a form of experiential 'mind programming', as one examiner put it, based on trial and error that enables the examiner to learn what a valid level two Galton feature looks like. It should be noted however that this 'mind programming' takes place without a 'ground truth' feedback system. This means that the feedback received is based solely on the subjective opinions of peer review, rather than categorical validation of examiner decision making.

For example, an expert weather forecaster will get feedback of every decision made to confirm or disprove hypotheses, it either rains the next day, as forecasted, or it doesn't. This feedback loop does not exist in fingerprint examination. To prove the hypotheses surrounding uniqueness and individualisation would empirically require examining all fingerprints of all living people as well as those no longer living to ensure the conclusion reached were indeed correct, something that is obviously impossible to achieve.

Assessing ridge flow and structure is also perceived as easier now than would have been the case early in the careers of many examiners. With experience the examiner appears to gain confidence in decision making ability. So by the time they are considered expert and able to testify they appear to be able to consolidate those decision making skills and to be more assertive in their opinions. Of course, whether this is evidence of a cognitive improvement in the individual examiner based on learning and experience, or whether this has more to do with the cultural social psychological nature of group dynamics and peer review pressures is a matter of conjecture and worthy of future investigation.

Examiners reported it was easier with experience to locate focal areas, or targets. As one examiner put it, 'it just pops out now'. It was suggested that inexperienced examiners looked at the entire print holistically and that it took time for these examiners to come to a conclusion, but that with experience examiners are able to 'just concentrate on areas of interest'. Those interviewed who were early in their careers suggested finding and locating areas of friction ridge skin difficult, especially palm, whereas more experienced examiners who participated, suggested they were able to narrow down search more effectively. This insight appears to suggest that with expertise comes a form of 'chunking' (Gobet et al 2001), whereby the 'expert' is taking cognitive shortcuts to help the examiner locate and memorize certain key areas within a fingerprint more efficiently to aid the comparison process, rather than try to take in all the information available which would impact on the ability to process information quickly and efficiently.

Participants felt that today there is less emphasis on understanding how to search for fingerprints because AFIS removes some of that need. Newer examiners, it was suggested, just don't see as many fingerprints as the pre AFIS generation examiners and that this was impacting on the ability of examiners to understand and interpret the fingerprints put before them for examination. As one examiner put it, AFIS, 'takes away the ability to make the decision…you are already given the score…people get used to the ident being at the top of the queue'. 'Examiners don't have the ability on AFIS to decide the search limits and are willing to accept the decision of AFIS, be it a non match and just stick it on the shelf'. 'There is very little motivation now to try harder….we have lost ownership of casework'. 'If you are just doing AFIS work you will not build up the sort of experience you need to make decisions'.

This is strong evidence suggesting an unbalanced cognitive distribution between the human and the forensic technology available suggesting that the interaction between the two entities is not maximising the potential of fingerprint bureau to reliably search, identify and report on fingerprint evidence and will require further investigation in this chapter.

What is the most challenging aspect in latent fingerprint examination?

Most examiners, even with experience, still find interpretation of fingerprint distortion challenging. The types of distortion described by participants included low quality and faint impressions as well as a type of distortion caused by movement that result in a thickening or thinning of ridge detail. It was suggested experience is still no protection against erroneous decision making where there are poor quality prints to examine. It seems those 50/50 calls, ('is it a match, or isn't it'?) still require the examiner to somehow reach the 'tipping point' of 51/49, do I see a match? Or should I 'walk away'? Many examiners alluded to the gray area between identification and a non match, 'where you would want to see one more ending ridge to be sure', 'in the gray area its tough', said one examiner. This need to reach a 'tipping point' in the examination, this 'winner take all' approach, or gambler's bias often serves to increase the effects of escalating commitment (Staw 2004) causing decision makers to refuse to withdraw from a situation, and to continue to throw resources, effort and time after a potentially hazardous fingerprint comparison that should be left well alone, leading in this instance to the potential for erroneous identifications of fingerprints. Challenges presented to examiners by really bad prints, where there is simply a lack of information, but where the examiner has a 'gut feeling' that there is a match to be made, are problematic in that it appears even experienced examiners are unsure or uncertain as to where their individual 'tipping point' or 'decision threshold' is, or even where it should be! The challenges are greatest apparently when you 'need to find a match' and maybe when it is appropriate 'to give up'. Examiners reported faint prints are problematic where the contrast is very poor, and in terms of finding target features to work with and it was suggested that level 1 patterns such as arches make target searching more difficult. It was felt that 'marginal prints' provided the most problems when analyzing latent finger marks as this was when distortion was most prevalent in terms of interpretation and determining feature confidence.

What makes some comparisons more difficult than others?

Some participants felt that rolled prints have different types of distortion properties to that of plain (slap) impressions. With poor and 'ugly prints' participants suggested that even though the comparison was difficult, sooner or later the decision had to be made, a feeling that the decision must either be yes (match) or no (no match). It was suggested that examiners sometimes feel 'obliged to make decisions'. This 'obligation' is clearly an example of peer pressure as well as an example of cultural background pressures associated with working for law enforcement agencies and requires further investigation.

Examiners suggested the difference between different fingerprint capture processes can also make a difference to the ease or difficulty of the comparison. Ink or live scan, or is the person a bricklayer, smoker etc, will all apparently provide different challenges to the examiner. Participants discussed how ‘noise in the artifacts’ will impact on ability to interpret information. Digitally captured latent images were seen as hard to compare to a ten print card when pixelisation of features causes difficulty in interpreting detail. It was also suggested that poor quality ridge structure made it difficult to assess what are real or false characteristics.

Examiners also suggested that they found palm searching more problematic as it is more difficult to orientate the ridge flow. It was suggested there is a bigger area to examine and the scope of search is much larger. Arches seemed to cause more problems too as ’there are less defined areas’. A point of reference to begin comparison is lacking and locating target features to compare becomes more difficult it appears. Some participants felt that arches are more difficult to assess for this reason because ‘you don’t have a reference point for where details are’; ’certain features seem to repeat themselves on arches too’. It was also suggested that since the advent of ‘live-scan’ there appear to be more anomalies toward the top of the print that cannot be explained away easily whereby ridge detail can be distorted to a point of either creating of erasing entire units of ridge detail. When assessing the presence of anomalies between the latent fingermark and the ten- print exemplar the examiner seeks to assess, ‘Is the problem with the mark or the print?’

Latent (print) Orientation

The correct orientation of latent finger marks was described as a bit like looking at a map of the earth, ’it can be cut up like a jigsaw but as soon as you see a continent you know where it fits on the globe’, ‘you know where it belongs’. ‘No matter how you cut it, how you stretch it, it will always be the same globe’. Orientation was defined by observation of ridge flow and observation of deltas and other ridge detail, and was based on experience and training, a judgment based on the experience of understanding ridge flow and where features regularly appear. Orientation, it was suggested, is usually obvious, and there is usually a ‘most likely way it is oriented’. Apparently after seeing thousands of prints you get to ‘know where the front of the car is’. Print orientation is based on experience but also a mix of ‘flipping around’; ’a bit like looking at jigsaw puzzles’. ‘You have a mind that thinks in bits and pieces’; ‘do things flow right’. When asked about orientating a latent finger mark one participant alluded to treating the ridge flow like ‘a puzzle’ to be solved, and often the use of ‘mental rotation of ridge flow and target features helps to define orientation’, based on experiential trial and error. Correct orientation of the latent was described in terms of ‘mental gymnastics’ whereby the features and ridge flow were somehow ordered mentally.

In short, participants mentioned skills for orientation of latent finger marks including visual acuity, the ability to hold the detail in the mind and to draw from the database in the mind and match up likely ridge flow patterns etc. This seems to confirm the concept, as described by Shepard et al, of the use of mental imagery to create mental representations and transformations in the determination of pattern similarities (Shepard et al 1982).

Deciding where to concentrate your analysis of the (latent) finger mark.
Target searching involves looking for groups or clusters of unique similarities. Participants discussed looking for targets in the same space and relationship, not counting points, but rather working holistically. Examiners, it seems, look at any features that stand out. This could be ridge flow, or it might be lakes or sweat pores. Some participants discussed looking for target detail in the core area if at all possible. It was suggested that level two details such as ridge endings or bifurcations will usually stand out before other detail when initially analyzing a print. Though ‘smaller more unique features like lakes and crossovers will always draw the attention of the examiner’ and provide an ‘obvious target point’ from which to begin the comparison process.

Examiners described a preference for working sequentially round a print looking for information. Examiners alluded to seeing fingerprints as ‘maps’. ‘There will be lakes, areas where there is nothing to see, furrows and roads to be followed’. ‘The quality of the map helps to define the direction in which to go’. It was suggested again that the comparison process is like ‘putting together a jigsaw’. ‘It is not necessarily the detail of each individual feature that adds weight, but rather the way one piece of the jigsaw fits in with another piece’. ‘You look for features and shapes’; ’if there was an all black puzzle it would not bother me’ said one examiner, ‘I am only looking for the edges and shapes of the individual pieces’.

Participants suggested that most areas of good detail ‘stick out immediately’, and that ‘it is an automatic process’. Other participants suggested however that they ‘recognize certain patterns and target groups that are more unusual’. Participants continued to assert that they felt that arch patterns presented problems to examiners because the ‘opportunity to focus on target groups of features is more difficult’. Indeed participants discussed spending more time on arches, suggesting that in local proficiency tests that arches take longer to assess than others, because there are no obvious target features.

Cognitive abilities

Examiners interviewed considered the difference between good examiners and less good examiners in a multitude of ways. Abilities discussed included the speed of comparison and speed of decision making, given that it was assumed that both types of examiner will be accurate. It was also suggested that a good examiner 'will be able to interpret limited information', and 'display a balance between risk taking and decision making'. Participants felt that the best examiners can observe many things at one time, for example, one participant knew they could do the job (fingerprint analysis) because they had taken tests in the past that included looking at lists of information and assessing accuracy.

There was some sense that where computer systems generate respondents on AFIS that some better examiners will go right to the comparison on screen, while others want to see other latent finger marks in the case in hard copy format before making a decision.

General Discussion:

The interviews conducted indicated that there are some very interesting skills and cognitive abilities associated with effective fingerprint examination work. It is apparent that there is a perceived need for sound visual acuity, as well as an ability, whether innate or taught that enables examiners to filter out image 'noise'.

It seems feature searching in the fingerprint domain may be guided to categorically-defined targets (Wolfe 1994, Yang et al 2009), that is to say in a fingerprint, the target of feature clusters will usually include 'lakes', 'crossovers' and bifurcations and ridge endings' and will represent the target class of the fingerprint, drawing the examiner in when the examiner begins to assess features for use in any comparison process.

This would explain why examiners often refer to features 'standing out'. It may be that the arch fingerprint patterns, however, have relatively fewer strong 'target' features compared to a whorl pattern and thus the available clustered reference points in the target class of an arch pattern is fewer than those found in a whorl, thus it can be hypothesised that examiners will find it harder to discriminate effectively and home in on features found in an arch fingerprint pattern, compared to features found in a whorl pattern. Certainly the search for such features may take longer in an arch pattern as opposed to a whorl pattern. This is an interesting area for future research.

Signal detection theory as it is commonly called, is a means to quantify the ability to discern between 'signal' (real features) and 'noise' (artifacts caused by distortion and ambiguity). According to the theory, there are a number of determiners of how a detecting system (fingerprint examiner) will detect a signal (ridge detail), and where their individual threshold levels for detecting such signals will be. Changes to the clarity v ambiguity of a fingerprint (noise) will impact upon an examiner's ability to see details within a fingerprint, which in turn will affect their ability to discern patterns and features (Wickens 2002). When the detecting system is a human being, experience, expectations, physiological state (e.g. fatigue) and other cognitive factors like pre-conceived ideas about the value of digital technology and the quality of images such technology produces for comparison purposes can affect the determination of this threshold, making it plastic, thus, a tired examiner or even a demotivated examiner may have an impaired ability to detect a signal relative to a rested or more motivated examiner.

The introduction of automated fingerprint matcher technology appears to be a major influence on fingerprint examiners. Some suggested that ‘hits seem easier because there seems to be far more information in the respondents’. Conversely, some suggested that ‘digital latent images are hard to compare to ten-print cards, as pixelisation causes difficulty in interpreting detail’. The differing attitudes exhibited to the use of AFIS and digital technology and its impact on fingerprint examination may be associated with the ability (or otherwise) to detect the signals within the fingerprint material that would alert the examiner to real fingerprint minutiae. Such ability (or inability) to detect signals in certain patterns of fingerprints when wanting to consider using AFIS would act as a motivator or demotivator depending on their experience in using such systems and impact on attitudes to the use of the technology.

Some examiners felt that where computer systems generate respondents on AFIS that some examiners will go right to the comparison on screen, while others want to see other latents in the case in hard copy format before making a decision. Some examiners suggested finding it hard to use AFIS because ‘you cannot bend and crease an on screen comparison’. It seems tactile interaction with the fingerprint material (perhaps a working copy photograph) is important as it may relate to the ability to match fingerprints more efficiently through the use of mental rotational strategies. Roger Shepard and Jacqueline Metzler (Shepard et al 1971) originally discovered this phenomenon. Their research showed that the reaction time for participants to decide if a pair of items matched or not (perhaps a fingerprint match) was linearly proportional to the angle of rotation from the original position. That is to say in the fingerprint domain, the more a fingerprint has been rotated (perhaps distorted) from the original, the longer it takes an individual to determine if two fingerprint images are a match. In assessing the use of AFIS technology, where the ability to fold and rotate the fingerprint material to facilitate the comparison process is limited, or sometimes impossible, it could be that the efficient use of mental rotational strategy by the examiner is curtailed by the limitations of the technology, thus slowing down the decision making process and demotivating the examiner from the use of such technology.

There was some discussion of the type of latent prints and pattern types that cause most problems to examiners. It was apparent that palm marks were deemed the most difficult type of ridge detail to compare. Arches also caused more problems for finding focus features because examiners found ‘there is no reference point for finding features’, suggesting that when assessing unique features for comparison in a plain arch pattern that features ‘tend to look very similar and the target features are harder to establish’. One examiner discussed how ‘tight whorls and arches tend to slow down the comparative process, there is not enough ridge flow going on to enable easy feature target attribution’…’nothing really stands out’. While one participant felt that arch patterns present problems to examiners because the opportunity to focus on target groups of features is harder, it was also felt that ‘AFIS systems appear to have the same problems also, because there is no focal point’ for the algorithms to home in on. Thus, it seems, the ability to detect signals may be inhibited or degraded in such patterns as arches and palm, not only to the detriment of the human examiner, but also for the AFIS technology also, that may not always be able to effectively cognitise the features within a fingerprint.

In understanding how signal detection (Wickens 2002) may impact upon the fingerprint examiner the relative costs and benefits within the fingerprint domain as fingerprints are compared must be assessed. Accuracy is highly valued, so examiners that perform correct rejections and identify the correct hits will be favoured. The same can be said of artificial intelligence like AFIS matchers that use algorithms to calculate match probabilities. The better, more accurate matchers will survive, while others less accurate will cease to be viable competitors. In an ideal world we would move towards maximum accuracy in both categories, that is to say 100% accurate positive outcomes as well as 100% accurate negative outcomes. In reality this doesn't happen because correctly identifying hits and correctly rejecting false alarms become contradictory goals. At a certain point in order to increase the identification of hits we have to increase sensitivity to the signal, however, this will inevitably lead to a higher degree of risk of erroneous judgments, to picking up false signals. To decrease the risk of erroneous judgments sensitivity must be lowered. So there must necessarily be a trade off between increasing correct identifications of a signal (fingerprints) and decreasing false alarms (misses and erroneous matches) by setting the optimal threshold. It can be argued that the use of AFIS technology in the Brandon Mayfield case, as well as strong contextual and motivational influences on the examiners, contributed to a catastrophic uplift of sensitivity to false signals to a point where the accurate detection of signals became more difficult for the AFIS technology as a result of the algorithms being unable to detect the appropriate and correct signals in a poor quality print, which, when aligned with the human examiners’ inability to also accurately discriminate between signal and noise in the arch fingerprint pattern presented to them, due to the contextual and other top down influences they were subjected to, led to a greater chance of error occurring.

Aligned to the above discussions on fingerprint pattern interpretation and abilities to detect ‘signals’ in the print, it was also discovered that digit determination is more difficult because

‘AFIS takes away the need to be accurate with the determination of finger choice’. ‘AFIS takes away decision making skills from examiners’. It was even suggested that ‘modern examiners are de-skilled’. One examiner stated, ‘AFIS can make you a bit lazy; you are looking at enlarged images all the time’. In addition, some examiners felt that AFIS ‘takes away the ability to make the decision…you are already given the score…people get used to the ident being at the top of the queue’. ‘Examiners don’t have the ability on AFIS to decide the search limits and are willing to accept the decision of AFIS, be it a non match and just stick it on the shelf’. ‘There is very little motivation now to try harder….we have lost ownership of casework’. ‘If you are just doing AFIS work you will not build up the sort of experience you need to make decisions’.

The fact that examiners felt that AFIS technology removed some of the decision making responsibility from the human examiner, which in turn made the human examiner ‘lazy’ is noteworthy. For example, if, as stated, examiners get conditioned to believe that ‘identifications’ are always at the top of an AFIS respondent queue, then does this act as a primer for expectation when examiners are assessing a respondent list of fingerprint exemplars for a potential match against a crime scene mark? That is to say will examiners in these conditions over confidently assess features in noise if the respondent print is in 1st place, and conversely, will respondents lower down the list be assessed less carefully? Dror et al (2012) contributed to this discussion. They presented 3,680 AFIS lists (a total of 55,200 comparisons) to 23 latent fingerprint examiners as part of their normal casework. They manipulated the position of the matching print in the AFIS list. The data showed that latent fingerprint examiners were affected by the position of the matching print in terms of false exclusions and false inconclusives. Furthermore, the data showed that false identification errors were more likely at the top of the list, and that such errors occurred even when the correct match was present further down the list.

The qualitative assessment of fingerprint examiners has highlighted that fingerprint examiners exhibit preferences for certain fingerprint pattern types, which appear to discourage analysis of fingerprints that have ambiguous ridge flow properties and that are hard to mentally conceptualise, such as palm and arch type fingerprint patterns, and that in addition to these concepts around fingerprint utility, they also profess strong views on the relative merits, or otherwise, of the use of AFIS technology. While examiners receive continuous training and development on the best practice usage of AFIS technology in order to maximise the effectiveness of search (for example the best way to manually encode Galton detail on a latent mark), there is apparently little attention, if any, given to the overarching methodology of bureau best practice with AFIS interaction in mind. There is certainly no training given on how AFIS may potentially affect human judgements. Now it will be important to see whether these qualitative findings are replicated empirically in real operational case work.

A Detailed Review of Live Case Work to Understand the Interaction of the Human Fingerprint Examiner with Technology

In order to understanding further how fingerprint examiner search strategies may be affected by the examination of different fingerprint pattern types as well as to better understand the interaction of the human examiner with AFIS technology in an operational environment, it was necessary to conduct a detailed review of live case work to ascertain whether there were any empirical trends that may underpin the conclusions drawn from the qualitative study conducted previously.

Method:

Access permission was granted to an ISO 9001 accredited fingerprint bureau in order that a technical review of live processed casework could be conducted. A total of 213 cases were selected at random over a 6 month period from archive case work from the year 2009. The casework was randomly selected (1 in every 10 cases) from the casework of 13 fingerprint experts who all worked for the nominated bureau in 2009. Each case was then examined closely and fingerprint evidence within each case was analysed to see which search strategies were employed given a particular fingerprint pattern type. All the case work had been independently reviewed to verify that all finger marks considered were both manually comparable with known ten-print exemplars and was of sufficient quality and that they could ultimately have been searched on an AFIS system if desired (that is to say the generally accepted AFIS threshold standard of 8 Galton points).

Results:

The nature of the samples obtained from the live cases meant that it was difficult to draw scientific conclusions about the data and the inferences to be drawn from it. The reasons for this arise from the random nature of the sample data and the inconsistent nature of evidential yield for each case observed, for example one case may have many evidence submissions, whereas other cases may have very few evidential submissions, and even then the propensity of fingerprint pattern types available to the practitioners varied from case to case. However, this investigation was conducted primarily to gather observations that would help to confirm anecdotal assertions made by examiners in the previous qualitative study around fingerprint pattern interpretation and AFIS interaction so as to guide future controlled research opportunities.

Table 17   Percentage of Available Marks by Pattern Searched on AFIS

		% Available Mark Patterns Searched on AFIS				
		Whorl	Loop	Fingertip	Palm	
Expert	A	71.4	25	0	66.6	
	B	63.63	60	25	50	
	C	80	100	0	0	
	D	66.6	75	33	0	
	E	50	66.6	0	0	
	F	100	40	0	35.7	
	G	100	70	25	71.4	
	H	100	60	8.6	50	
	I	37.5	50	0	35.7	
	J	100	80	20	83.3	
	K	0	100	0	35.7	
	L	100	100	0	0	
	M	66.6	75	40	35.7	
				No Data Assumed Average Palm		

An ANOVA analysis was conducted of the different examiners observed (see table 17) to ascertain whether there was any significant impact of the type of fingerprint pattern availability and subsequent search strategy employed on AFIS systems.

While there was no significant effect for examiner search strategy consistency (see table 17) between themselves (p = .45), whereby it appeared that the experts as a group tended to consistently agree in the relative way they preferred to search different finger mark patterns on AFIS. There was however a significant effect noted between different pattern types and searches conducted as a group of examiners toward searching particular fingerprint pattern types in preference to others (F = 17.40, p = < .01, eta = .52). Whorls and loops were clearly preferred for search on AFIS rather than the fingertip and palm patterns. In short, the examiners were unanimous in their liking for whorls and loops and unanimous in their dislike for searching fingertips and palm on AFIS systems. These findings can be observed in the following descriptive statistical analysis of the data.

Observations were made around the percentage of fingerprints searched given the context of knowledge of the crime committed. Major crime cases included such events as rape, murder and arson, whereas volume crime included burglary and vehicle theft. As already discussed, the random nature of the data mining exercise of live casework between major crime and volume crime do not permit us to draw any firm conclusions around the relative percentage of marks searched on AFIS from available finger marks. The 213 cases sampled and reviewed randomly provided more volume crime cases than major crime cases (future studies could look to sample equal numbers of volume crime and serious crime cases).

If the observed data (see Figure 44) whereby fingertips were consistently discarded in favour of loops and whorls, were to be validated in future research then this would suggest the possibility, at least, that the decisions made around the appropriateness or otherwise to search on AFIS may be influenced by the nature of the fingerprint pattern type or ridge flow configuration that when aligned to the seriousness of the case and other contextual influences, could result in important and useful fingerprint material being ignored in not just lower level criminal activity, but also in more serious criminal events such as rape and murder. This is important, not just because incriminatory evidence could be ignored and discarded, but, perhaps just as importantly, the potential for ignoring and discarding exculpatory evidence is highly likely as well.

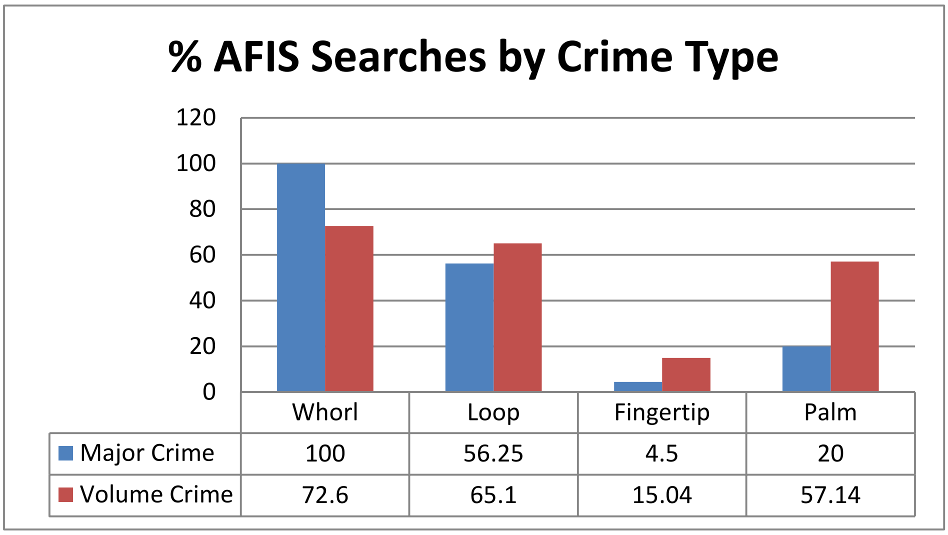
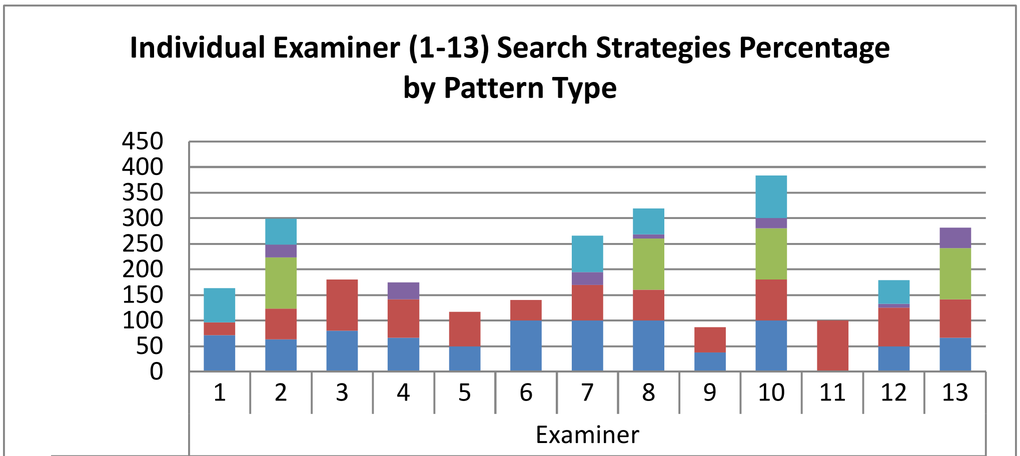


Figure 44 The percentage of AFIS searches carried out by examiners in instances of volume crime v major crime events.

Further research (which may involve other case review studies where an equal number of volume crime cases and major crime case could be observed) is required to see whether it can be shown that there is a strong correlation between major crime and volume crime types and the propensity to search different fingermark patterns on AFIS systems, and what impact, if any, cognitive technology has on the search strategies of fingerprint examiners. Whether the crime is based on volume, low level crime such as vehicle theft or whether it is serious crime such as murder, the tendency to avoid searching fingertips on AFIS systems is consistently observed in comparison to the preference by fingerprint examiners to search whorls, loops and palm.

It can be assumed from the data analysed above that the consistency of search approach employed by individual examiners is consistent in that there was broad consensus about the preference for searching whorls and loops on AFIS systems. This is further e videnced below (see Figure 45).



	1	2	3	4	5	6	7	8	9	10	11	12	13
Palm	66.6	50	0	0	0	0	71	50	0	83.3	0	45.5	0
Fingertip	0	25	0	33.3	0	0	25	8.6	0	20	0	8.3	40
Tent	0	100	0	0	0	0	0	100	0	100	0	0	100
Loop	25	60	100	75	66.6	40	70	60	50	80	100	75	75
Whorl	71.4	63.6	80	66.6	50	100	100	100	37.5	100	0	50	66.6

Figure 45 Individual examiner (1 to 13) search strategies percentage by pattern type

While there appears to be some degree of inconsistency of search strategy between examiners with regard to individual pattern types (see Figure 45), for example some examiners search whorls seemingly all the time, while others will search whorls 50% of the time, it is apparent that when considering the relative weighting apportioned to each pattern by all examiners, some apparent trends do stand out. Of all the pattern choices observed, it was the fingertip, or area of no defined pattern, that was least favoured for AFIS searching. While generally it was the whorl or loop pattern that was most favoured in terms of AFIS search strategy. This is echoed below (see table 18) when the 1st place preference by ranking is considered based on the data above.

Table 18 Examiner's 1st place pattern preferences for AFIS searches (Examiners 8 and 10 expressed equal 1st preferences)

	Ranking of Pattern Preference 1												
Examiner	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**
Whorl	1					1	1	1		1			
Loop			1	1	1				1		1	1	
Tent		1						1		1			1
Fingertip													
Palm													

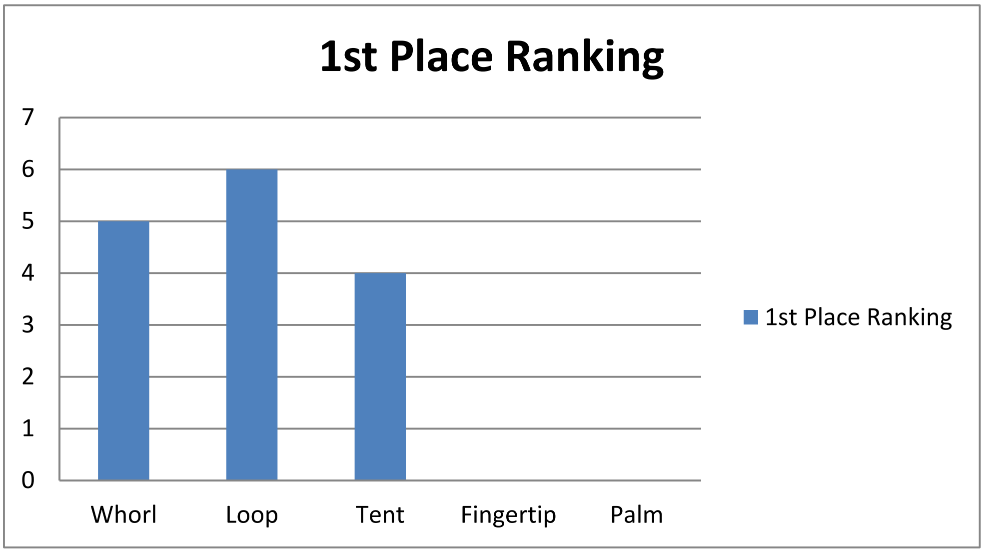


Figure 46 Examiner's 1st place pattern preferences for AFIS searches

The observations in Table 18 and Figure 46 graphically illustrate the apparent preferences of fingerprint examiners when adopting choices for searching on AFIS systems. Based on the data analysed thus far of live case work analysis and through directly talking with practitioners, it can be hypothesised fingerprint examiners have a strong preference to search for Whorls and other clearly defined patterns on automated systems rather than configurations of friction ridge skin that lack clearly defined patterns or ridge detail that provide a target for the human examiner to focus on. In addition, the context of the presence of cognitive technology as well as case knowledge may be operationally impacting upon the cognitive processes of fingerprint examiners as they apply ACE-V in the operational setting. The next study conducted attempted to confirm the hypothesis that examiners prefer loops and whorl type patterns over arch and fingertip patterns, not just as a result of AFIS interaction, but based purely on their observational preferences based on quality determination. A novice examiner control sample was recruited to explore whether there was any similarity between the intuitive opinions of a novice over the experienced opinion of a trained expert.

The Relationship between Fingerprint Pattern and Difficulty Appraisal by Fingerprint Examiners

Following on from earlier investigations in this chapter, this experiment was conducted in order to better understand the relationship between fingerprint pattern and the interpretation by examiners of relative difficulty associated with the different fingerprint patterns observed. It can be hypothesised from assessment of the operational case work, as well as the views expressed by examiners earlier in this chapter that fingerprint examiners least favour arches, finger tips and ambiguous ridge flow associated with palm when assessing preferences for AFIS search and comparison. The next investigation looks at the way fingerprint examiners interpret relative difficulty associated with different fingerprint patterns. In a controlled experimental situation it was hoped to replicate the inferences drawn from the earlier investigations in this chapter. Namely that by asking examiners to place quality determinations on the various fingerprint patterns put before them that they would instinctively place a higher difficulty tariff on arch and fingertip patterns. In addition, it was hoped to further understand whether culture and experience has any influence on these decisions. For this reason a control group of novices was assessed to see whether the experts applied different search difficulty criteria to that of novices.

Method:

In this experiment, 16 fingerprint examiners where selected from different fingerprint Bureaux, in addition to 20 novice control participants taken from the student population at Bournemouth University. Whether in the expert or novice control part of this experiment, each participant was presented with a booklet of 20 fingerprints that comprised a total of 5 examples of a particular pattern type. There were 5 arches, 5 whorls, 5 loops and 5 examples of where there was indeterminate pattern and where only ridge flow was present (fingertips). Each of the 20 examples was independently assessed by a fingerprint expert who was not part of the study, and was deemed to be of sufficiently high quality for search and comparison purposes against suspects with a view to making an individualisation.

Figure 47 Example of Whorl

Figure 48 Example of Arch (the Brandon Mayfield erroneous latent from the Madrid train bombing)

Figure 49 Example of loop

Figure 50 Example of a fingertip

Each set of 20 fingerprints were randomly ordered for presentation to different participants so that no two participants were presented with the same order of fingerprints. This ensured sufficient counterbalancing to avoid any presentational or experimental biasing (Martin 1996). Participants were asked to assess each of the finger marks and to provide a ranking score on their assessment of the quality of the finger mark with regard to suitability to compare against a potential suspect's fingerprints. A scale of 1 to 5 was employed, 1 being the highest quality and 5 being the lowest quality.

One of the finger mark examples, a pattern known as an arch (see Figure 46) that was shown to participants was in fact the latent finger mark from the Madrid train bombing that was erroneously identified to Brandon Mayfield by the FBI. Introducing this example into the experiment was to provide still further understanding of what characteristics of this particular finger mark provided the catalyst for such a catastrophic error.

Results and discussion:

An ANOVA was conducted on the average scores (each data point had a value of 1 to 5, 5 representing the poorest quality and 1 the best quality) provided by all experts and novices (see Figure 51) against all examples within each pattern classification. So for example for all experts looking at the arch patterns (16x5 data points), they on average gave arches a difficulty rating of 2.04.

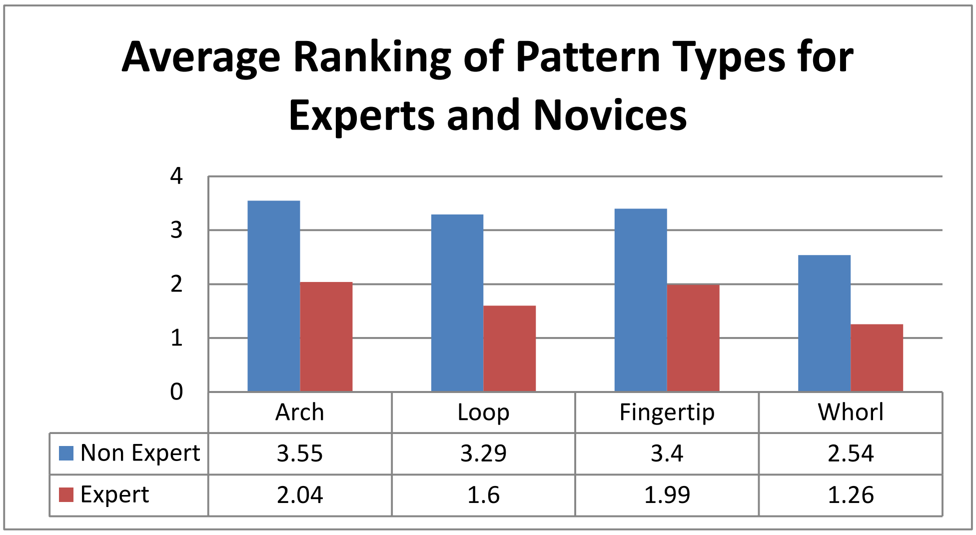


Figure 51 Average scores for each pattern class by experts and novices

There was a significant difference between the relative observations between novices and experts (F = 290.15, p =<.01, eta = .82) in the way they observed the different finger mark patterns. The experts consistently assessed each fingerprint pattern at a lower difficulty calibration than their novice counterparts, but this differential was not uniform across all pattern types. For example, there seemed to be less difference (or closer consensus) between experts and novices when they assessed the whorl pattern. There was also a significant effect (F = 21.46, p = < .01, eta = .18) observed between different fingerprint patterns in that some patterns appeared to have a greater impact on difficulty determination by participants. It was apparent that even though there was sufficient ridge detail to both search and compare individual examples to suspects, the assessment overall by both the expert fingerprint examiners and the novices was that they found the arch and finger tip patterns the most difficult and the loops and whorls the easiest.

Further statistical analysis was conducted on the experts and on the novices separately to assess how they assessed fingerprint patterns as groups of individuals. A two factor with replication ANOVA was conducted for both experts and novices to assess their relative scoring for groupings of arches, loops, whorls and fingertips. Their was a strong interaction noted for the experts against pattern type (F = 8.33, p = < .01, eta = .22) and further interrogation of the data seems to signify and confirm that experts in assessing the different fingerprint patterns scored the loops and whorls far more favourably than their collective assessments of the fingertip and arch patterns (see table 19).

Table 19 Expert analysis of different fingerprint pattern difficulty

SUMMARY	A	B	C	D	E	Total
Arches						
Count	16	16	16	16	16	80
Sum	52	19	16	33	43	163
Average	3.25	1.1875	1	2.0625	2.6875	2.0375
Variance	2.466667	0.1625	0	1.129167	1.695833	1.783386
Loops						
Count	16	16	16	16	16	80
Sum	31	34	25	16	22	128
Average	1.9375	2.125	1.5625	1	1.375	1.6
Variance	0.8625	1.316667	0.6625	0	0.25	0.749367
Fingertips						
Count	16	16	16	16	16	80
Sum	23	36	38	34	28	159
Average	1.4375	2.25	2.375	2.125	1.75	1.9875
Variance	0.395833	1.133333	1.583333	1.583333	1	1.202373
Whorls						
Count	16	16	16	16	16	80
Sum	25	17	25	18	16	101
Average	1.5625	1.0625	1.5625	1.125	1	1.2625
Variance	0.6625	0.0625	0.395833	0.116667	0	0.29731

In addition, when the data of the novices was analysed there was a striking similarity in the findings, namely a strong interaction between pattern types (F = 13.77, p = < .01, eta = .26) which also mirrored the assessment of the expert examiners in that the novices too, favoured the whorl and loop patterns to the arches and fingertips. These findings were consistent with the data assessed in the study of the live casework, which also suggested that expert examiners had distinct preferences for whorl and loop patterns over arch and fingertip patterns (see table 20).

Table 20 Novice analysis of different fingerprint pattern difficulty

SUMMARY	A	B	C	D	E	Total
Arch						
Count	20	20	20	20	20	100
Sum	92	45	56	87	75	355
Average	4.6	2.25	2.8	4.35	3.75	3.55
Variance	0.252632	1.25	1.326316	0.555263	1.25	1.704545
Loop						
Count	20	20	20	20	20	100
Sum	70	64	76	52	67	329
Average	3.5	3.2	3.8	2.6	3.35	3.29
Variance	1.105263	1.326316	1.115789	0.884211	0.871053	1.177677
Fingertip						
Count	20	20	20	20	20	100
Sum	49	86	67	76	62	340
Average	2.45	4.3	3.35	3.8	3.1	3.4
Variance	1.313158	0.536842	0.765789	1.115789	1.147368	1.333333
Whorl						
Count	20	20	20	20	20	100
Sum	64	54	50	61	25	254
Average	3.2	2.7	2.5	3.05	1.25	2.54
Variance	1.010526	1.378947	1.210526	1.102632	0.302632	1.442828

Next, the data for the arch pattern group only was scrutinised for both the experts and novices. The arch pattern grouping included the Arch (A1 see figure 52) that was in fact the Brandon Mayfield mark erroneously identified by the FBI.

Figure 52 Arch A1 (Mayfield mark)

In the novice group ANOVA analysis revealed that the arch examples presented to the novices were rated with a variety of difficulty ratings for each of the individual fingerprint arches ($F = 23.92$, $p = < .01$, eta = .48). The same can be observed for the expert group too ($F = 23.74$, $p = < .01$, eta = .42). Interestingly in both instances, experts and novices alike both assessed arch A1 as the most difficult in terms of quality with a view to comparing against a suspect. This is interesting as mark A1 was the Brandon Mayfield mark. Strikingly, while the novices as a group showed no statistically significant differences between their relative consistency as individual examiners in assessing the arch pattern finger marks ($p = .11$) , the experts, in contrast exhibited inconsistency from one expert to the other in terms of their rating for the different arch patterns ($F = 4.76$, $p = < .01$, eta = .32). In other words, even with training and experience, the ability for fingerprint experts to assess the arch pattern consistently was seemingly less reliable as a group than that of the novices. Analysis of the other pattern types did show a significant inconsistency between the novice participants (whorl eta = .26, loop eta = .32 and fingertip eta = .30). However, the expert examiners were similarly inconsistent as a group when assessing the other pattern types for quality (whorl eta = .25, loop eta = .49 and fingertip eta = .55). Given the size of the effects observed though, it would appear on balance that the experts, if anything, are somewhat more inconsistent within their group than that of their novice counterparts.

The inconsistency between the experts is worrying and may indicate cultural differences between different laboratories or even different laboratory teams as well as other factors such as training and assessment criteria. This should be investigated further in future research.

Discussion:

While the experts and novices agree in the overall rating of what both groups perceive as either easier or more difficult pattern types, it is apparent that the expert examiners seem to have a lower calibration of their assessment of finger mark difficulty in that consistently the experts rated all categories of fingerprint pattern type at a lower difficulty level. This phenomenon may have much to do with the very nature of the examiners experience and expertise. Examiner expertise may result in more confident and efficient decision making, resulting in an overall lowering of the difficulty calibration for each fingerprint assessment.

A number of computational models have been developed in cognitive science to explain the development from novice to expert. In particular, Herbert Simon and Kevin Gilmartin proposed a model of learning in chess called MAPP (Memory-Aided Pattern Recognizer). Based on simulations (Simon et al 1973) they estimated that about 50,000 chunks (units of memory) are necessary to become an expert, and hence the many years needed to reach this level. More recently, the CHREST model (Chunk Hierarchy and REtrieval STructures) has simulated in detail a number of phenomena in chess expertise (eye movements, performance in a variety of memory tasks, development from novice to expert) and in other domains (Gobet et al 2004). So given that the fingerprint experts were able to rationalise their assessment of difficulty for each fingerprint pattern at a lower scaling threshold to that of novices may be because the ability to rationalise difficulty was associated with the computational power required by the human to reach a conclusion, that is to say it could be that the expert examiners reached their conclusions more quickly and efficiently than the novices (who also reached the same conclusions overall) because they were able to utilise cognitive shortcuts to draw upon previous experiential learning. The expert examiners could work more efficiently because they were able to draw upon memory and experiences, using 'chunks' of knowledge, rather than using all available cognitive power to reach conclusions.

An important feature of expert performance seems to be the way in which experts are able to rapidly retrieve complex configurations of information from long-term memory. They recognize situations because they have meaning. But because they have retrieved this information from a 'silo' of memories and experiences in order to be more efficient cognitisers, this also means that they are unable to recall the more detailed rules and regulations around how they reached conclusions. For example, to drive a car you are an expert motorist, you perform tasks such as gear changes automatically. However, if you ask a motorist to describe in detail their drive to work, they will find recalling such small details as changing gear very difficult, if not impossible to do. The mind is not constructed to process vast amounts of information simultaneously and develops coping strategies to enable humans to function. If one was to ask a fingerprint expert how they came to a conclusion, such as 'how do you know this is an identification?' perhaps the answer may be, 'because it just is'. If the examiners who took part in this study were to be asked 'how are you able to assess these fingermarks as relatively more easy as non experts?', a similar answer may be forthcoming, namely 'I just can'. While this may seem an unsatisfactory set of answers, perhaps this is to be unfair to the fingerprint expert as this is a true reflection of how expertise is categorised. There is increasing evidence that experts recognize situations based on experience of many prior situations. They are in consequence able to make rapid decisions in complex and dynamic situations. It is stated by Dreyfuss that if one asks an expert for the rules he or she is using, one will, in effect, force the expert to regress to the level of a beginner and state the rules learned in school. Thus, instead of using rules they no longer remember, the expert is forced to remember rules they no longer use (Dreyfuss et al 2005).

The nature of expertise is that experts are more efficient and faster cognitisers who are able to draw upon vast amounts of experiential learning in chunks, or 'silos' of information held in long term memory. Novices use more cognitive power in the same domain which results in slower processing speed potentially in order to reach final conclusions, even if the overall decision and outcome may be the same. In the domain of fingerprint expertise this is very useful as experts can quickly interpret information in a fingerprint and make judgments on utility, pattern type; AFIS search usefulness and even decisions around individualisation. But if these ‘chunks’ of information held in memory come with a ‘context’, such as knowledge around previous decisions and comparisons in the fingerprint domain, then it is not just the positive memories, but also negative memories that will be processed in blocks of data. Thus, it could be that pattern types in fingerprints will be ‘chunked’ and categorised, associated with difficult comparisons or as patterns that do not search well on AFIS systems and thus perfectly useful marks for search and comparison may be inappropriately negatively assessed given the cognitive shortcuts used by the expert examiner and be contextualised, guiding perception and judgment to make rash and sometimes false assumptions.

Experience and expertise can be characterised by the use of cognitive shortcuts to speed up task processing. More efficient task processing based on experience and prior knowledge leading to speedier decisions may explain the relatively lower difficulty ratings by experts as compared to novices in assessing the different fingerprint patterns. However, it is not so clear how the experts and novices appeared to have differently calibrated scaling thresholds, but yet overall the decision outcomes were broadly the same, namely the relative dislike of arches and fingertips.

It can be hypothesised that the nature of expert as opposed to novice task processing in the fingerprint domain may also be dictated by relative ability to confidently detect the ‘signals’ (features) in fingerprints against a background of ‘noise’ (distortion and superimposition of fingerprints for example). Signal detection theory as proposed by Green and Swets in 1966, used the signal detection paradigm to test an observer's ability to discriminate ‘signals’ from ‘noise’ by responding "yes" to signals and "no" to noises (Swets 1966). In signal detection, any uncertainty that tends to make decisions difficult is called "noise" (Marteniuk, 1976). So it would seem that fingerprint experts and novice examiners alike both perceive the ‘noise’ in fingerprints that indicate to both expert and novices that some fingerprint patterns are more difficult than others, that is to say appear to have properties that make the detection of signals more problematic. But it must be asked, why are the fingerprint experts applying a lower threshold for noise (difficulty) than the novice examiners?

There have been several studies conducted on signal detection in relation to sport (Allard & Starkes, 1980; Starkes, 1987; Starkes & Allard, 1983). Allard and Starkes (1980) were the first to use the signal detection paradigm to analyze perceptual skill in volleyball players because a player must be able to ignore attack patterns of the opposing team. This experiment examined both volleyball players (experts) and non-players (novices). Both sets of subjects were compared for speed and accuracy in detecting the presence of a volleyball in a rapidly presented slide of a volleyball situation. Players and nonplayers did not differ in accuracy of response, but players were much faster in responding for both game and non-game situations. If we apply the rationale behind this study to research within the context of fingerprint examination, then it can be hypothesised that fingerprint experts and novices will come to the same conclusions (accuracy), but that the fingerprint experts will be faster and perhaps perceive that their task in relation to novices easier, thus accounting for the lower threshold for difficulty (noise). Such speculations are interesting but further research in signal detection in the fingerprint domain is required to ascertain if different fingerprint pattern types provide specific challenges to cognitive processes and the ability to detect signals within them. By understanding the nature of expertise, the nuances of signal detection theory and memory and cognition it may help to formulate better training and recruitment strategies and help predict which novices may be more suited to fingerprint examination tasks than others. Ongoing research will help to address some of these issues.

Expert and novice differences may be the keys to unlocking the "black box" of information processing in the fingerprint domain. It is through examination of what separates experts and novices that researchers are provided with an indication of how these groups differ in the different stages of information processing. Gaining a better understanding of these processes can ultimately lead to a better understanding of ourselves as human operators, and how we interact in the environment. This information could be especially useful in the environment of fingerprint examination.

It was noteworthy that both the expert fingerprint examiners and the novices both assessed the arch ‘A1’ as the most problematic and the mark considered most difficult within the arch class of prints examined. It is important to note that mark ‘A1’ was in fact the mark erroneously identified to Brandon Mayfield by the FBI as part of the Madrid train bombing investigation. This indicates that novices appear to have the same intuitive ‘gut’ feeling about the poor visual quality of the Brandon Mayfield mark as the expert examiners. Signal detection research in the fingerprint domain may provide clues as to why it is fingertips and arch patterns, specifically the Brandon Mayfield mark erroneously identified by the FBI cause fingerprint experts to assess these ridge detail configurations more unfavourably than others, and may provide clues as to whether there are some examiners who are more adept at detecting signals within the prints observed that make them more reliable examiners and perhaps more immune to visually challenging fingerprint distortions that might fool an examiner into making an incorrect judgment. Only by understanding the way the human mind interprets and formulates conclusions around visual pattern recognition, will it be possible to design and implement robust and reliable training and recruitment tools that will not only calibrate expert performance (providing that all important feedback) but also facilitate valuable predictors of which novice examiners may become the best examiners of fingerprints in the future.

One of the cornerstones of the human cognitive system is that it has limited computational capacity and resources, and therefore, it cannot process all the information that is provided as sensory input to the brain (Knudsen, 2007; Posner, Snyder, and Davidson, 1980; Sperling, 1960). The result is that the cognitive system selectively allocates attention and cognitive resources. This process requires us to constantly engage and disengage attention and shifting attentional focus and cognitive processing to different segments of the visual input (Wright and Ward, 2008; Posner and Petersen, 1990).

The ability to allocate attention to the important and most crucial information defines much of the human cognitive system and intelligence. This process of selectively and wisely knowing where to focus attention gets better and more refined with expertise (Dror, 2012). Selective attention is one of the cognitive ways of achieving expert level performance (e.g., Wood, 1999). As one becomes a greater expert, they get more selective, paying only attention to the important pieces of information, at an ever increased rate and accuracy.

While a novice is still trying to absorb the information and make sense of it, the expert has already focused on the critical information (e.g., Valk and Eijkman, 1984), processed it, and solved the problem. The process of selection is critical. For example, expert radiologists selectively process X-ray films according to clinically relevant abnormalities (Myles-Worsley, Johnston and Simons, 1988; Valk and Eijkman, 1984). This results in efficient and effective processing. This explains why the experts and novices may have reached similar conclusions, but because the experts had more knowledge, they were able to perform tasks more quickly potentially and thus perceive them easier? This will require more research to better understand this assertion.

This experiment has confirmed the hypothesis derived from the first two investigations that fingerprint examiners not only intuitively feel (through conversation) that arches and fingertip patterns provide the most challenging fingerprint comparisons. But that this is now both validated in empirical studies of live data and through a controlled study where it was discovered that while expertise and experience clearly provide more confidence and a lowering of difficulty ratings for fingerprint assessment tasks, the overall ‘gut feeling’ for the difficulty of a fingerprint is mirrored by experts and novices alike. This may indicate that in developing ongoing recruitment and training tools that if such tools can be tested and validated on fingerprint examiners, then it should serve as a robust predictor of potential pattern recognition ability in the novice population. Next it will be important to conduct a further investigation to now understand the relationship between fingerprint pattern and AFIS search strategies between experts and novices.

The Relationship between Fingerprint Pattern and the Consideration by Fingerprint Examiners to Search on AFIS

This investigation was conducted in order to better understand the relationship between fingerprint pattern and the interpretation of multiple fingerprints within casework as consideration is made to search against an AFIS system. To further understand whether culture and training has any influence on these decisions, a control study was also conducted to analyse relative data generated between expert fingerprint examiners and novices.

Method:
Fifteen fingerprint examiners where selected from different fingerprint Bureaux, as well as 20 novice control participants taken from the student population at Bournemouth University.

Whether in the expert or control part of this experiment, each participant was presented with a booklet of 5 fingerprint sets, each of which comprised of a total of 4 basic pattern configurations. In each set of fingerprints presented there was 1 arch, 1 whorl, 1 loop and 1 example of where there was indeterminate pattern and where only ridge flow was present. As in the previous experiment, each of the examples was independently assessed by a fingerprint expert who was not part of the study, and was deemed to be of sufficiently high quality for search and comparison purposes.

Figure 53 An example presentation set which includes the arch (furthest left) which was erroneously identified to Brandon Mayfield by the FBI as part of the Madrid bombing investigation

The stimuli sets were ordered into packets of 5 for presentation to each of the different participants. The order of pattern type on each presentation set was changed within the packets so that not only were the order of pattern sequence different for each trial, but also the order in which these trials were presented to participants was similarly randomised so that no two participants were presented with the same order of stimuli. This ensured sufficient counterbalancing to avoid any experimental biasing.

Participants were asked to assess the stimuli presented to them (see example in figure 53) and to provide a ranking score on their assessment of their search preferences with regard to the different fingerprint patterns. A scale of 1 to 4 was employed, 1 being the highest search preference and 4 being the least favoured for AFIS search. One of the arch fingerprint examples within the trial sets was in fact the latent print from the Madrid train bombing that was erroneously identified to Brandon Mayfield by the FBI. Introducing this example into the body of the experiment was to provide still further understanding of what characteristics of this fingermark contributed to the FBI error.

Results:

An ANOVA was conducted on the average ranking scores (each data point had a value of 1 to 4, 4 representing the least preferred pattern type for search on AFIS and 1 the most preferred) provided by all experts and novices against all examples within each trial set. So for all experts looking at the arch patterns (15x5 data points), the average ranking within the trials put before the participants was 2.82 (see table 21).

Table 21 Average ranking score for each pattern type by experts

	Arch	Loop	Tip	Whorl
Experts	2.82	2.27	3.09	1.58
Novices	2.92	2.49	2.85	1.74

There was a significant difference between the relative observations by novices and experts for different pattern types: ($F = 33.77$, $p = < .01$, eta = .97). As with previous findings earlier in this chapter, there was a strong disposition for all participants to favour the whorl and loop patterns for search on AFIS while both experts and novices rated the fingertips and arch patterns least favourably. There was no significant effect observed ($p = .6$) for relative differences between expert and novice in so far as if one group found a pattern to be favoured then the other group did also to the same relative degree and vice versa. This implies a consistent level of observation and decision making between the experts and the novices in their estimation of relative pattern preference in this task.

This is further evidenced by assessing the between expert rating for different pattern types. ANOVA revealed no significant effect for how different fingerprint experts rated different pattern types overall (Arch, Loop, Whorl and Fingertip = $p > .05$). This was not replicated by the novice participants who while there was no significant difference between novice examiners for the loop, fingertip and whorl patterns, specifically displayed between examiner inconsistencies in the way they apportioned preference ranking for the arch patterns (F = 2.15, p = < .01, eta = .15). The broad consensus between expert examiners and the other novice participants seems to indicate that the intuitive feelings of the novices was just as valid as the perceived expert opinion of the more experienced examiners.

Table 22 Order of preference for AFIS search by expert and novices.

Order	Pattern
1	Whorl
2	Loop
3	Arch
4	Fingertip

In assessing the data (see table 22 and 23), the apparent correlation in pattern rating between experts and novices is noteworthy in that for each combination of trials placed before experts and novices their assessment of what they would least want to place onto an AFIS system to search was identical. This was even the case in sample set 'TWAL' where experts and novices both rated the fingertip and loop equally as poor in ranking order for search on AFIS (see table 22 and 23).

Table 23 Matrix of most favoured ranking and least favoured ranking by pattern type for experts and novices.

Experts	Most Favoured AFIS Search	Least Favoured AFIS Search
ALTW	Loop	Arch
LTWA	Whorl	Fingertip
TWAL	Arch	Fingertip and Loop
WALT	Loop	Arch
ALTW	Whorl	Arch
Novices	Most Favoured AFIS Search	Least Favoured AFIS Search
ALTW	Fingertip	Arch
LTWA	Arch	Fingertip
TWAL	Whorl	Fingertip and Loop
WALT	Loop	Arch
ALTW	Whorl	Arch

Figure 54 Example presentation sequence Arch (Mayfield), Loop, Fingertip, Whorl

An ANOVA confirmed that there was a strong interaction between the different pattern groups of arch, loop, fingertip and whorl (F = 83.69, p = < .01, eta = .74) and that there were clear differences in search preferences by expert examiners dependent upon the pattern type available (see figure 54). This was replicated by the novices (F = 30.04, p = < .01, eta = .71) who similarly had the same preferences and dislikes for searching certain fingerprint patterns in the context of an AFIS search.

A separate ANOVA of the arch pattern data specifically indicated that there was a significant difference in how experts ranked the different arch patterns (see figure 55) (F = 210.62, p = < .01, eta = .92) during this experiment and that strikingly, nearly all experts (14 out of the 15 who participated) rated the arch (A1 the Mayfield mark) as the least favoured arch pattern that they would wish to search on an AFIS system.

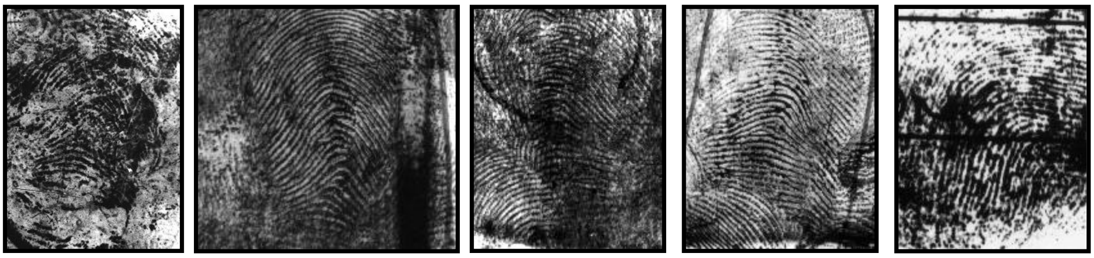


Figure 55 Arch Patterns A1, A2, A3, A4 and A5 (A1 is top left is the Brandon Mayfield mark)

Specifically, given a hypothetical case where the four pattern choices presented included an arch pattern, and the Brandon Mayfield mark in particular, not only were arches one of the least favoured AFIS search choices, but that also, it was clearly seen by experts that the Mayfield mark was the least favoured mark in its presentation pack. When this is combined with earlier data analysis on relative difficulty, where the Mayfield mark was considered the most difficult arch encountered for suspect comparison purposes, perhaps the assessment that the Mayfield mark was also the least popular for AFIS search is not unsurprising. However, this raises interesting, and alarming questions on the utility of the Mayfield mark for either suspect comparison or AFIS search.

General Discussion Based On the Studies within This Chapter

To maximise the potential of automated fingerprint identification systems they must be used to full capacity (Charlton 2003). From the data collected and analysed in the interviews with fingerprint experts, the analysis of operational casework as well as the controlled studies of fingerprint experts and novices, it is apparent that fingerprint experts overall prefer to search loop and whorl fingerprint patterns over other fingerprint patterns such as arches and fingertips.

The evidence suggests expert examiners have a preference for searching loops and whorls on AFIS systems and this appears to echo the preferences of examiners for whorls and loops when assessing relative quality ratings for different fingerprint patterns. In short, not only do examiners rate the whorls and loops more favourably but that they also carry that philosophy forward when deciding which fingerprints to search on AFIS technology. Indeed, it is also possible that the cognitive technology of AFIS systems may also exert a biasing context to the decision to search or not to search based on the expert examiner's prior knowledge of AFIS technical capabilities and its potential strengths and weaknesses (Gold 2004, Dror et al 2010, and Dror et al 2012).

The decision by examiners to avoid searching certain types of fingerprints on AFIS systems means by definition that the technology is being underused, which is neither judicially effective, nor cost effective. It would be interesting to further examine this hypothesis in more detail in future research. It is possible that it is the very anticipation by the examiner of interacting with cognitive technology that is also biasing the examiner by influencing the decisions around AFIS search based on prior experiential knowledge, expectations and positive versus negative feedback of prior AFIS usage which in turn will have a motivating or demotivating (Charlton et al 2010) influence dependent upon those experiences. If, as already discussed examiners perceive that technology like AFIS takes away their independence and threatens their very expertise, then use of such systems will immediately be the subject of significant demotivation leading to poor efficiency and performance in using such systems.

Cognitive technology may have an impact on the analysis of finger marks and ultimate impact upon search rational on AFIS systems. But more research is required to better understand the relationships between search strategies for the use of AFIS, and why it appears that patterns such as arches and fingertips are not searched on AFIS in proportion to whorls and loops. One consideration for future research may be to investigate fingerprint expert analysis techniques in more detail and to assess their skills within the domain of signal detection theory (SDT). Signal Detection Theory (McNichol 1972, Coren et al 1994 and Swets 1964) is a way of conceptualising understanding of the human ability to discern between genuine visual stimuli and background ‘noise’. When the detecting system is a human being, experience, expectations, physiological state (e.g. fatigue) and other contextual factors can affect the threshold applied to detecting such signals. For instance, a sentry in wartime will likely detect fainter stimuli than the same sentry in peacetime. In the case of fingerprint analysis, this would equate to the ability of examiners to assess genuine ridge flow against a background of substrate ‘noise’ and other potential distracters which could be caused by distortion and even the matrix used (powders and chemicals etc). Such difficulty in discriminating between ‘bottom up’ interpretation of visual stimuli would be amplified by the presence of ‘top down’ factors such as context in relation to both case details, the presence of suspects as well as biases arising from interaction with cognitive technology.

Another observation from the set of investigations in this chapter was the apparent relative consistency between experts and novices around their preferences for certain fingerprint patterns with regard to AFIS search strategies and overall quality assessments. In both instances novices broadly agreed with their expert counterparts in preferring loops and whorls as opposed to arches and finger tips. It can be speculated that this observed phenomenon may also have its root cause within SDT. Firstly, depending on the rotation within a print, a fingerprint can be more or less difficult to recognize (effect of viewpoint). Secondly, features can be more or less superimposed by other finger marks of ridgeflow from other sources, which can impair detection performance (effect of superposition). Thirdly, the number and type of other finger marks can challenge visual search and processing capacity (Schwaninger et al 2004). Such noise and distortion may make the interpretation and understanding of fingerprint patterns where there is a less well defined ‘target’ or reference point, as in a finger tip of an arch pattern where there is no core and no delta area, that may make it more difficult to discriminate ridge flow and features, making them less attractive for consideration for either search or comparison. Therefore further research should be conducted that may provide more data on SD performance between expert fingerprint analysers and novices to better understand and investigate causal links between SD abilities of fingerprint examiners and pattern complexity in fingerprints.

Motivation can play a large part in decision making and it is possible that preconceived ideas about the ability of AFIS to search for arches or finger tips, as well as experience of comparing such ridge detail on a monitor screen may sway examiners to avoid searching such patterns on AFIS systems. Motivated decision makers can bias their judgments to support desired conclusions about AFIS technology, accuracy or effectiveness (Boiney et al 1997). Clarity and accuracy of object perception has been found to be related to an observer's 'interest' in perceiving them (Vernon 1962). So should an examiner really not want to search finger tips on AFIS, that is have no ‘interest’ in doing so, then the visual perception of such finger patterns will be biased against seeing enough to search in any event.

The scope and size of the set of investigations within this chapter were relatively small, and future studies will need to consider more subtle variations in fingerprint pattern, orientation and noise quality, as well as to conduct more controlled experiments around AFIS search conditions around time pressure and the extent to which searches are persisted with, as well as interactions with Human Computer Interfaces (HCIs) to gain better understanding of cognitive technology and contextual as well as perceptual biases in the fingerprint domain. The findings thus far indicate that such phenomenon are impacting upon fingerprint examiners and effecting overall search efficiency.

The domain of fingerprint examination is becoming increasingly reliant on the use of technological systems in the desire to provide evidence and intelligence to Police investigators in more efficient and less labour intensive ways which ultimately result in more rapid investigative processes. Whether it will ever be possible for technology to completely replace human intervention is debatable. However, given that technology will inevitably play an ever larger role in fingerprint examination as well as other forensic science domains, the question is not so much whether there is a need for human and technology partnerships, but rather where the boundaries between the humans and technology should be and where those interfaces would best reside on a continuum to maximise the effectiveness of the partnership. Fingerprint examination is a complex undertaking that requires skill and ability on the part of the human expert. There are considerable challenges to technology systems in replicating and augmenting such expertise. The term 'cognitive technology' encapsulates and describes how the technology in the work place takes some of the cognitive burden away from the human, where such technological systems interact with the human, in essence doing the thinking for the human actor within a task or process. The interface between where the human and technology share cognitive load is vitally important if the effectiveness and efficiency of such a partnership is to be maximised (Dascal et al 2005). Many technological innovations in the forensic domain have not maximised their potential because they have been designed, developed, and implemented without consideration of the human element (Dror 2006). Indeed, the notion of technology and appropriate human interaction is echoed by a report authored by Dr Carole McCartney et al (McCartney et al 2010) on behalf of the Nuffield Foundation in a report entitled 'The Future of Forensic Bioinformation'. Within this report it is suggested that there are significant 'cognitive issues' for individual examiners in understanding and using 'the operations of different proprietary systems' in the forensic domain.

Technology is better at comparing vast amounts of information efficiently, and it is not susceptible to psychological and cognitive vulnerabilities. However, the data analysis within the studies undertaken here suggest that human intervention, both by experts and novices, all too often results in subjective and arbitrary assessment of fingerprint suitability for AFIS search and comparison utility that prevent the technology from fulfilling its cognitive potential by denying the technology access to fingerprint material that not only is searchable by the technology, but ultimately capable of finding a finger print match where the human declined to act. This has serious implications for not only efficient and accurate forensic science, but also raises ethical and social issues around the extent to which law enforcement is culturally biased to seek out proof of guilt, rather than look for exculpatory evidence which would be potentially available should an examiner search unidentified marks after identifying a suspect on other marks found from a crime scene.

Technology offers many opportunities, but it also has vulnerability. Its inflexibility and passive information processing can often seem counterintuitive to human cognition. It seems then that both human expertise and technology both possess very strong and crucial elements, however, both also have weaknesses.

Dror suggested that rather than conceptualising them as competing and fighting for supremacy, it is better to find the best way for their integration and cooperation. The key to constructing the most efficient and effective system is through understanding the characteristics of human cognition and technology, and then to integrate the advantages that each has to offer. Through the correct balance a mutually complimentary contribution can be achieved (Dror 2005 and 2007).

In the fingerprint domain there have been significant advances over the past 20 years, not only in the use of automated fingerprint matcher systems, but also in the way fingerprint evidence is captured and transmitted to the end user, the fingerprint expert. This technology has brought about improvements to timeliness and can now provide fingerprint matches in seconds that might have taken years previously. Cognitive technology has improved to some degree efficiency and effectiveness in the domain of fingerprint examination. The question really is can such cognitive technology be still further improved to augment current systems by a better understanding of how humans and technology interact so that the interface between the two entities can be placed appropriately along the cognitive continuum (Dror et al 2010) so that both technology and human can co-exist without detrimental cognitive factors having an effect from one to the other.

In UK forensic laboratories and agencies, it would be useful to apply theoretical knowledge around cognitive technology and distributed cognition, to suggest practical ways in which human and technology could co-exist to produce more efficient and effective forensic solutions. For example, if, as it seems based on the evidence within this chapter, that human cognitive vulnerabilities are preventing certain types of searchable finger marks being searched and compared on AFIS systems, then perhaps it would be better to offload the cognitive burden for such searches to the more objective and bias immune technological systems that would be capable of providing an objective assessment of the relative merits, or otherwise of particular finger marks. One such method of enabling the movement of cognition away from the human actor to the technology would be to move to what is termed 'lights out processing'. This is a method whereby human intervention would be negated, hence, no need for lights as no humans are around to do the work. In this scenario, AFIS would search all finger print material irrespective of pattern or quality thresholds, immune to the contextual top down influences such as knowledge of suspects or how serious the crime was. The human would intervene at the point of decision by the technology. For example, an AFIS hit would then require human intervention to 'validate' findings and prepare court evidence etc. Of course, this type of scenario raises still more potential issues around cognitive distribution whereby this action may solve one set of cognitive weaknesses on the part of the human, but may introduce other cognitive biases through overuse and trust in AFIS systems (see Dror et al 2012).

Forensic evidence can now be analysed and communicated to investigating officers very quickly and greatly increases the opportunities of making an early arrest, and ultimately recovering stolen property. The end result for victims of crime will not only be that the suspect is identified, but they may also get back their possessions which can be irreplaceable. The sheer speed and adaptability of the technological system must however, be tempered with considerations of ethical and cognitive issues that impact on efficiencies around the human and technology working in partnership. Not only must the cognitive distribution between human actor and technology be appropriate and efficient, but it must also be seen to be acting in the best interests of due process and justice. This means using technology in ways that maximises the ideals of objective and impartial investigation.

In assessing the data arising from this chapter, it is apparent that while fingerprint examiners generally work to a very high standard and work with best intentions to search crime scene evidence methodically and with dedication, it would appear that both the top down context of the pattern type being analysed, as well as the knowledge around technology and its benefits and possible disbenefits, is having an impact upon the ability of examiners to conduct objective and comprehensive fingerprint searches, and that also, bottom up information processing of visual stimuli, the ability to detect signals, if you will, is somehow inhibited by the nature of the features observed within certain types of fingerprint pattern.

While more research and understanding is required to establish in more detail the root causation of some of the observed phenomenon, it is proposed within this thesis to offer a potential solution to this problem through the use of a new strategy for fingerprint examination that maximises the benefit of the human examiner and the use of cognitive technology such as AFIS systems. It is proposed that fingerprint laboratory processes should be modified to encompass a methodology for working with fingerprint technology which can be termed 'Phased Cognitive Engagement'. PCE can be seen in other domains such as the medical profession. MRI scans, X-Rays, and even ultrasound provide diagnostic and empirical measurements at various stages of the treatment cycle. Such technology carries out cognitive work for the human, be it a nurse or a Doctor and takes on some of the cognitive load of such specialists to enable them to focus more of their cognitive energy on effective treatment and care planning and hopefully to be more efficient and accurate in their ability to care for the patient. Such cognitive technology can also offer objective analysis that helps to highlight mis-diagnosis on occasions and thus can help to formulate human error rates in the medical profession. This is a facet of PCE that would be very helpful in the forensic domain and fingerprint examination in particular, where the known error rate is hard to establish for a variety of reasons. PCE can help developers, fingerprint laboratory managers and practitioners to plan future development of laboratory technology that maximises the potential of evidence as well as maintains human interaction to the benefit of the wider judicial system. For example, PCE as a concept can be introduced in response to the evidence found in this chapter. If it is accepted that examiners are apparently ignoring the very real and usable fingerprint material associated with fingertips and arches that may provide both incriminatory as well as exculpatory intelligence, then it must be the case that humans are failing to maximise forensic search potential. By passing the cognitive burden along a continuum to the technology so that it is the technology that becomes the arbiter of quality and search availability, then this may lead to enhanced search efficiency and improve results to the benefit of the judicial system, and more importantly, provide better justice. This can be done in many operational scenarios in fingerprint laboratories now. For example, in the UK FISH (Forensic Information Scanning Hub) is already deployed to speed up the remote transmission of fingerprints from crime scenes into bureau using wireless technology. Such technology could be developed to facilitate movement of the evidence received via this medium directly into AFIS systems without the need for human intervention so that the cognitive burden for assessing the fingerprints is passed straight to the AFIS system which will then automate the assessment and comparative process (lights out technology). Should a match be forthcoming from the AFIS system, the cognitive burden can be redeployed to the human for them to carry out the task of comparison and evaluation which is better suited, it can be argued, to the human expert system. If the data within this chapter is replicated on a much larger scale set of investigations, then the concept of PCE may not only set a template for the fingerprint community, but perhaps other forensic domains also.

PCE is one potential solution. There needs to be an evaluation of how examiners are recruited into the fingerprint profession. Both recruitment, and also training and accreditation, must be carried out with a view to a cognitive co-existence with not just AFIS technology, but also other digital and technological tools at the disposal of the modern fingerprint examiner.

History has shown that successive Police administrations over the decades have been happy to continue to develop technology without consideration of the wider cognitive implications. The development of FISH as well as AFIS technologies to move to a joined up and unified entity will have to be carried out with the human and the technology working in tandem, each aware of the potential strengths and weaknesses of the other. This will involve not just a partnership between developers and end users, but should also include discussions with training specialists, cognitive psychologists and recruitment personnel to ensure that cognitive technology can work in harmony with humans in the forensic domain and ensure optimum performance, accuracy and efficiency in the interests of not just speedy, but proportionate and appropriate justice. It remains to be seen whether the criminal justice system is capable of using technology 100% safely (Charlton 2005) given the need for much more research to ensure that not only is AFIS fit for purpose, but that the humans who use such technology, as well as the processes that evolve to accommodate human and technology interaction are similarly fit for purpose.

Chapter 8:

Scientifically Based Management and Standards

Many disciplines in forensic science rely on the interpretation of evidence by the forensic examiner themselves. In forensic evidence, such as fingerprinting, firearms, bite and tyre marks, handwriting, hair, bloodstains, and shoe prints, the forensic 'instrument' is to a large extent the human expert who observes visual patterns and makes, by and large, subjective judgments of similarity or dissimilarity based on the visual presentation of that evidence within a law enforcement or laboratory setting. Human decision making plays an important part in many professional domains, including medical practice, the military, aviation, and law enforcement. Such professional domains have been engaged in years of research to scientifically guide and improve their work and performance. If it is assumed that fingerprint experts receive (in the UK at least) the same level of training then it should be the case that examiner conclusions should be consistent (Charlton 2002). But this thesis has demonstrated repeatedly that this is just not the case. There is subjectivity and bias in the profession that cannot be ignored. If we accept that subjectivity is a core trait of the fingerprint examination process then it is important that recruitment and testing tools are developed to not only predict future development of novice recruits and their potential visual abilities but also to minimise the vulnerabilities to cognitive biases.

Researching cognition and decision making that takes place in these various domains has helped to establish scientifically based management and standards. These include, for example, proficiency testing, screening and selection during recruitment, best practices and protocols, and training. All these play a major role in professional domains that rely largely on human cognition and decision making, and can (and should be) based on data, scientific understanding and research.

In contrast to other domains, forensic science lacks such scientifically based management and standards, especially when it comes to human factors and cognition. One example of the inappropriate application of unscientific management standards in the fingerprint domain is the way in which new recruits are recruited and the tests that are used to assess the potential suitability of individuals to perform such tasks. Many of these tests are based on manual visual tests using real fingerprints; some tests involve the use of other psychometric assessments. Primarily though, the overarching basis for using such tests is usually because they 'seem to work' and through word of mouth recommendation such tests are passed from one bureau to the other with scant regard for the scientific validity of the tests or even how appropriate they are to maximise the chances of finding and recruiting those who have the best innate abilities to perform the very important task of matching patterns and recognising fingerprints that may potentially send individuals to prison for life. Even the presentation methodology for such tests may be inappropriate. Many strategies for the recruitment of trainee fingerprint examiners involve the administration of pen and paper exercises that try to assess the relative abilities of potential recruits in their visual acuity and problem solving skills. These include tests that require the person taking the test to match fingerprints or patterns. Even if we put to one side the rights or wrongs of the benefits of such tests to accurately predict abilities (since this study is the first tranche of work within a scientific programme as yet to validate such tests), such tests may have little relevance in the modern working laboratory environment where the vast majority of all bench work is now carried out using automated identification systems that require different aptitudes and cognitive abilities to maximize the use of such technology. Recruitment tests are still rooted in the Victorian era, when they should be designed and modified to represent the modern 21st Century requirements of forensic science. The testing of new recruits in the domain of fingerprints currently makes about as much sense as training a pilot to fly a transatlantic jet airliner in a 1920s two seated bi-plane.

In this chapter it is hoped to lay the foundations for building successful 21st century standards for testing and a framework is proposed for such work to progress. This chapter focuses on how to develop tests for screening and selecting candidates during recruitment. By providing this framework, it will be possible to exemplify an approach that needs to be adopted in other aspects of forensic science, from proficiency testing to best practices. Introduction of such standards is a difficult process; however, it is part of the evolutionary process that will see progressive growth of various forensic disciplines into full scientific entities. This change within forensic disciplines has parallels to the industrial revolution. At first weavers, carpenters, engineers, or fingerprint examiners who produced high quality clothes, rugs, furniture, machines and forensic services performed these tasks predominantly in a manual and piecemeal fashion. With the industrial revolution, as the production of such items and services moved to mass production (in fingerprints this is akin to the development of AFIS systems perhaps), standards had to be developed and tested, processes specified and controlled, production and human resources managed. Thomas Edison produced lamps and electrical systems, Ernst Werner von Siemens developed the dynamo and Joseph Swan also developed lamps and electrical systems. Each electrical system was developed independently and there was no standard electric lamp, or light socket for that matter. Today, standards exist to the point where at least individual countries have the same standard (for example standard lamp socket and base IEC 60061), and where different standards do apply, then at least there are defined standard measures to facilitate interoperability (the travel plug). ‘The adoption of standards marks an important stage in the passage from a scientific novelty to a full commercial product’ (Holmstrom 1947). If we were to apply this argument to the fingerprint domain then it can be seen how important standards are in the evolution of a forensic science such as fingerprint examination from a position of useful but unvalidated novelty, to a finalised and complete entity that is underpinned by both scientific validation and testing.

The result of such a revolution is streamlining, efficiency and accountability, but at the price of losing some of the individual variation (as well as creativity) that existed before. Once systems of standardisation take over, setting standards, controlling and monitoring then the future evolution of scientific method, engineering processes (fingerprint examination processes) etc, are observed within the context of those pre-defined standards. .

As forensic science, and fingerprint examination science in particular, grows and advances, increasing its scientific basis and efficiency, it is important that advances are made in a systematic and scientific way, based on validated research. It is also important that the introduction of those standards is through scientific necessity rather than operational expediency. There is a tendency in severe crimes to lower the standards of identification because it is an important case. This is called the "gravity standard". The seriousness of a case (the gravity of the crime) is used as an argument to put the ruling standard aside. In those cases the normal standard never goes up but always down (gravity). This is not scientific and not objective; it is also with regard to the responsibility to society questionable. One could instead argue that with severe crimes the standard should be higher because the consequences of a possible mistake are higher. The gravity standard is also an example of circular reasoning, "I know it’s him so the standard is too high". Whether this is a fixed empirical standard or a personal one within the holistic approach is irrelevant, to change it with an eye at the nature of the case is scientifically wrong. It is also in conflict with the basic rules of independence and objectivity.

When it comes to human cognition (a critical element in many forensic disciplines) forensic science lacks the expertise and has little-to-no framework of how to develop standards, proficiency testing and management, all based on scientific research and data. These notional standards are too often based on ‘experiential knowledge’, which is very important and insightful, and must play a part in guiding the research, but by itself is not adequate. How to use such experiential knowledge in scientific research, how to develop scientifically based standards, testing, and management, is the topic of this chapter. Rather than addressing this issue theoretically and in the abstract, we can consider a specific example: recruitment. Through this process it will be possible to present the scientific framework forensic science needs in the area of human factors and cognition, so it can advance and prosper.

Scientifically based tests for recruitment

How do you set out to develop scientifically based tests for recruitment and why are they needed? Is the effort and cost worthwhile? Forensic science has always been a popular career option attracting many applicants per position advertised.

Since there are dozens (sometimes over a hundred) applicants per position, there is a significant opportunity for selection and screening to get the best person for the job. Recruitment and selection today is very poor, and definitely not scientifically based, and forensic laboratories often acknowledge that they select people who they later deem to be not the best and most talented for this field (and sometimes totally inadequate). This reflects a huge waste of resources. In the UK, where mandatory training exists through the National Police Improvement Agency, there are requirements for trainees to attend various residential training modules over the life cycle of a three to four year training programme. Each residential course costs in the region of £2,200 plus accommodation to the employer. The residential courses alone cost the employer about £4000 in total. There are three UK mandatory courses amounting to £12,000 in course costs. On top of this must be taken into account loss of productivity of existing experts through trainee mentorship, as well as costs of time and effort of bureau trainers. The estimated cost of bringing a UK officer through training to expert status is approximately £50,000 over a 4 year training lifecycle. If forensic managers could select the right people for the job that would save a great deal of funds at a time when fiscal constraints on public service provision have never been under such close scrutiny. One region in the UK (South East) must find forensic savings of £20 million over the next few years. It can be seen immediately that effective training and retention of new staff can contribute to a substantial saving over time to this forensic budgetary deficit. This is not only because it will minimize the number of people selected and trained, who then drop out of the profession (or worse, they are not suitable, but remain in the profession), but selecting the right people with the right cognitive skills has additional benefits, such as making training easier and faster.

Background research and survey

The first step of developing tests for requirement is to learn what is already available in forensic, as well as other professional domains that primarily rely on human cognition. This consists of extensive background research into this field. A literature review revealed a large number of studies on expertise, cognitive abilities of experts, and how these can be tested and quantified. Below are highlighted and referenced some of these findings, and provide concrete examples of how cognitive abilities are tested and used to help select candidates during recruitment.

Cognitive ability testing has a long history. Over 100 years ago Binet (1905) developed the approach for standardized tests to assess an individual's abilities. The Binet Scale laid down the foundation for IQ (Intelligence Quotient) tests (Linden and Linden, 1968; Kamin, 1995). The American Association for the Advancement of Science listed IQ tests as one of the most significant scientific discoveries of the 20th Century, along with nuclear fission, flight, and the transistor. The Binet Scale was developed to detect children with low cognitive abilities, who needed to be placed in special schools. Standardized assessment tests were then used extensively in World War I to help the US Army screen new recruits (see Figure 56 for an example). The Army Alpha and Beta tests were developed by Yerkes (1919, 1921) to determine the suitability of soldiers to specific job requirements (McGuire, 1994).

The Alpha and Beta tests set out the whole endeavor of scientifically maximizing performance by selecting the right people to do the jobs they are most capable of. The philosophy was that different people had different characteristics and abilities, and these must be taken into account, so as to best match people to what they can do. This, of course requires not only being able to test and quantify certain abilities, but also to understand and specify the abilities that are needed to fulfill specific tasks and jobs. Different professions require different abilities. This is obvious when one considers what distinguishes accountants from interior decorators, but the observation applies to all specialized professions... special abilities enable people to excel in occupations that depend critically on specific mental processes (Dror, Kosslyn, and Waag, 1993).

Such formalization of job requirements, along with tests that measure people's cognitive suitability to those requirements, has been well developed in the military domain where many jobs require very specific skills, (e.g., Guilford, Fruchter, and Zimmerman, 1952; Miller, 1962; Dror, Kosslyn, and Waag, 1993). These are critical for accomplishing missions safely and successfully, which is paramount in the military domain.

The importance of a good fit between cognitive abilities and job requirements has now been recognized and developed in a whole range of domains. From sports (Gary, 2009) and chemistry (Bodner and Guay, 1997; Pribyl and Bodner, 1987; Carter, LaRussa, and Bodner, 1987) to medical surgery and radiology (Caminiti, 2000; Waywell and Bogg, 1999; Westman, Ritter, Kjellin Torkuist, Wredmark, Fellander, and Enochsson, 2006) and graphics and engineering design (Yue 2007).

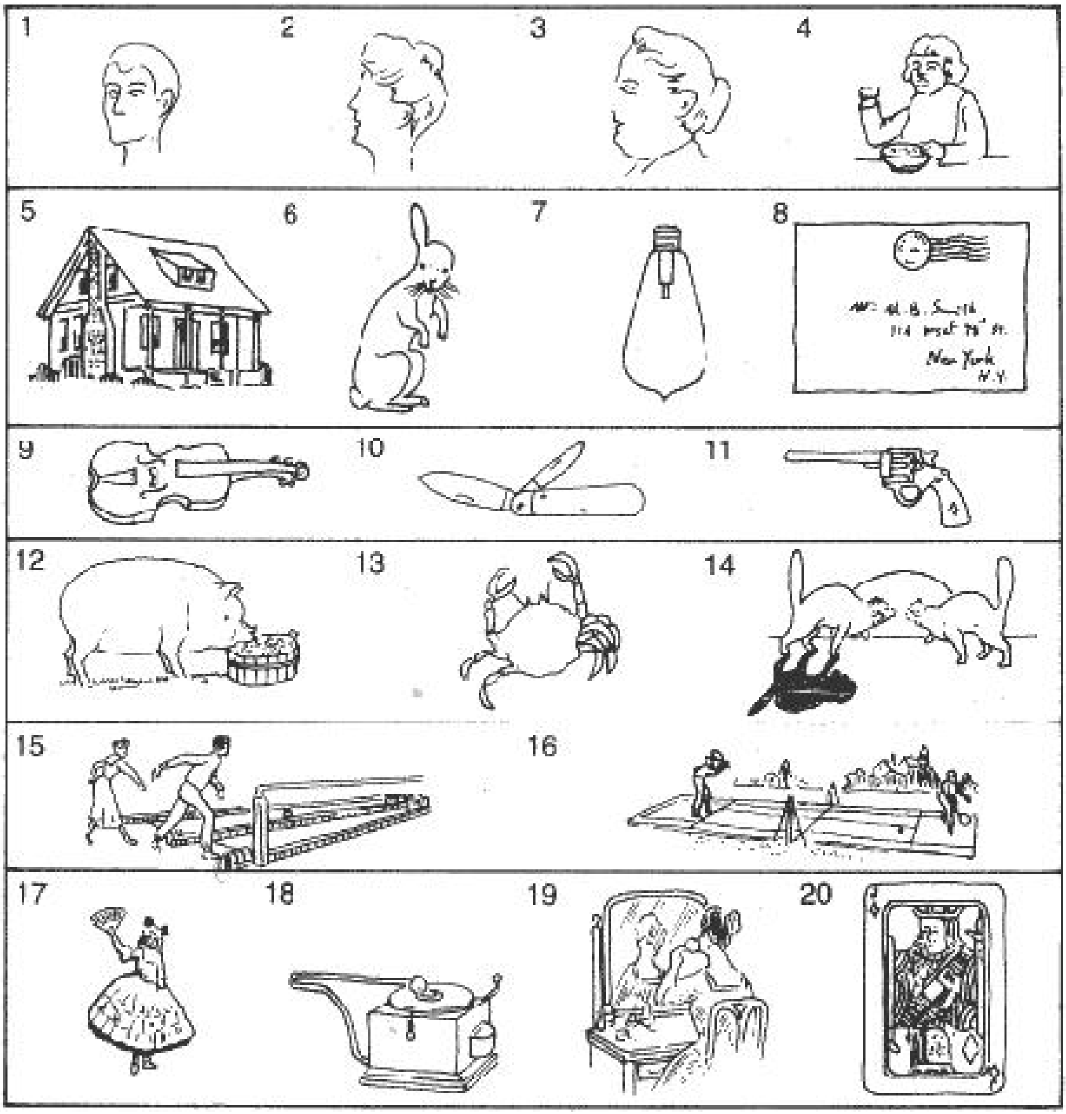


Figure 56   An example from the Army Beta test. In this 'Picture Completion Task' new Army recruits were required to determine what was missing from each picture.

Selecting the right people for the job not only allows the recruit to achieve professional expert performance (including minimizing errors), but also makes training more efficient. It improves both the levels of achievements and the time required for training (e.g., Zamvar, 2004; McClusky, Ritter, Lederman, Gallagher, and Smith, 2005; Cuschieri, Francis, Crosby, and Hanna, 2001). Therefore, it is cost effective because it minimizes the resources needed for training, as well as allows individuals to reach the levels of performance needed with relative ease. From an organizational point of view, it saves expenditure and effort in dealing with people who find it hard to do the job they are required to do, or in worse cases, with people who just cannot perform at the levels required (dealing with the people as well as with the consequences of the backlogs they created and errors they may have made). It is unfair to require people to do jobs they are not able to do. This is as much about a duty of care to the individual as it is about the needs of the manager. A person who is ill suited to a particular role will find such environments stressful and unpleasant. Hence, it is vital to properly screen and select people to the jobs they can do best.

Tests of cognitive abilities must be developed scientifically, objectively, and require proper validation (e.g., Faulkner, Regehr, Martin, and Reznick, 1996; Dolgin and Nontasak, 1990; Martin, Regehr, Reznick, MacRae, Murnaghan, and Hutchison et al., 1997). There are many complex issues in selection of people during recruitment, if one is going to scientifically try to optimize screening practices (e.g., Borman, Hanson, and Hedge, 1997; Borman, Hedge, Ferstl, Kaufman, Farmer, and Bearden, 2003). Some of these issues relate to validity and adequate reliability, and others to learning curves and error analysis (e.g., Ruff, Light and Parker, 1996; Borman et al., 1997, 2003)

This chapter focuses on cognitive abilities needed to reach accurate and reliable conclusions when examining fingerprints. However, there are other issues to consider during the selection of candidates that are not covered in this chapter. These issues include interpersonal skills, motivation, susceptibility to bias, and different aspects of personality (e.g., Merlo and Matveevskii, 2009; Powis, 2009; Dror and Charlton, 2006; Schneider-Kolsky, Wright, and Baird, 2006, Charlton, Fraser-Mackenzie and Dror 2010). However, most of these currently lack reliable tests, relative to quantifying cognitive abilities (Neisser, Boodoo, Bouchard, Boykin, Brody, Ceci, Halpern, Loehlin, Perloff, Sternberg, and Urbina, 1996).

The tests developed in other domains try to explicate and measure specific cognitive abilities needed to perform the certain tasks required in the specific field of expertise. For example, spatial ability testing for astronauts shows that mental imagery, as a subset of cognitive ability, is critical for successfully controlling a robotic arm, during Space Shuttle and International Space Station missions (Menchaca-Brandan, Liu, Oman, and Natapoff, 2007). Research has shown that the cognitive ability to inhibit an ongoing action in response to a signal from the environment is important in baseball batting (Gray, 2009). The SSRT metric (stop-signal reaction time, based on batting inhibition model) captures and quantifies this ability, and therefore can be used for screening and selecting players (for details, see Gray, 2009).

Many such abilities and tests exist in the medical domain, and specifically in surgery, and their importance (as well as their limitations) is discussed in the literature (e.g., Zamvar, 2004; Caminiti, 2000; Tansley, Kakar, Withey, and Butle, 2007; Waywell and Bogg, 1999). Most critical in any such tests is their predictive value for performance at the job (e.g., in the surgical domain, Caminiti, 2000).

However, these examples of baseball batting and surgical competence relate to expertise that is characterized by technical skills, as they require and extensively rely on the ability to perform an action, executing a motor command. If cognitive abilities are important for such expertise and these areas can benefit from cognitive testing at recruitment, then forensic and other domains that are much more cognitively oriented, such as requiring visualization and pattern matching, can benefit to an even greater extent. Take for example, the reliance on spatial visualization in technical graphics and engineering design (Yue, 2007). Even chemistry requires visual cognitive abilities, for example, visualizing three-dimensional structures of molecules from two-dimensional representations (Bodner and Guay, 1997; Pribyl and Bodner, 1987; Carter et al., 1987).

Domains that seem especially close and relevant to forensic impression and pattern evidence (and specifically to fingerprinting) are domains that particularly focus on image perception and interpretation. For instance, X-ray security operators that examine images for forbidden objects. Candidate selection relies on testing applicants during recruitment to measure their cognitive abilities that underpin identification of certain objects. For example, the test offered by Renful Premier Technologies provides “data needed to make informed hiring decisions based on testing candidates’ ability to be an X-ray operator.”

A management tool for use in the recruitment and selection of trainee Air Traffic Control (ATC) candidates is FEAST (First European Air Traffic Controller Selection Test). The test is used in Europe by civil and military Air Navigation Service Providers (ANSPs). The FEAST test package is a professional state-of-the-art web-based testing tool which improves the quality of selection decision making by ATC recruiters. FEAST also contributes to the cost efficiency of the overall recruitment and selection process and to the goal of reducing the costs associated with failure of ATC trainees.

The highly skilled job demands and high training costs, as well as the consequences of incompetent performance and error have made aviation in general, and military aviation specifically, par excellence domains for development of screening tests (e.g., Dror, Kosslyn, and Waag; 1993; Boehm-Davis, and Hansberger, 1997; Street, Chapman, and Helton,1993).

Examples from the Air Force Officer Qualifying Test (AFOQT) and the Marine Corps Aviation Spatial Apperception Ability Test are provided in Figures 57, 58 and 59.

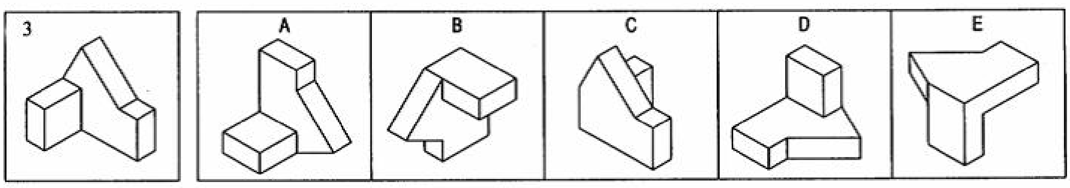


Figure 57 Regardless of orientation, which of the images A to E matches image 3 on the left

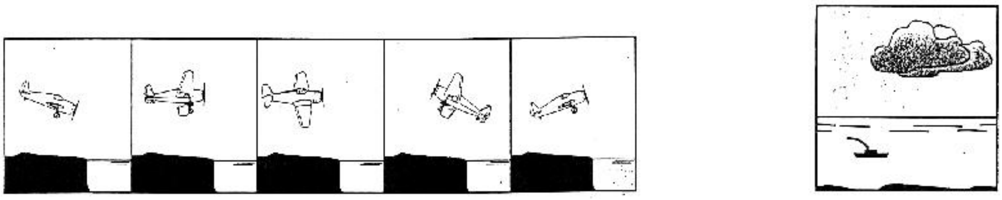


Figure 58 If looking straight ahead through the windshield of an aircraft cockpit and seeing the image on the far left, which of the five aircraft sketches most nearly represents the position or altitude of the plane and the direction of flight from which the view would have been seen.

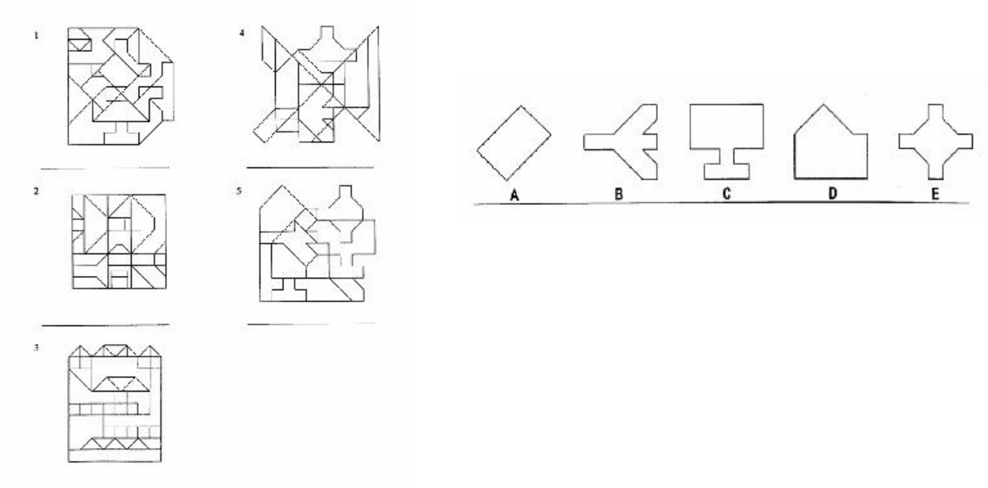


Figure 59 For each of the five shapes (A-E) on the right; determine the shape on the left (1-5) that it is present within.

Research, Survey and Evaluate Existing Tests in Forensic Fingerprinting

As has been seen in a number of professions, testing cognitive abilities enables better screening and selection during recruitment. This in turn produces higher quality experts, with improved job performance and minimisation of error, as well as reducing the need for training. As part of this scientific approach a survey was conducted to look for the existence of cognitive tests (or guideline for them) for use during recruitment of fingerprint examiners. Data was obtained from dozens of sources, including professional bodies and a variety of fingerprint laboratories (varying in size, accreditation and country).

Most laboratories do not use any cognitive testing during recruitment. Many require a university science degree, however, such agencies had no data to suggest that applicants with science degrees are in any way better able to perform fingerprint examination than candidates that had degrees in other domains (such as art or design), or better than candidates who did not have a university degree at all. The professional bodies, such as SWGFAST and the IAI, had no guidelines as per cognitive testing of candidates and how to screen applicants based on their abilities (both in terms of the cognitive profiles needed to excel in fingerprint examination and in terms of how to measure such abilities).

Some laboratories do use some element of cognitive testing. These include use of 'off the shelf' general cognitive testing (see Figure 60 for an example). Some laboratories use actual fingerprint comparisons (sometimes taken from proficiency testing), as well as some cognitive tasks (see 61 for example). Some laboratories do the testing themselves, whereas others rely on commercial companies to conduct testing during recruitment. These commercial companies are either general testing companies that offer cognitive testing to a whole range of clients, or companies that specialize in providing services in the forensic domain, one of which is testing during recruitment.

All these tests fall into two categories: One category contains a whole range of existing general cognitive tests, 'off the shelf'. The other category contains tailor made tests. The problem with the tests in the first category is that these tests do not address and tap into the specific cognitive abilities needed for fingerprint examination. See for example Figure 60, it contains standard cognitive tasks. The first (left panel) measures visual mental rotation (Dror, Ivey, and Rogus, 1997; Cooper and Podgorny, 1976; Shepard, and Metzler, 1971). Although the ability to mentally rotate images may well be important for fingerprint examination, the test used (see Figure 60, left panel) tests three dimensional rotation rather than two dimensional rotations in the plane. The former may not be needed and may even be irrelevant for fingerprinting. Similarly, the other test in Figure 60 (right panel), measures a different type of mental imagery ability (Shepard and Feng, 1972; Shepard and Cooper, 1982). However, here too, this cognitive ability is not relevant to fingerprinting. That is to say while the ability to mentally rotate imagery is important in the fingerprint domain, the tests themselves may not be appropriate to specifically assess these abilities. Hence, although many of the general cognitive tests have been well developed and validated by cognitive psychologists, they are not built and tailored to test the specific and highly skilled abilities needed particularly for fingerprinting expertise.

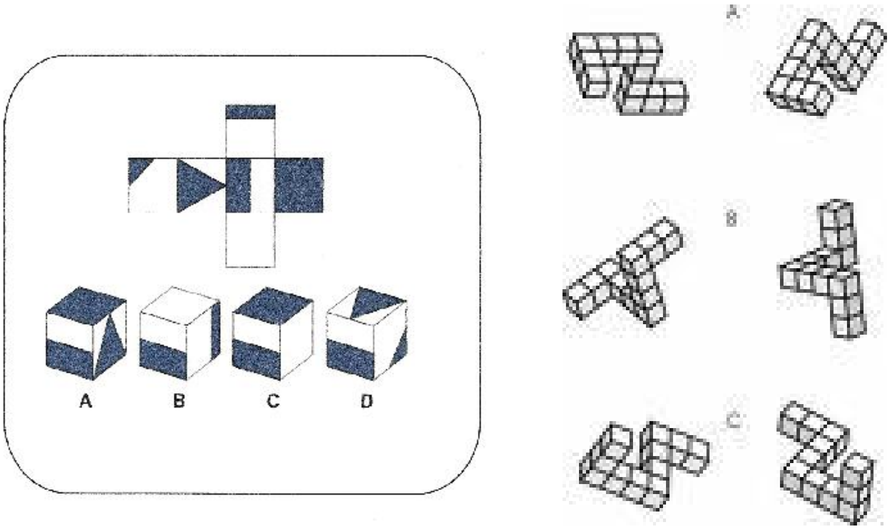


Figure 60 Examples of standard 'off the shelf' general cognitive tasks used for testing during recruitment of fingerprint examiners.

Three-dimensional visual mental rotation (left panel), requires to determine which pair of objects (A-C) are identical, regardless of 3-D orientation. Paper folding (right panel), requires to determine which of the four boxes (A-D) will result from folding the image above the boxes. Both tasks (left and right panels) test different aspects of image mental transformation, which is one process involved in mental imagery. Other processes include image generation, image retention, and image inspection (see details on these processes and ways to measure them in Dror and Kosslyn, 1994).

The second category of tests includes a variety of tasks developed by various fingerprint laboratories and commercial companies that provide forensic services. These tests have not been scientifically developed, nor are they properly validated. Some use fingerprints for the tests, but there has been no objective quantifiable justification for using these specific prints, and it is not clear what is being tested and how best to score the tests. Other, non-fingerprint, tasks are used in some of the testing during recruitment, but it is clear that these tasks are not properly developed to meet minimal standards of testing and validation (see Figure 61 for examples).

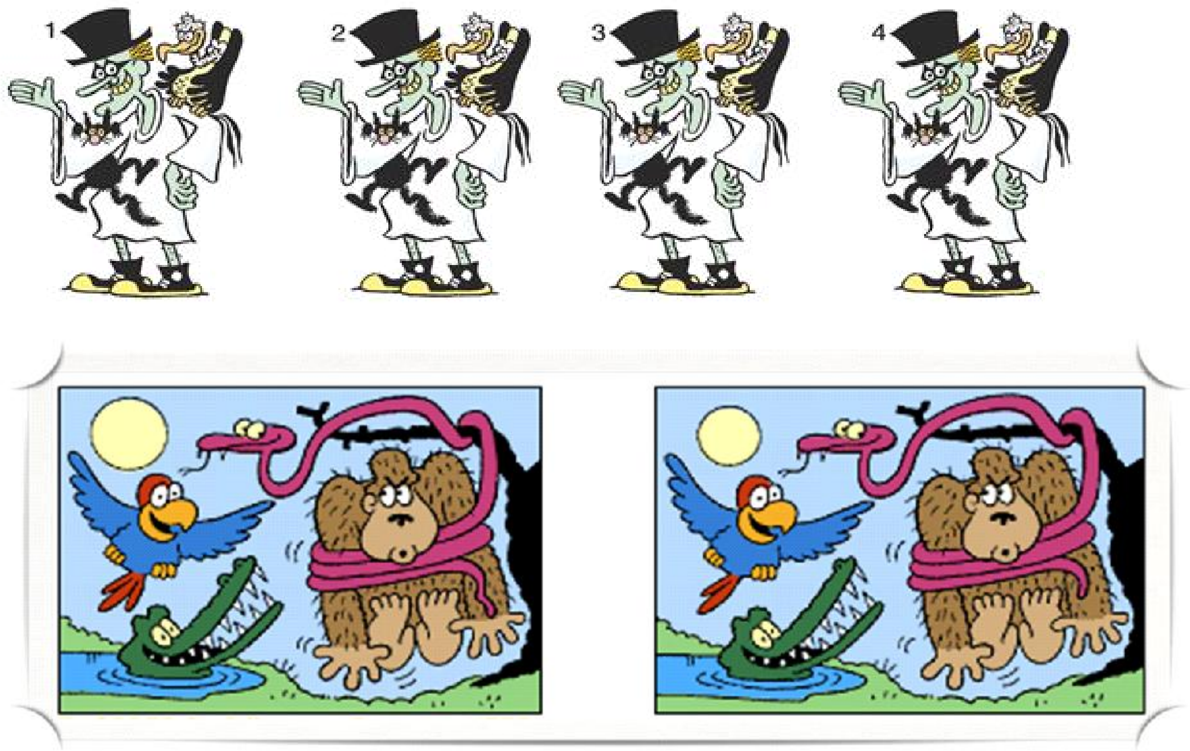


Figure 61 Examples of tests that were specifically chosen and used for fingerprint examiners recruitment tests. Which two of the four pictures (1-4) are identical (top panel), and what are the differences between the two images (bottom panel).

In the worldwide survey of cognitive tests used during recruitment, one test was found that did not fall within the two categories specified above. This was the Form Blindness Test (see Figure 62). This test was developed specifically for the forensic domain, in particular for document examiners. It received some recognition because it appeared in a published paper in the journal of the International Association of Identification (IAI), the Journal of Forensic Identification (JFI). The article presents the Form Blindness Test and promotes its use (Byrd and Bertram, 2003). The original Form Blindness test was developed quite some time ago (Osborn, 1939). The Form Blindness test is one of the most reasonable cognitive tests currently used during recruitment of fingerprint examiners. Although it is a step in the right direction, it falls short to be included in the cognitive tests that need to be developed. The reasons for this are as follows:

It is built to test defects, not for quantifying special abilities, talent, that is required for fingerprint examination. This is not only a conceptual and theoretical difference, but it has practical implications. For example, 'defect' testing, such as visual acuity and this Form Blindness Test, should only test for minimal threshold requirement (which if you do not reach, you should not be a fingerprint examiner, and if you do reach, then you have passed the test). Such 'defect' testing does not provide scores that enable to judge and rank the relative talent of the candidates.

The test was developed over 80 years ago. This is decades before cognitive psychology had even emerged, and very long before brain scans and other cognitive neuroscience tools and methodologies had been developed. In the past 80 years our knowledge and understanding of the human brain and the cognitive system, and in particular visual cognition, has increased substantially. This test does not take into account any of the scientific findings and insights from the last 80 years.

It lacks proper validation even in the area it was designed for implementation and use, i.e., document examination. Even more so it has not been scientifically validated in the domain of fingerprinting.

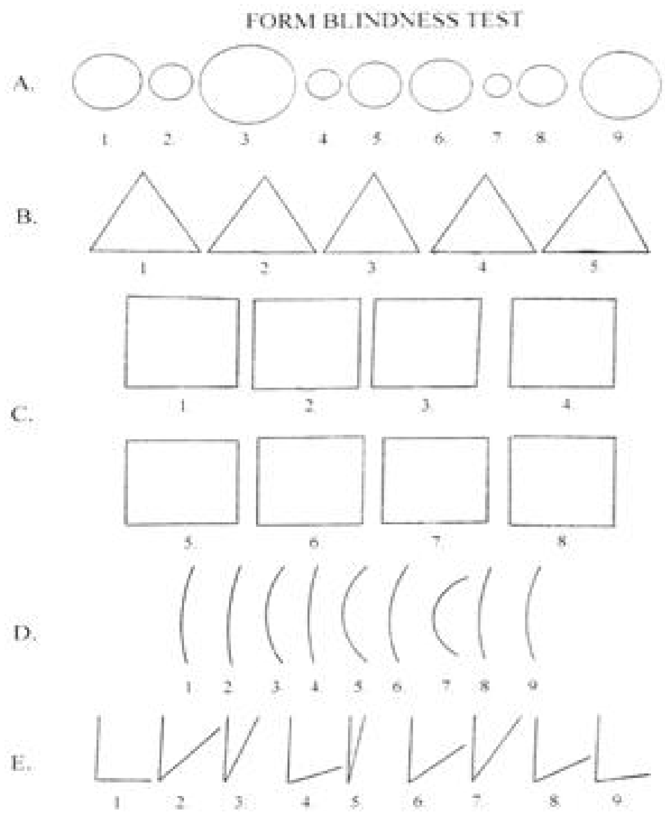


Figure 62 Form Blindness Test, developed in 1939 for document examiners. From A to E which two examples most closely resemble one another).

Qualitative Interviews with Fingerprint Examiners to Establish Cognitive Profiles of Fingerprint Examiners

Experiential knowledge is very important and valuable. The knowledge of fingerprint practitioners (mainly examiners, but also managers and people who train fingerprint examiners) was sought to understand and gain insight into the cognitive profiles of examiners and what it is that makes them good at what they do and what characterises their skills and abilities. Below are a summary of general issues, trends and conclusions. The interviews were semi-structured in design and were constructed in such a way as to encourage latent print examiners to volunteer information pertaining to their daily workload that would allude to the cognitive processes impacting and influencing their day to day activities. Laboratory managers and trainers were also interviewed, so as to gain further insights and points of view.

Method

Each interview was digitally recorded to enable accurate recording of the data produced and to enable focus on conducting the interview, rather than taking full notes 'on-line'. Each recording was destroyed as soon as was practicable to complete the data recording, and no longer than 48 hours post interview (no audible records now exist of these interviews). Each interview lasted approximately 30 minutes and was held in a variety of settings, which provided seclusion, privacy and quiet that promoted free discussion. Each participant was given a coded identifier to anonymize their data. In addition to the consent forms each participant was invited to fill out a short questionnaire which asked participants to provide training and career information as well as hobbies and past times. Before assessing the detail within the interviews, it would be interesting to analyze the detail within the questionnaires, as this too, may yield interesting facts about the cognitive make up of fingerprint examiners.

Within the body of the answers to the questionnaire proper, there was evidence amongst the participants of many hundreds of hours of on the job training as well as continual peer review of standards and procedures. Many of the participants worked within accredited laboratories via ASCLD and personal accreditation via the IAI certification programme. Interestingly, many cited as a hobby some sort of musical interest, playing a specific instrument, while others professed to painting, or art as a general background interest. The assertion that musical ability, or at the very least a creative predisposition, goes hand in hand with latent print examiner proficiency has long been an anecdotal discussion point, but the data collected as part of this chapter does seem to reinforce this conjecture and may be worthy of further research however is out of scope for this study.

In assessing the interview data it was important to concentrate on the headline cognitive abilities. To aid this it was useful to tie in various comments made by examiners to certain question types. This helped to focus an understanding of the examiner thought processes associated with areas of the fingerprint science. The interviews conducted contributed to understanding of the fingerprint examiners skills and abilities as evidenced in previous chapters, but the opinions of the experts and managers interviewed are worth reiteration. In previous chapters the interviews served to highlight examiner perceptions on specifically AFIS usage and pattern choices, here, it was important to document broader cognitive profiles of experienced fingerprint examiners.

Experience, length of service and what is easier now than early on in careers.

Some participants felt that it was harder to differentiate individual Galton details when first starting out in their career, and that it was now possible, with experience, to focus in on detail even in light (faint) prints, establishing whether a feature was a ridge ending or bifurcation. Vague detail was seen as problematic at the start of one's career. Many examiners alluded to seeing more detail today in difficult prints, than they could earlier in their careers, with examiners having more confidence to call matches in difficult prints as their experience grew. With experience, it appears examiners now find it much easier to assess distortion and to determine substrate and matrix material. In contrast, those just starting out in their careers found it problematic to analyze the arrangement of ridge flow and features.

Why everything seems easier to the participant today was described by one participant in the following way, 'the eye is used to seeing ridges, bifurcations and dots'. In other words, it seems that there is a form of experiential 'mind programming' based on trial and error that enables the examiner to learn what a valid level two Galton feature looks like. It should be noted however that this 'mind programming' appears to take place without a 'ground truth' feedback loop. This of course places emphasis on the nature of the feedback received to aid training being based on subjective opinions of peer review, rather than categorical validation of examiner decision making.

Assessing ridge flow and structure is also perceived as easier today than would have been the case early in the careers of many examiners. With experience the examiner appears to gain confidence in decision making ability. By the time they are considered an expert and able to testify in court, they appear to be able to consolidate those decision making skills and to be more assertive in their opinions. Of course, whether this is a cognitive trait based on learning and experience, or whether this has more to do with the social psychological nature of group dynamics and peer effects is a matter for future investigations. However, even as a cognitive trait, it seems to affect decision thresholds, rather than modify cognitive abilities.

Examiners appear to find it easier with experience to locate focal areas, or targets. As one examiner put it, 'it just pops out now'. It was suggested that inexperienced examiners looked at the entire print holistically and that it took time, but that with experience they are able to 'just concentrate on areas of interest'. Indeed those early in their careers also found locating areas of friction ridge skin difficult, especially palm, whereas now they able to narrow down the search more effectively. This insight appears to suggest that with expertise comes a process of taking cognitive short cuts, whereby the expert is taking cognitive efficiency processes that enable the examiner to locate and memorize certain key areas within a fingerprint to aid the comparison process.

Examiners discussed the ability with experience to differentiate light and dark in a latent finger mark and to be able to filter out distortion and noise. In fact some examiners were most insistent that the ability to have gray scale acuity was vital for the better examiners to distinguish contrast within a print and to differentiate between genuine features and background noise from either the matrix used or the surface substrate from which the latent print recovered.

Participants felt that today there is less emphasis on understanding how to search for fingerprints because AFIS (Automated Fingerprint Identification Systems) removes some of that need. Newer examiners, it appears, just don't see as many fingerprints as the pre AFIS generation examiners. As one examiner put it, AFIS 'takes away the ability to make the decision…you are already given the score…people get used to the ident being at the top of the queue'. 'Examiners don't have the ability on AFIS to decide the search limits and are willing to accept the decision of AFIS, be it a non match and just stick it on the shelf'. 'There is very little motivation now to try harder….we have lost ownership of casework'. 'If you are just doing AFIS work you will not build up the sort of experience you need to make decisions'. This relates to issues of distribution cognition, will require further investigation in the future and is beyond the scope of this particular study (see previous chapter).

What is the most difficult and challenging aspect in latent fingerprint examination?

Most examiners, it seems, even with experience, still find interpretation of distortion challenging. The types of distortion described included low quality and feint impressions as well as a type of distortion caused by movement that result in a thickening or thinning of ridge detail. In addition, experience is still no protection against decision making where there are poor quality prints. It seems those hard to call comparisons still require the examiner to somehow reach the tipping point, or to just 'walk away'. Many examiners eluded to the gray area between identification and a non match, 'where you would want to see one more ending ridge to be sure', 'in the gray area it's tough', said one examiner.

Challenges presented to examiners by really bad prints, where there is simply a lack of information, but where the examiner has a 'gut feeling' that there is a match to be made, are still problematic in that it appears even experienced examiners are unsure or uncertain as to where their individual 'tipping point' or 'decision threshold' is, or even where it should be. The challenges are greatest apparently when you 'need to find a match' and maybe when it is appropriate 'to give up'. It was felt that 'marginal prints' provided the most problems when analyzing latent finger marks as this was when distortion was most problematic in terms of interpretation and determining feature confidence. Distortion still provides the most challenges to many examiners, especially mirror image prints (reverse direction) and reversals (reverse color where the ridge flow goes from a positive representation where the examiners sees black ridges to that of a negative where the ridges are in fact the furrows and vice versa).

How long does a comparison on average take to complete?

The general consensus from the interviews indicate that the easier comparisons as defined by the examiners as 'roadmaps' where the latent looks almost like a rolled impression against a good quality ten print card, will take a matter of a few seconds to a few minutes. In contrast, the really difficult comparisons can take hours, or even days, requiring high levels of dedication and concentration to reach conclusion. While the comments of examiners must be taken at face value, it is unclear what mechanisms are impacting on the decision making process when the really difficult comparisons do indeed take days. For example, does motivation and the need to solve high profile crime place a contextual influence on the time spent to see features in a fingerprint comparison? If so, then the length of time spent on a comparison may or may not aid the attainment of both accurate and credible outcomes. This will need further research outside the scope of this research.

What makes some comparisons more difficult than others?

Some participants felt that rolled prints have different types of distortion to that of plain impressions. With poor and 'ugly prints' participants suggested that sooner or later the decision has to be made (a feeling that the decision must either be yes or no). The difference between different capture processes can also make a difference to the ease or difficulty of the comparison. Ink, live scan, is the person a bricklayer, smoker etc., will all provide different challenges to the examiner. For example, older people tend to display flattened out and less well defined ridge detail, and bricklayers have worn and scratched ridge detail that creates a lack of quality and clarity in the observed ridge detail.

Participants mentioned 'noise in the artifacts' will impact on ability to interpret information. Digitally captured latent images were seen as hard to compare to a ten print card when pixelization of features causes difficulty in interpreting detail. It was also suggested that poor quality ridge structure makes it difficult to assess what are real or false characteristics. Other problem examinations involve background noise, and general distortion caused by the substrate surface and the media used to recover the latent material….i.e. powders and different chemical treatments etc. For example, finger marks developed using a chemical process such as ninhydrin, will often result in broken up ridge detail that requires the ability of the examiner to effectively 'fill in the gaps' left by the missing information.

Latent (print) Orientation

The correct orientation of latent finger marks was described as a bit like looking at a map of the earth… 'it can be cut up like a jigsaw but as soon as you see a continent you know where it fits on the globe', 'you know where it belongs'. 'No matter how you cut it, how you stretch it, it will always be the same globe'. Orientation was defined by observation of ridge flow and observation of deltas and other ridge detail, and was based on experience and training, a judgment based on the experience of understanding ridge flow and where features regularly appear. Orientation is usually obvious, and there is usually a 'most likely way it is oriented'.

Apparently after seeing thousands of prints you get to 'know where the front of the car is'. Print orientation is based on experience but also a mix of 'flipping around'… 'a bit like looking at jigsaw puzzles'. 'You have a mind that thinks in bits and pieces…do things flow right'. When asked about orientating a latent one participant alluded to treating the ridge flow like 'a puzzle' to be solved, and often the use of visual mental rotation of ridge flow and target features helps to define orientation, based on experiential trial and error. Correct orientation of the latent was described in terms of 'mental gymnastics' whereby the features and ridge flow were somehow ordered mentally. In short, participants discussed skills for orientation of latent finger marks including visual acuity, the ability to hold the detail in the mind and to draw from the mind's database to match up likely ridge flow patterns, etc.

Deciding where to concentrate your analysis of the (latent) finger mark

Target searching appears to involve looking for groups or clusters of unique similarities. Participants discussed looking for targets in the same space and relationship, not counting points, but rather working holistically. Examiners, it seems, do not always assess individual features in a systematic way, but rather they look at any features that stand out. This could be ridge flow, or it might be lakes or sweat pores. Some participants discussed looking for target detail in the core area if at all possible. Level two details such as ridge endings or bifurcations will usually stand out before other details when initially analyzing a print, though smaller more unique features like lakes and crossovers will always draw the attention of the examiner and provide an obvious target point from which to begin the comparison process.

A good comparison involves not just looking at similarities but also assessing areas of dissimilarity. Examiners described a preference to work sequentially round a print, working around a print looking for information. Though these observations seem at odds with earlier observations whereby examiners suggest they are drawn by features that stand out. Further research will hopefully clarify this apparent disparity and more clearly define how examiners conduct visual navigation during assessment of friction ridge skin. Examiners alluded to seeing fingerprints as ‘maps’. There will be lakes, areas where there is nothing to see, furrows and roads to be followed. The quality of the map helps to define the direction in which to go.

Participants generally felt that arch patterns present problems to examiners because the opportunity to focus on target groups of features is more difficult. Indeed participants, it seems, spend more time on arches, asserting that in local proficiency tests that arches take longer to assess than others, because there are no obvious target features. It was also suggested that subtle details associated with third level detail is assimilated almost unconsciously. Sometimes these features pop out at you. There is also evidence that participants relate ridge detail to associated known objects, like hooks and dolphins, etc. Something the examiner has seen before to create the association. The desire to associate ridge detail with known objects in the real world may be an important clue in ascertaining how better examiners are able to memorize and match ridge detail.

Tolerances of feature confidence are based on what information is in print and how it is related to the features under consideration. Participants discussed the notion of having a database of images in the head from which to draw templates of expected feature configuration which aided search strategies. A process was used to develop knowledge of the latent which was described by one interviewee as ‘mind programming’.

Participants discussed the memorization of target features, then holistically applying those features to the known print, gradually narrowing down the comparison till achieving a match. Examiners needed sufficient volume and uniqueness of ridge detail, while all the time accounting for distortion, assessing relative confidence in features based on the immediate area of interest. Interestingly, it appears the examiners do not consider pattern types, but rather consider the latent print ‘holistically’. One participant suggested that examiners do not categorize pattern type mentally, but rather prefer to draw sketches to aid the determination process of selecting target features. This individual considered that good examiners will be adept at drawing and retaining vital information through such a process. This assertion may be pertinent when it has been noted already that there appears to be strong evidence of creative and artistic predisposition whereby selection of target features by drawing and retaining information may be aided by such abilities.

Examiners interviewed assessed the difference between good examiners and less good examiners in a multitude of ways. Abilities discussed included the speed of comparison and speed of decision making, given that it was assumed that both types of examiner will be accurate. It was also suggested that a good examiner ‘will be able to interpret limited information’, and ‘display a balance between risk taking and decision making’.

Participants felt that the best examiners can observe many things at one time, for example, one participant knew they could do this job (fingerprint analysis) because they had taken tests in the past that included looking at lists of information and assessing the accuracy of those lists.

It was felt some examiners seem to see things others cannot. They see things differently. Some examiners, it seems, have more drive and desire to hunt out and search and a higher tolerance for thoroughness and an ability to complete work. There was some sense that where computer systems generate respondents on AFIS that some better examiners will go right to the comparison on screen, while others want to see other latent finger marks in the case in hard copy format before making a decision.

Task analysis, cognitive processes, and their measurement

The interviews, along with the literature review, provide insights to the cognitive processes that underpin work carried out by latent fingerprint examiners. How can this knowledge be fully laid out as foundations for hypothesizing and testing these ideas? To achieve that it was necessary to carry out two additional steps:

Conduct task analysis.

Integrate the task analysis with the interviews and literature review, and consider all the combined data together, so as to draw conclusions about the cognitive processes that will establish the cognitive profiles. These will be empirically tested.

To conduct the task analysis each and every step latent print examiners take in their work was considered. Each step from a cognitive perspective was analysed, thus examining what cognitive processes are required to accomplish each step successfully. Two steps that seem very similar from a non-cognitive perspective may in fact rely on very distinct and different cognitive processes; conversely, two steps that seem very different from a non-cognitive perspective may in fact rely on a single cognitive process.

A whole set of cognitive neuroscientific research establishes the cognitive processes involved in various tasks (Dror and Thomas, 2005; Gazzaniga, 2005; Posner, 2004). Scientific research in this area relies on three different lines of research, all combined together. The different lines of research are neuroscience, computaional investigation and behaviour.

Neuroscientific research that looks at brain activations while people perform different tasks (e.g., Sergent, Zuck, Lévesque, and MacDonald, 1992; Corbetta, Miezen, Schulman, and Petersen, 1993; Kosslyn, Alpert, Thompson, Chabris, Rauch, and Anderson, 1994; Sergent, Ohta, and MacDonald, 1992).

In this line of evidence, for example, functional Magnetic Resonance Imaging (fMRI) or Positron Emission Tomography (PET) monitor brain activity and examine which parts of the brain are activated; specifically looking if different brain regions, or one specific area, processes information relating to various tasks. Figure 63 shows a brain scan revealing how different brain areas are active (top panel), and how using this method shows that different cognitive processes are involved in making categorical spatial judgment and in metric spatial judgment (Hama, Raemaekersb, Wezelc, Oleksiakc, and Postma, 2009; Kosslyn, Thompson, Gitelman, and Alpert, 1998).

Computational investigations (Marr, 1982) that utilize computer simulations and modeling, while using computational architectures similar to the brain, such as parallel distributed processing, connectionism, and neural networks (e.g., Bower, 1990; Crick, 1989; Cook, Früh, and Landis, 1995; Gluck and Thompson, 1987; Grossberg, 1987; Kosslyn, Chabris, Marsolek, and Koenig, 1992, 1995; Jager and Postma, 2003; Kosslyn, 1987).

In this line of evidence, for example, computer neural network models computationally investigated the cognitive processing underlying object features (such as shape, color, and texture) and spatial features (such as location and motion). The computer simulations tested the relative computational similarities between these different cognitive abilities, so as to understand to what extent they share similar cognitive processing (Rueckl, Cave, and Kosslyn, 1989; Jacobs, Jordan, and Barto, 1991).

A variety of behavioural experimentations include requiring participants to multi task and measure cross-task interference (e.g., using the 'dual task paradigm'), tracking the eyes of the participants while they perform a variety of tasks, and observing the behavioural deficits of brain damaged patients (e.g., Laeng, 1994; Meadmore, Dror, Bucks, and Liversedge, in press; Mesulam, 1981; Robertson and Delis, 1986; Liversedge and Findlay, 2000; Rayner, 1998; Rullen, Reddy, and Koch, 2004).

In this line of evidence, for example, it was determined the working memory (short term memory) has distinct cognitive processing for visual information and for auditory information. The former is processes by the Visuospatial Sketch Pad, and the later by the Phonological Loop (Baddeley and Hitch, 1974; Baddeley, 1992).

All these different lines of research come together in cognitive neuroscience to provide a clear understanding of how different cognitive processes underpin human performance (Dror and Thomas, 2005).

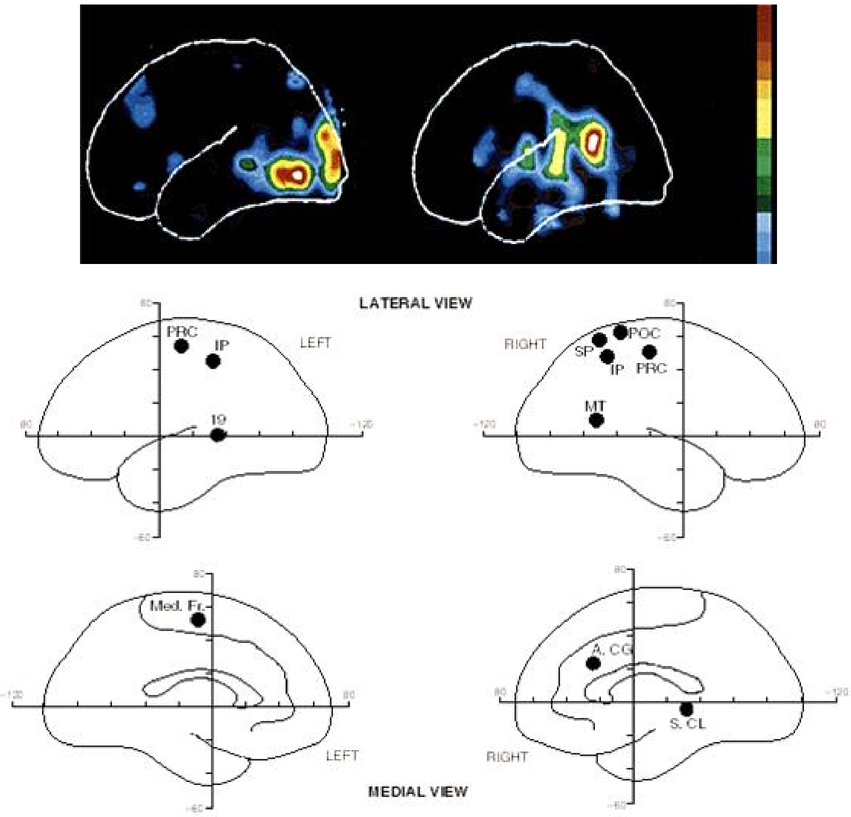


Figure 63 A PET brain scan (top panel) different brain regions are activated when using ability to make categorical vs. metric spatial judgments (bottom panel; Kosslyn, Thompson, Gitelman, and Alpert, 1998).

The task analysis was based on interviews, literature review, and expertise in forensic fingerprinting. The next step was to cognitively analyse each step that examiners take in fingerprinting. Although most latent fingerprint examiners follow the ACE–V steps (Analysis, Comparison, Evaluation, and -Verification), there are variations of how these are carried out (Dror, 2009).

Therefore, with such differences it was hard to assess across the board the various steps latent print examiners perform in their work. The NIST Human Factors Expert Working Group on Friction Ridge Analysis made the task easier, as they have devised a flowchart that captures and specifies the various steps most examiners use in their work (see Figure 64).

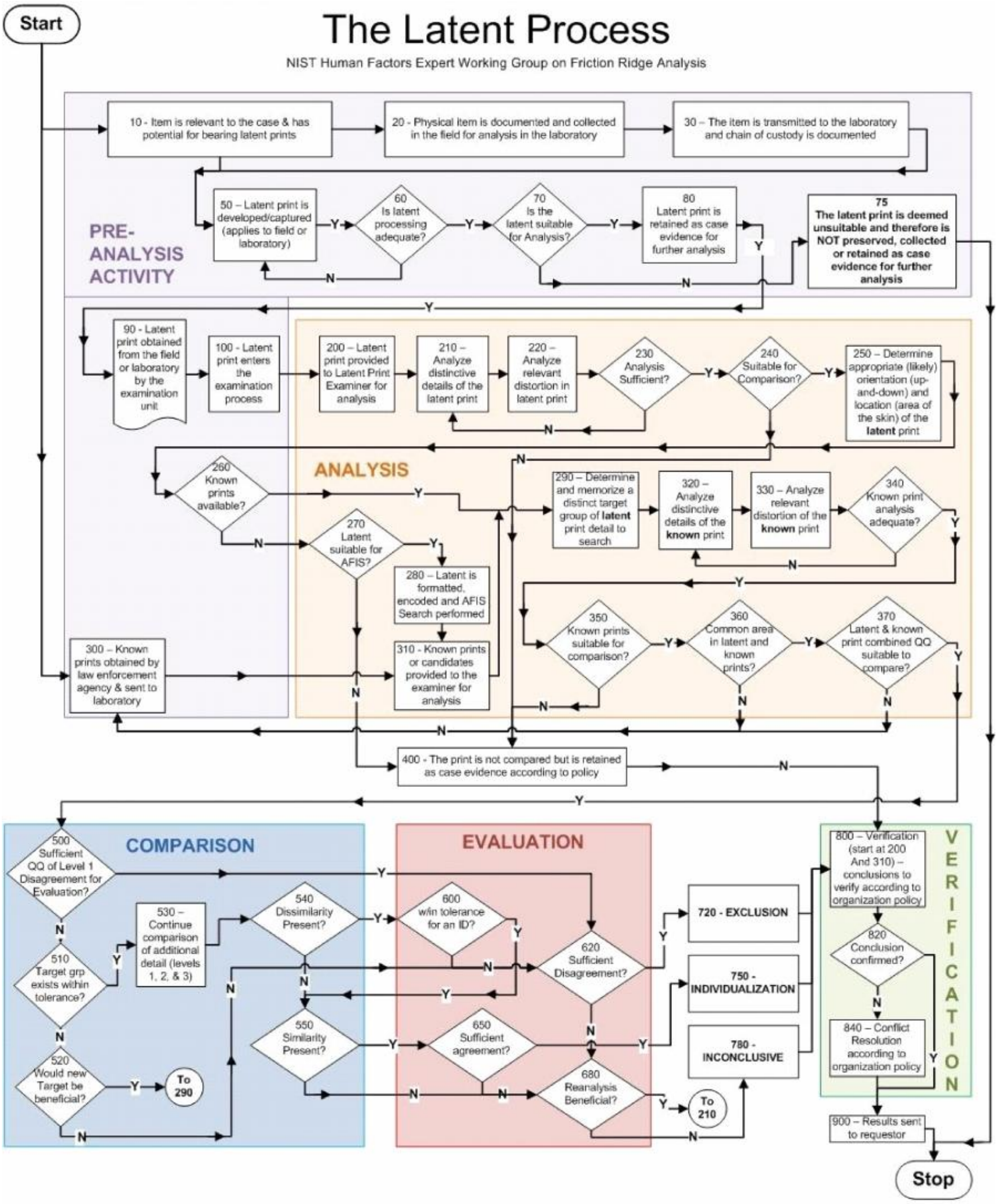


Figure 64 The various steps most latent print examiners use in their work.

The results of the task analysis revealed that latent print examiners rely predominantly on the following cognitive processes:

Attention allocation
Visual mental imagery
Inspection
Rotation
Transformation
Dealing with and filtering noise
Perceiving and comparing
Curvatures
Width
Features
Visual search

Attention Allocation

One of the cornerstones of the human cognitive system is that it has limited computational capacity and resources, and therefore, it cannot process all the information that is provided as sensory input to the brain (Knudsen, 2007; Posner, Snyder, and Davidson, 1980; Sperling, 1960). The result is that the cognitive system selectively allocates attention and cognitive resources. This process requires us to constantly engage and disengage attention and shifting attentional focus and cognitive processing to different segments of the visual input (Wright and Ward, 2008; Posner and Petersen, 1990).

The ability to allocate attention to the important and most crucial information defines much of the human cognitive system and intelligence. This process of selectively and wisely knowing where to focus attention gets better and more refined with expertise (Dror, 2012). Selective attention is one of the cognitive ways of achieving expert level performance (e.g., Wood, 1999). As one becomes a greater expert, they get more selective, paying only attention to the important pieces of information, at an ever increased rate and accuracy.

While a novice is still trying to absorb the information and make sense of it, the expert has already focused on the critical information (e.g., Valk and Eijkman, 1984), processed it, and solved the problem. The process of selection is critical. For example, expert radiologists selectively process X-ray films according to clinically relevant abnormalities (Myles-Worsley, Johnston and Simons, 1988; Valk and Eijkman, 1984). This results in efficient and effective processing.

In latent print work, experts must allocate attention to the pieces of information that will help them reach a conclusion; points of similarity as well as discrepancies. Figure 65 illustrates such a potential test.

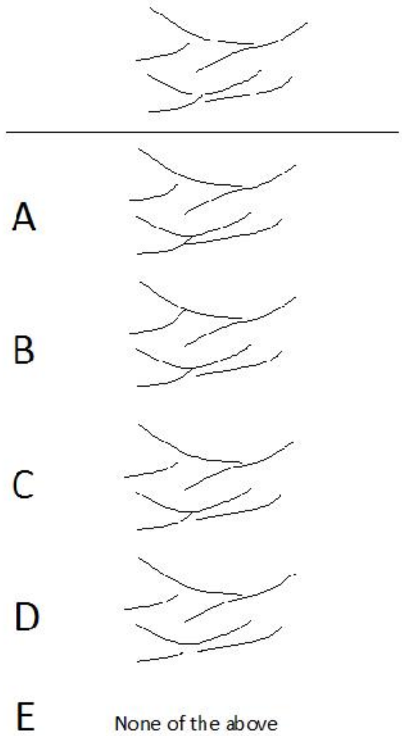


Figure 65 Potential Test. From A to E, which is the closest match to the exemplar provided?

Visual mental imagery

Visualizing images in the 'mind's eye' is a common cognitive process; in fact most people use mental imagery to help them perform everyday tasks. For example, when trying to recall the colour of a car, one may create a 'mental picture' of the car and 'look' at it; or when shopping for a new sofa, one might visualize how one's room would look if the sofa were placed against a particular wall. Imagery can be used not only in memory and reasoning, but also in other cognitive tasks. One of the most fundamental discoveries about imagery is that it has a complex underlying structure (Dror and Kosslyn 1994).

Mental imagery, however, is not a single cognitive process, and is comprised from a whole set of mental abilities. Some of these are especially important and characteristic of tasks that are important for experts in certain domains, such as Air Force pilots (Dror, Kosslyn, and Waag, 1993), musicians (Aleman, Nieuwensteina, Böckerc, and Haana, 2000), and designers (Kavaklia and Gerob, 2002).

For latent fingerprint experts, there are three visual mental imagery functions that are especially important:

Image Inspection: This process enables to visualize and inspect an image (for details, see Cooper and Podgorny, 1976; Dror and Kosslyn, 1994; Dror, Kosslyn, and Waag, 1993; Busey and Vanderkolk, 2005). Often latent print examiners compare friction ridges by visually examining one and comparing it to another they are visualizing.

Image Rotation: This process enables to imagine an image in different orientations (for details see, Shepard and Metzler, 1971; Dror, Ivey, and Rogus, 1997). Visual mental rotation is often used by latent fingerprint examiners when they have a small latent mark and they are trying to find the corresponding area on a 10-print.

Image Transformation: This process enables to mentally reconstruct and modify visual images (for details see Shepard and Cooper, 1982; Shepard and Feng, 1972). Distortions often affect friction ridge, and latent print examiners need to imagine the print without the distortions, thus mentally modifying and recreating a 'distortion free' image in their mind.

Figure 66 presents some tests that may be suitable for testing these abilities with latent print examiners.

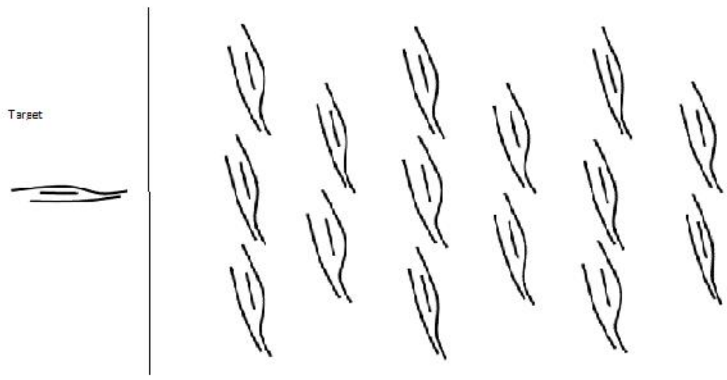


Figure 66 Ability Tests, which shape to the right most closely matches the exemplar on the left?

Dealing with and filtering noise

Visual information provided to the cognitive system is rarely ‘noise free’. Along with the ‘signals’ which contain the sought out information needed to perform the task, the input also contains ‘noise’. For the cognitive system to make sense of the signals (e.g., recognize patterns, evaluate and judge the significance of the information, and interpret the meaning of it), it must first deal with and filter out the noise (Busey and Vanderkolk, 2005; Lu, Lesmes, and Dosher, 2002; Harmon and Julesz, 1973; Solomon and Pelli, 1994; Dosher, and Lu, 1998). Experts are very good at ‘seeing through the noise’ (Busey and Dror, 2011).

The work of latent print examiners often involves working with images that contain noise. Noise can be as a result of the substrate from which the latent finger marks are retrieved, for example most kitchen work tops are stippled and grainy and desk tops often have a wood grain patterning even if a laminate is used rather than real hard wood. Noise can also be as a result of the matrix used to enhance the latent print (see figure 67 and 68)

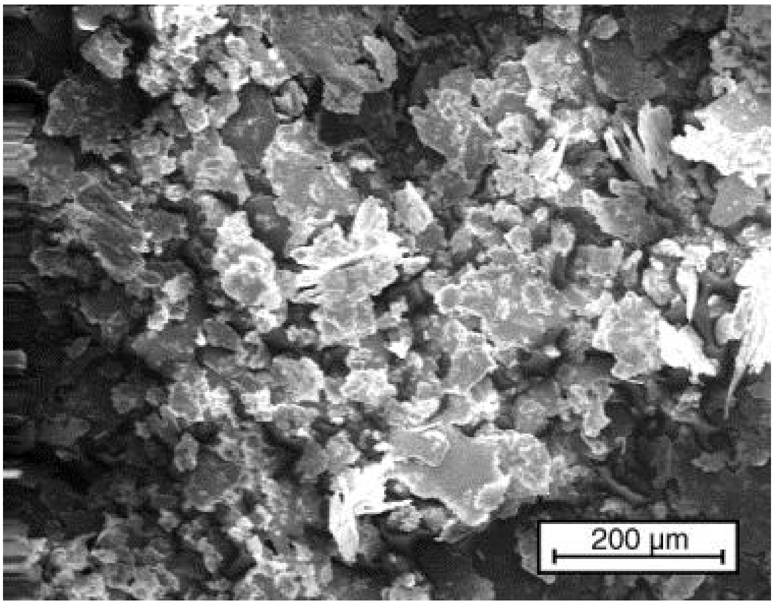


Figure 67 Flake-shaped aluminum metal powder particles (magnified 100 times)

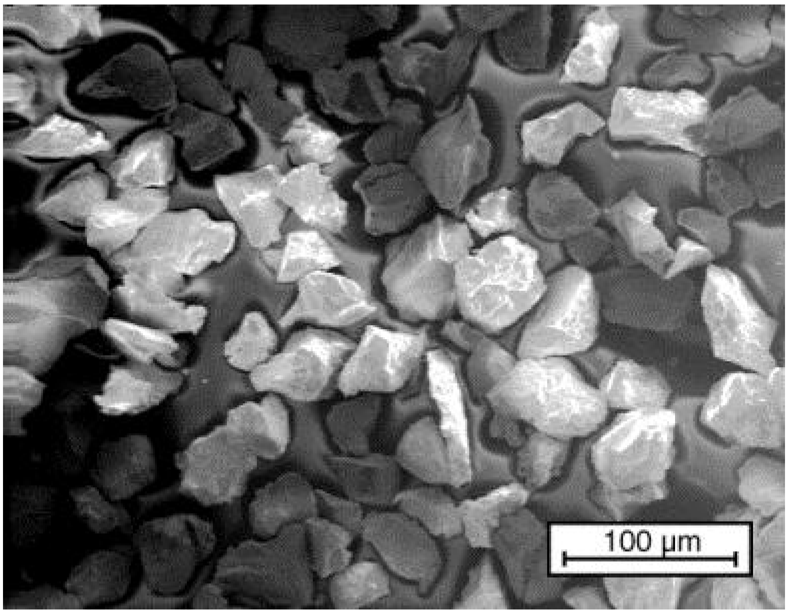


Figure 68 Granular-shaped ferro-aluminum metal powder particles (magnified 100 times)

Aluminium powder, magnetic powders and other powder compounds have different granular shapes and properties that will impact upon their ability to adhere to surfaces and record finger marks. The use of these powders will create levels of distortion just through their application. Latent prints are rarely ‘noise free’. Latent print examiners ability to deal with and filter such noise seems to play an important part in their work. Figure 69 represents a potential test for this with latent print examiners.

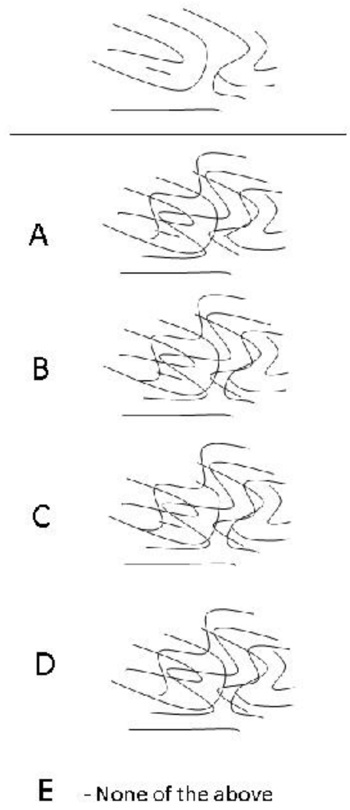


Figure 69 Potential Test for Latent Print Examiners, from A to E, which contains the configuration of the exemplar provided?

Perceiving and Comparing

One of the fundamental abilities of the human visual system is to perceive visual information, encode its characteristics, and compare them against other images. These visual characteristics are many, and depend on what is unique and informative for the task at hand (Dror, Stevenage, and Ashworth, 2008; Kundel and Nodine, 1983; Ashworth and Dror, 2000; Biederman, 1987). Some of the most relevant visual information used by latent print examiners are:

Curvatures: The ability to perceive, judge, and differentiate between different curvatures.
Width: The ability to perceive, judge, and differentiate between different widths.
Features: The ability to perceive, judge, and compare different minutia and other features that appear within friction ridge.

Figure 70 represents a test that may be used for testing these abilities.

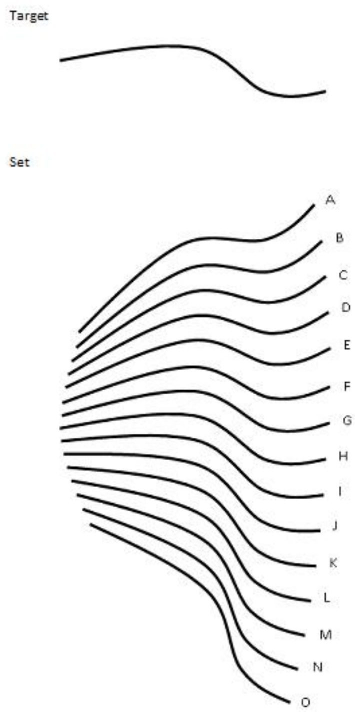


Figure 70 Test for testing perception abilities, from A to O, which most closely resembles the exemplar provided?

Visual Search

Before perceiving and comparing curvatures, width, and features; dealing with noise; and allocating attention (see above), one must first scan the visual input and search it, so as to detect regions of interest within the image (Wolfe, 1994; Czerwinski, Lightfoot, and Shiffrin, 1992; Nobre, Sebestyen, Gitelman, Frith, and Mesulam, 2002). This is an important ability for latent print examiners, as they need to visually search the friction ridge to find which regions have the information most valuable for them, i.e., containing information that they can use to reach a conclusion. Figure 71 presents a test that may be applicable to test this ability with latent print examiners.

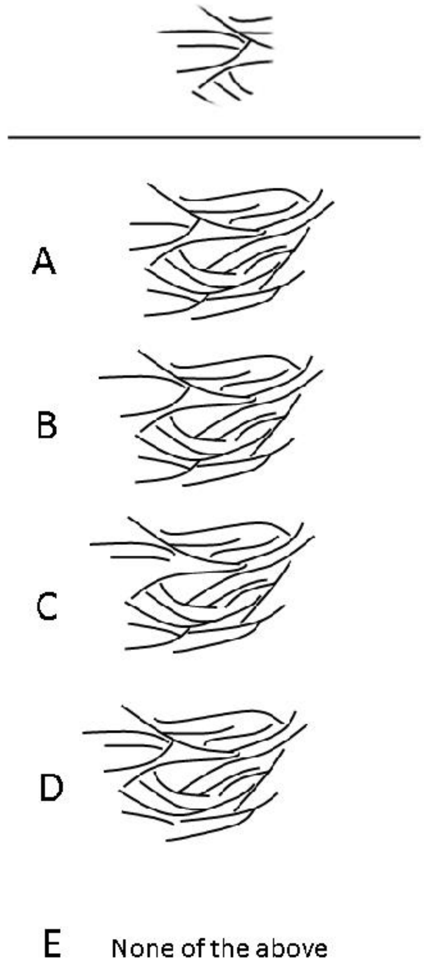


Figure 71 Potential Test for Latent Print Examiners, from A to E, which contains the configuration seen in the exemplar provided?

Further Steps

These cognitive processes constitute some of the candidates for the cognitive profile of latent print examiners. They will be used in the future to generate tasks and tests, which need to be empirically tested. The test will first need to be piloted to calibrate difficulty so as to avoid floor and ceiling effects. Then, they need to test if they can isolate and quantify the cognitive process, and whether the cognitive processes in question correctly characterize latent fingerprint examiners. After completing all these tests, and selecting the proper test based on statistical analysis, they must also then be cross validated via a new set of blind data collection. This will enable the development of scientifically based testing for recruitment. Doing such research and development requires time, effort and funding, however it is very cost efficient. Once such a test is properly developed, saving and efficiency will grow in the field, as well as a better understanding of the cognitive abilities needed for latent print examination, which will aid other efforts in the domain.

Summary and Conclusions

An industrial revolution is needed in fingerprint science so standards and management of forensic science is efficient, streamlined and conducted effectively. This process should be based on data and scientific understanding and research. This is especially important in the area of human cognition and decision making because it plays such a critical role in many forensic domains, and has been largely neglected in forensic sciences. Such research can lay the foundations to scientifically develop proper proficiency testing, training, testing during recruitment, and best practices and protocols. In this chapter it has been described how such a scientific endeavour should take place. Testing for screening and selection during recruitment was used to exemplify and illustrate in a very practical way how this should be done. It is hoped this will serve as a blueprint for other developments in forensic science.

Chapter 9:

Group Networks and Dynamics Associated with Fingerprint Analysis and Decision-Making (Discussion)

Thus far in this thesis it has been predominantly the intention to assess the cognitive functions of the human mind and how these impact upon expert performance in the fingerprint domain. While this is the primary focus of this thesis, it might also at this juncture be worth discussing briefly some of the other social psychological influences that may impact upon not just decision making, but also perhaps verification and arbitration tasks as well.

The domain of forensic science is arguably one of the most highly scrutinised and debated areas of law enforcement. It can be a source of reliable and rapid intelligence, or, at a stroke, can destroy a police prosecution because of bad practices or poor interpretation of the results obtained. In recent years latent print examination has been criticised by both lawyers and academics for its perceived lack of scientific reliability (Cole 2001) and a failure to accept the concept of human error rate. There have also been public concerns raised over instances of high profile erroneous identifications by forensic practitioners, including latent print examiners. Notable examples of where the science of fingerprints has been undermined by such errors of judgment are now evident as a global phenomenon. The forensic community should not assume that such problems exist solely because of inappropriate execution of procedures or the technical shortcomings of individual practitioners. Here are just a few examples of where poor execution of procedure or poor judgment have resulted in catastrophic misinterpretation of data, leading to wrongful arrests and damaged reputations.

The Chiori case of erroneous identification at the Metropolitan Police had at its root cause a culture of hierarchical peer pressure to conform. Rank in itself can be destructive to the individual integrity of the verification process (Charlton 2002). Senior officers can often 'intimidate' staff, the assumption often being made is that a senior fingerprint officer must because of their higher rank necessarily be a better perceiver of detail and as a result a 'better' examiner. As stated by Charlton in 2002, 'all examiners are equal, but some are more equal than others' (adapting a quote from Animal Farm by George Orwell). Indeed, often debate is controlled and even stifled by senior managers in the domain to 'protect the profession' (Charlton 2001).

The Shirley McKie and Brandon Mayfield cases demonstrated how the decision-making abilities of experienced examiners could be compromised. In the cases discussed, as well as the cognitive biases and vulnerabilities mentioned thus far, other external environmental conditions either within the bureau, or external to the workplace, may possibly have conspired to produce the right conditions for such mistakes to occur. In any work place there is always the possibility of such conditions prevailing.

Hierarchical dominance can lead to the need of a junior member of staff to agree with decisions or actions of others because it either is too difficult to disagree through fear of ridicule or a sense of inadequacy, or because they may lack experience to counter a superior's conclusions. While procedures and processes within fingerprint bureau are designed to offset these human traits, it is by no means certain that this has been universally achieved at the present time. Even the way casework is presented and the atmosphere and context in which decisions are made can be influential in determining the potential validity of an individualisation.

Research within this thesis has provided demonstrable proof that the context in which an examiner makes judgments can influence the conclusions of even the most experienced practitioners. The context in which evidence is assessed has been shown to induce contrary decisions by examiners on specific mark to ten-print comparisons, even where the examiner concerned has already assessed the same latent finger mark evidence previously (Dror and Charlton 2005 and 2006). It is inevitable that there are many factors that can potentially contribute to decision-making errors. Another factor may be that of human interaction and group dynamics in the workplace.

Within the ACE-V process, verification represents the validation of the findings of scientific peers (Ashbaugh 1991). This phase may be vulnerable because of individual decision making variables such as competence and training, but it may also be vulnerable from wider human influence, such as context and other cognitive vulnerabilities, including group and social interaction.

A simple way to explain this concept is to imagine getting a group of work colleagues to stand within a circle. With a large ball of string, ask a question of a colleague, it could be 'count backward from 10 to 0'. Holding on to the string, throw the ball of twine to the person to whom you asked the question. This person will then ask a question of another colleague and throw the ball of twine to this person and so on and so forth. Now, what does this silly game demonstrate to us as academics and practitioners? Well, one would hope and expect that in a utopian social design, where all the group members have equal status, respect, similar personalities and mutual affection for one another, that the pattern observed of the string as it navigates around the circle will be similar to a spider's web that can be observed in any garden. That is to say the web will be beautifully symmetrical and elegant, where the decision as to who might receive the ball of string next was decided on in a non-specific and democratic process. However, what may well happen, is that the observed web resulting from this game looks distorted and will have concentrations of web (string), around certain areas of the circle. This arises because of in-built prejudices based on friendships, alliances, trust and respect. It is the latter pattern that may exist in most fingerprint bureau.

Many reading this chapter may recognise this phenomenon within their own sphere of influence. Specifically in fingerprint bureau there will be those fingerprint examiners that 'have a good eye', or are deemed 'less than effective', or are considered 'arrogant' or 'rude' or 'unsociable'. Any such attribute will potentially impact upon the dynamics of the workplace, the dynamics of verification. If the 'web' of interaction is skewed, the potential for unreliability in the verification process is likewise skewed. This, after all, is why some erroneous identification may arise. This phenomenon can best be described by alluding to a concept known as 'the buddy system'. This is where examiners will have 'preferred' colleagues to verify their bench work. This is at the core of what may be termed a poorly structured 'web of professional reliability'. Unless there is a near perfect social web within the workplace then it could be verification processes are skewed by 'having favourites'?

It could also be argued that non-random selection of those examiners with 'a good eye' for the purposes of verification might generate a more trustworthy and reliable product? While there is some merit in this argument, the foundation of the dynamics of bureau demographics must lie in a generally random and generic validation system that avoids 'picking and choosing' who might check another's work. It is surely better to aspire to generally raise the technical competence of all the examiners through sound training and well designed procedures, rather than rely on a few who are perceived, at least, as being better than everyone else.

So what might we consider as an ideal working unit, and how should it be established when a group of examiners are working well together to provide trustworthy verifiable results?

Kurt Lewin was one of the founders of the movement to study groups scientifically. He coined the term 'group dynamics' to describe the way groups and individuals act in a given set of circumstances (Lewin 1947). Lewin suggested that neither nature (inborn tendencies) nor nurture (how experiences in life shape individuals) alone can account for individuals' behavior and personalities. Lewin suggested that both nature and nurture interact to form an individual persona. This idea was presented in the form of Lewin's Equation for behavior $B=f\ (P, E)$. While the formula should not be seen as a mathematical absolute, the formula offers an elegant explanation for human interaction. It states that Behaviour (B) is a function (f) of the Person (P) and his or her Environment (E). In other words, when considering a laboratory environment, one cannot divorce the individual from the group when examining the likely effectiveness of decision making of one examiner over another. It is important to note that it is both the external environmental conditions such as social interaction and other life experiences, as well as more personal tendencies such as lifestyle and beliefs that shape who we are and how we perform. Such factors provide the fertile ground from which variable performance grows. In considering how to minimise the effects of both external and more personal influences in decision making as described by Lewin, we need to look at models that provide for a more robust professional environment, where such influences, while ever present, can be transparent and mitigated.

In the mid 1960s Bruce Tuckman proposed a methodology for assessing and nurturing group dynamics to maximise team effectiveness (Tuckman 1965). Called the 'Forming, Storming, Norming, Performing' model of team development, Tuckman maintained that these phases are all necessary and inevitable in order for a team of professionals to grow, to face up to challenges, to tackle problems, to find solutions, to plan work, and to deliver results. When developing an effective team (and this is just as applicable to any bureau or lab that engages in fingerprint examination), it is necessary to 'form' the team. The team looks at some of the problems, assesses the challenges and agrees on likely goals that need to be achieved. Such a challenge could be to deliver an effective verification process for evidence and a robust arbitration procedure. Tuckman observed that at first the team members behave independently of one another, and remain relatively ignorant of the issues and problems to be addressed and are self interested and exhibit selfish self centered tendencies. During the 'storming' phase there is an element of conflict as ideas compete for attention and team members confront one another over conflicting ideas and philosophies. The storming stage is necessary to the development of the team. It can be unpleasant and even painful to members of the team who are averse to conflict. Tolerance of each team member and their differences needs to be accommodated and highlighted. Without tolerance and patience the team will inevitably collapse. Tuckman suggested that this phase is a dangerous phase in effective team building and policy formulation and can become destructive, lowering motivation if allowed to get out of control. Supervisors of the team at this time may need to be prescriptive in their guidance with regard to decision-making and professional behavior.

Hopefully, the team will enter the norming stage in which staff adjusts their behaviour to each other as they develop work habits that make teamwork seem more natural and fluid. Generic rules and standards are agreed upon, along with professional values, professional behavior, shared methods, where even taboos (erroneous identifications perhaps) are discussed. During this phase, team members begin to develop trust in each other. During this phase caution should be exhibited as teams may at times lose their creativity if the norming behaviors become too strong and begin to stifle healthy dissent, which should be avoided. In this phase of development, individuals take more responsibility for making decisions and for their professional behaviour.

Assuming the transition has been successful, and the management of the team has been effective, eventually the move toward 'performing' is observed. Teams begin to function as a unit without inappropriate conflict or the need for external supervision. Team members become interdependent and are motivated and knowledgeable. The team members are now competent, autonomous and are able to handle the decision-making process without supervision. Any dissent is expected and allowed through means acceptable to the team.

An environment exists in some laboratories where open and transparent discussion is the norm, where structured arbitration and honest debate is encouraged, and where hierarchical dominance of individuals is replaced by self regulation, self critique and a healthy questioning disposition. Such a working environment can offset and mitigate against the worst effects of 'group think'. This in turn can facilitate a move away from the age-old problem in the fingerprint community of the 'buddy system' and hierarchical bullying. Many laboratories are moving toward a working environment that reflects this culture. However, such change must be underpinned by an understanding of how we as human beings react, not only to each other, but also to our working environment and technological tools such as AFIS.

It is inappropriate to design educational packages, technological advances or laboratory processes without taking the human being into account first. Individuals are complex and there is rarely a simple solution to any given problem.

This chapter has only addressed one very small feature of all the possible social psychological influences within the examination processes. It is also acknowledged that there are many other technical skills that must be trained and understood by the individual, such as ACE-V and the concepts of biological uniqueness and embryology. However, the science of fingerprint analysis cannot be divorced from the mind in this regard, and unless there is a healthy awareness of the mental, as well as the technical shortcomings of latent print analysis, and the way in which humans and their interactions impact upon this process, then full appreciation of the science associated with fingerprints is incomplete.

At a time when there has never been so much scrutiny of the fingerprint profession, it is incumbent upon practitioners and the wider management hierarchy to acknowledge problems and shortcomings, not by assuming a defensive posture, but by actively embracing such criticism and looking to the future.

Critics of fingerprint analysis as a forensic tool, argue that notions of scientific validity and practitioner error rates have been inadequately addressed. The first of many such steps to address these issues, would be to accommodate those in academia who, while identifying problems within the science, are at least willing, and able to work with the forensic community to bring awareness and understanding to some of these issues. This chapter has only discussed in detail one possible solution to maximizing the potential of examiners within the laboratory and avoiding the pitfalls of group conformity and to create a culture and a philosophy that is more attuned to scientific endeavor. This is how the profession can minimize risk, engender public trust and deliver accountability and scientific respect.

Chapter 10: Thesis Summary

Is it possible to have scientifically validated Standards in fingerprint examination? Or are such standards doomed to be based on expediency?

The fingerprint profession has spent decades contemplating the scientific background of fingerprint analysis and the technology that is associated with examination and comparison. In doing so, the true tool of fingerprint examination (the human) has been largely ignored as an important factor for accurate and reliable forensic science.

The body of work within this thesis has provided evidence that fingerprint examiners are highly motivated individuals who have strong feelings about the service they provide and their desire to 'catch the bad guys'. This predisposition to aligning their skills and effort to supporting the prosecution is not in the spirit of impartial scientific examination of the evidence. Further research has highlighted the examiner's susceptibility to biasing contextual information slanting decision making outcomes, especially where the finger marks being compared are of poor quality and ambiguous in appearance. This thesis has also shown that examiners interact with technological tools in an inconsistent way that results in evidential potential being lost or ignored because of the failure to distribute cognitive processes effectively between the human and the technology. These cognitive failings are compounded by other research within this thesis that highlights inconsistency from one examiner to the next in the way they analyse the fingerprint material.

The fingerprint profession (Charlton 2005) has neglected to notice that while there are computers associated with automated fingerprint identification, the human examiner has not been assessed with respect to its fitness for purpose given the high cognitive demands placed upon the fingerprint expert in the modern policing environment (Charlton 2005). Some practitioners in the domain have understood the need to consider the human examiner and their skills and motivations as an important factor in serving more effective forensic science. Jeff Gold at the 2003 National Fingerprint Conference suggested that changes to best practice in the UK fingerprint Bureaux could only be achieved by skilled, motivated and well trained staff (Charlton 2004). Fiscal constraints and the need for radical changes to working practices and processes require, more than ever, that the practitioner and the mind is understood so that required changes made to implement technology and or new processes such as reducing verification checks from 3 to 2 and lights out processing can be validated scientifically before introduction (Charlton 2004, 2006).

The entire forensic community is now working within an environment where public scrutiny and that of the judiciary demands that the scientific principles of both the applied scientific values of uniqueness as well as that of expert human performance will be simultaneously challenged in the future and the community must collectively work with academia and other partners to provide reassurances that evidence is both safe and can be relied upon (Charlton 2002, 2004). Indeed, the profession must now consider removing itself from its position of insularity and embrace the wider scientific community to seek answers to the very important questions facing the domain today (Charlton 2002). This process cannot be achieved overnight and will take an evolution to some degree of acceptance of some of the arguments put forward in this thesis which may ultimately, as a groundswell of research and validation emerges collectively manifest itself as a forensic revolution (Charlton 2005).

Fingerprint examiners must be dispassionate and removed from the centre of the argument (Charlton 2006) in deciding upon the outcome of a fingerprint examination. This requires on the part of the examiner controlled objectivity immune from the pressure to make a decision that is anything other than a reflection of a sound scientific method.

In many countries the individualization of a single finger mark from a crime scene to a suspect can lead to the conviction of a perpetrator of a crime, and in some cases can mean long terms of incarceration and even execution. The burden and responsibility placed upon a latent print examiner is great. In the compilation of evidence and the ultimate presentation of that evidence in the courts, the fingerprint expert is a figure of respect and authority. The judicial system places trust in the ability and competence of the expert witness. Jurors place their faith and confidence in the word and opinion of the individual presenting fingerprint evidence. To demonstrate that expert witnesses are credible they must be sure and resolute in their decision making skills, and be certain in the knowledge that what is presented as evidence is as reliable as possible.

A fingerprint expert is indoctrinated from the earliest days of training to be certain in their decision, never waiver, and to understand that there is no gray area in a fingerprint comparison. The conclusion reached can only ever be individualization, or exclusion, or if the clarity within the comparison is so poor as to make the decision inconclusive. Either way, the fingerprint expert is expected through training and experience to be able to differentiate between finger marks of varying quality thresholds. The expert is expected to reach an absolute conclusion that when explained to the court will leave the judiciary and jurors in no doubt as to the accuracy of the evidence being presented. If we assume that the acceptance of fingerprint evidence, or any other form of forensic evidence is based on trust by the judiciary and public in the credentials of the practitioners and scientists, then what should happen to the weight of that evidence and the trust placed in it when the levels of reliability within the individual fingerprint examiner waivers?

Even if we assume the evidence in the courtroom is sound, and that the judiciary accept the evidence, we are still left with another issue, namely, the way it is presented may have an impact on the jury (Charlton 2003). It can be speculated that how evidence is presented (framed and in what context) will have an impact on jury decisions. Such notions require still further research and understanding in the fingerprint domain, especially as it considers moving toward a probabilistic approach to fingerprint evidence presentation, similar to that of DNA.

From my own experience it is troubling that fingerprint examiners are sometimes unable to make a decision on a particular fingerprint comparison on one particular day, only to be able to make a clear and cogent decision the very next day on the same piece of evidence. Such daily variables in ability to make a decision cannot be acceptable if it is acknowledged that fingerprint expert training, experience and adherence to strict examination protocol and procedure should negate such swings in decisiveness. If all known procedures (ACE-V) are applied diligently it must be conceded that other factors are relevant to the decision making process. Research thus far within this thesis, indicates that context, emotion and the use of cognitive technology, aligned with ineffective recruitment and calibration tools all have a strong influence on the fingerprint examiner.

Fingerprint expert examiners are not supposed to exhibit behavioral traits associated with uncertainty or self doubt. It is a sign of weakness in the eyes of many within the fingerprint profession to display anything other than absolute certainty. Through objective analysis using a methodology known as ACE-V (Ashbaugh 1999), that is to say through careful assessment, comparison and evaluation, the expert is expected to reach a conclusion as to individualization, or not as the case may be, with absolute certainty. These findings are then verified through repeatable procedure by fellow expert examiners who will confirm or counter the decisions of the original examiner's decision.

This thesis has demonstrated that there is huge scope for subjectivity in the processes and procedures of fingerprint analysis; there is also an element of subjectivity in the interpretation of minutiae within the detail that constitutes friction ridge skin. This conflict between objective analysis methodology and the visual subjectivity of the examiner has not until recently been explored in the domain of forensic science. This divergence of philosophy has contributed, in part, to recent controversies within the fingerprint analysis domain. Controversy that has led to the criminal justice system, academics, the public and the media questioning the very fabric of fingerprint evidence as a reliable and trustworthy science.

The case of Shirley Mckie in Scotland has sent shock waves around the world amongst the fingerprint and wider forensic scientific community and has indicated to an increasingly skeptical public the level of subjectivity that is applied to the science of fingerprint analysis. But a not insignificant number of individuals from the fingerprint community still assert that the finger mark is indeed that of the aforementioned former police officer. Michael Mansfield QC has called for a public enquiry stating that the science of fingerprints needs to be re-examined (O'Neill 2006). Academic critics of fingerprint science have also been vocal on the evidential reliability of the processes and methodology. Simon Cole, a criminologist at the University of California has stated that fingerprint matching is undoubtedly a valuable tool for catching criminals but nobody knows how often examiners make a wrong call (Bamber 2005). Dr Itiel Dror suggests the mind is not a camera. It is a dynamic machine that can distort what it sees. Perception is far from perfection (Dror 2005). If Dror is correct then how can a fingerprint examiner have total faith and trust in what they see? This premise may lie at the heart of the very subjectivity to which I have alluded to throughout this thesis.

Friction ridge skin is elastic and malleable. Deposition pressure can affect the way friction ridge skin may present itself to the examiner. Likewise, movement and slippage at the time a latent print is deposited on a surface will affect the appearance and quality of the crime scene mark. All such factors require interpretation and understanding by the examiner before a conclusion on individuality can be reached. Relative correlation between points of reference on a finger mark from a crime scene are used during the comparison process when comparing features on the crime scene mark with a known exemplar of a suspect taken under laboratory conditions. The relative configuration of such reference points is important to the identification process.

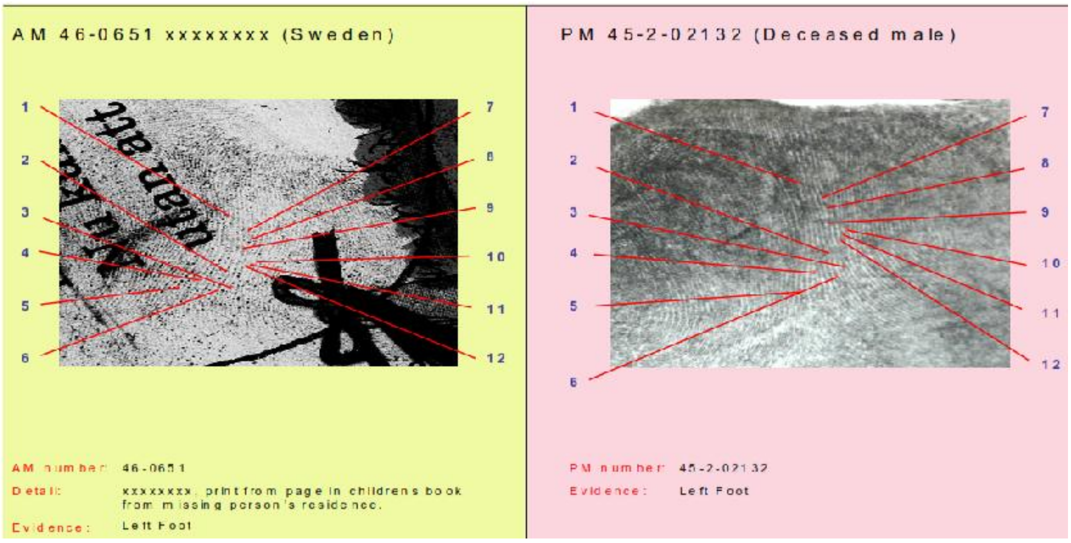


Figure 72 Example of a comparison chart using relative features in friction ridge skin from a known exemplar and an unknown source.

Ridge detail such as distance between individual features, ridge thickness and ridge counts between minutia all influence the decision of the examiner. Should the quality of a crime scene mark be very poor then it can be argued that the human eye could be fooled into seeing something within the print that isn't there? The context of the visual analysis can impact upon what the examiner may observe (see Figure 73).

Figure 73 Visual interpretation of the central figure in the chart above will depend on the visual context placed before the examiner.

While such visual illustrations are fun to look at, they do at their core have a very serious message for fingerprint experts. Namely that what is in the mind's eye is not always what is actually there. This is a fundamentally dangerous phenomenon for any scientists involved in visual interpretation, whether it is fingerprint recognition, facial recognition, or in trying to recognize a bomb in a suitcase at an airport security lounge. Such visual effects are in themselves enough to potentially influence the decision-making processes of fingerprint experts. But research in this thesis has also identified another potentially dangerous influence on the expert, namely contextual top down processing. Research in this thesis has suggested that emotions of individuals at the time of a fingerprint comparison could influence their decision thresholds, and ultimately, decision outcomes. It has been shown that fingerprint examiners are affected by top-down manipulations on emotional responses engendered by disturbing crime scene imagery. Where the fingerprints to be compared were classed as difficult, or ambiguous, then examiners were more likely to make affirmative judgment calls if associated within a disturbing context. Findings in research conducted thus far have highlighted that fingerprint experts are vulnerable to contextual influences. It will be important to try to ascertain how these influences manifest themselves, how such influences might be minimized or eradicated, and what steps could be taken to enhance training, accreditation and methodology in the future to improve the resistance of experts to such influences.

Future research will further investigate the training of forensic experts to identify how such training might be improved through better understanding of the psychological influences on the human being. By understanding how humans and especially forensic experts visualize images we can be better placed to offer enhanced training packages for fingerprint experts to alleviate many of the potential weaknesses in the methodology of fingerprint comparison. Finally, it will be important to investigate the decision-making processes of fingerprint experts within the real world domain where technology is placing an increasingly prominent role. It will be important to assess how new technologies are either strengthening or diluting the decision-making abilities of experts.

Another key strand of investigation not covered in this thesis will be the way in which forensic evidence is both served on the court and how the public and judiciary alike receive it. For example, should fingerprint evidence be presented using word of mouth only, or do visual presentations convey the message within the evidence more effectively.

Ditto et al (Ditto et al 1992) have shown that information consistent with a preferred conclusion is examined less critically than information that is inconsistent with a preferred conclusion. This is noteworthy when applied to the verification process for fingerprint matches. It could be that knowledge of a positive outcome or identification of a fingerprint may influence the verifier. It could be that they would look at the detail within the fingerprint less critically on verification if they prefer to see a positive outcome, ie, a murder case where the conviction of the ‘bad guy’ is paramount by the public, police and wider judiciary. This merits further investigation.

Within this thesis I have offered up challenges and maybe provoked thought. The aspirational target of enhanced quality and standards is a desire within the fingerprint domain and also within other forensic disciplines and sciences. Such ambitions cannot really be argued with. To define those standards and to ensure competency is achieved and improved upon, first you have to know what it is you are looking for in terms of competence before you can empirically measure it and act upon shortcomings. There then has to be a definition of what is meant by standards, whether the discussion had is about standards of the science or the standards of expediency or standards of the individual.

Credibility and reputation of fingerprint practitioners and the performance of laboratory and bureau staff is crucial and under constant scrutiny. Fingerprint identification remains one of the most important and valuable of all forensic disciplines and it is important to continually strive to minimize all potential erroneous identifications. While this is indeed a worthy aspiration, there are some impediments to the goal of scientifically based recruitment, training, and competency testing; namely that most fingerprint examiners are not trained researchers in behavioral sciences and have little interest or ability in this regard. The reasons for the dearth of research by fingerprint practitioners arises from a multitude of issues around the culture of the profession that sees itself as beyond reproach for the most part, a resistance to change, but perhaps just as importantly, there may be a concern among some aspiring researchers in the domain of the implications around attitudes and perceptions should the research produce conclusions that are contrary or against accepted practices. The fear of peer ridicule and even disciplinary action is often enough to dissuade those who might wish to explore new ways of working or counter previously held beliefs. This is inherently unscientific.

In summary, this thesis and the research conducted suggests a need to move toward and adopt a series of protocols and practices that will minimise the risk to examiners of susceptibility to the sort of vulnerabilities highlighted in this thesis. Some of these ideas are new, but some are re-affirmed here as it can be argued that some protocol while seemingly obvious in the wider forensic domain, appear to be ignored by the fingerprint profession specifically. These recommendations are not exhaustive but may include for example:

Table 24 Recommendations for the Fingerprint Profession

1	Contemporaneous supporting notes or materials should document the examination of fingerprints to make the interpretive process as transparent as possible.
2	Modifications to the results of any stage of latent print analysis (e.g., feature selection, utility, and assessment of distortion) after seeing a known exemplar should be viewed with caution.
3	Procedures should be implemented to protect examiners from exposure to extraneous (domain irrelevant) information in a case.
4	When comparing latent prints to exemplars generated through AFIS searches, examiners must recognize the possibility and dangers of incidental similarity.
5	A testifying expert should be familiar with the literature related to error rates, bias, cognition and the potential vulnerabilities within the fingerprint domain.
6	A structured approach that involves users throughout the entire design and implementation process should be followed when designing technology systems and the physical work environment.
7	Trainees should receive education in the scientific method, reading and understanding relevant scientific literature, communication skills, and methods for logically developing conclusions.
8	Management should establish policies and procedures for case review and conflict resolution, corrective action, and preventive measures. These policies should include tracking errors and the human factors associated with them.
9	Management should foster a culture in which it is understood that some human error is inevitable and that openness about errors leads to improvements in practice.

Given all the findings in this thesis, as well as other research being conducted to better understand the human condition as it relates to forensic science and the interpretation of pattern based evidence, only then, as a fingerprint expert, as I strive to understand my cognitive processes, can I begin to regain my own self-trust in my abilities to accurately interpret fingerprint evidence. By conducting such research as described in this thesis, the fingerprint profession as a whole can engage the trust of the public and wider criminal justice system that fingerprint science is as reliable and methodologically sound as it can be. If there is one overarching message to take from this thesis, it is that no human endeavour can ever be free from error. Fingerprint examiners have traditionally held the belief that methodology and the weight of time without close scrutiny by either the judiciary or academia have availed them the right to be immune from the rigors of scientific scrutiny.

However, this position has changed over the past 10 years and now those traditional values around the accuracy and efficiency of fingerprint examination have been laid bare. However, rather than sounding the death knell for the science, it may have ignited a renaissance within the profession, promoting new thinking, new approaches and new research to both underpin, as well as to enhance knowledge around cognitive processes to enhance both methodological as well as technological development. This is to be welcomed.

The fingerprint profession, the practitioners who work within the domain and the science upon which it is based has been around for decades, if not centuries. Only now is the profession beginning to understand the true dangers involved in the way forensic science is managed and carried out from a human perspective. The human is the tool, the instrument in the process that is most important. Without a true understanding of the nature of human endeavour, expertise and cognition will it ever be possible to find ways of doing more to guard against the potential for error in fingerprint, and other pattern recognition science.

The industrial revolution began in a haphazard and unplanned way. New ideas, new technologies and new practices reformed the way manufacturing was conducted. Some practices were good, some were bad. Some were safe, some were not. Some people even resisted change and were prepared to resist all attempts at changing their perspective and cultural attitudes. These ‘Luddites’ were prepared to face sanction and even prosecution to avoid changing and accepting the new ways. In the fingerprint profession, there are modern day equivalents, other such ‘Luddites’ who either will not, or cannot change their attitudes and accept that there are changes that may be needed. Similar to the latter day industrial revolution, technology and new ideas abound in forensic science. Law enforcement adopts these new ideas in a hap hazard way, just like the mill owners of 200 years ago. Some ideas are good, some are bad, some are safe and some are unsound. If the fingerprint profession is to evolve into the reliable science required by modern day policing and the wider society that demands justice and fairness in the way forensic science is conducted and deployed, then the Luddites within the profession must accept the changes are needed, are coming, and are necessary. The legacy of the 'industrial revolution' in the fingerprint domain will be a science that is robust in the face of close scrutiny, practitioners that are fit for purpose and able to perform the tasks asked of them with due diligence and sound knowledge, aligned with recruitment, calibration and cognitive awareness tools that will provide a better understanding of the vulnerabilities of the human condition.

In 1901 Sir Edward Henry said the following about fingerprint uniqueness, 'there is no sign (except in one case) of change (in human friction skin) through life', we are 'justified in inferring that between birth and death there is absolutely no change, in say 699 out of 700 characteristics' (on a hand). So much for the one dissimilarity doctrine?

If this thesis alludes to anything, it is that examiners must beware the 1 in 700 fingerprints that is susceptible to change through life and that the instrument of fingerprint examination, the human, is as vulnerable to cognitive phenomena as any other human being, which may result in the misinterpretation of the 1 in 700.

References

ACPO (Association of Chief Police Officers) (2006): National Fingerprint Manual, Issue 1, ACPO National Fingerprint Board.

Adler J. (2004) Forensic Psychology: Concepts, Debates and Practice. Ed. Williamson T. Oregon, US: Willan Publishing.

Ainsworth P B. (2002) Psychology and Policing. West Sussex, UK: Willan.

Ainsworth P B. (2005) Psychology, Law and Eyewitness Testimony. West Sussex, UK: Wiley.

Alam, M.S. Akhteruzzaman, M., & Cherri, A. K. (2004). Real-time fingerprint identification. Optics and Laser Technology, 36, 191–196.

Aleman A, Nieuwensteina, M.R., Böckerc, K.B.E. and Haana, E.H.F. (2000). Music training and mental imagery ability. Neuropsychologia, 38 (12), 1664-1668.

Allard, F & Starkes, J.L. (1980). Perception in sport: Volleyball.
Journal of Sport Psychology, 2, 22-33.

Anonymous Contributors from serving police officers and forensic scientists (1951). The Boy's Book of Scotland Yard. London: Burke.

Asch S. (1955) "Opinions and social pressure." Scientific American 193: 31-55.

Ashbaugh D. (1982) Ridgeoscopy the time is now. Fingerprint Whorld 8[30].

Ashbaugh D. (1985) The key to fingerprint identification. Fingerprint Whorld 10[40], 94-96.

Ashbaugh, D. R. (1991) "Ridgeology". Journal of Forensic Identification Vol 41 (1) ISSN: 0895-l 73X

Ashbaugh D. (1991) Ridgeology. Fingerprint Whorld 17[66].

Ashbaugh D. (1992) Ridgeology Pt 2. Fingerprint Whorld 18[68], 45-55.

Ashbaugh D. (1992) Ridgeology Pt 3. Fingerprint Whorld 18[67], 5-12.

Ashbaugh, D. (1999). Quantitative-qualitative friction ridge analysis: an introduction to basic and advanced ridgeology. Boca Raton, FL: CRC Press.

Ashbourn J. (2000) Biometrics: Advanced Identity Verification: The Complete Guide. London, Springer-Verlag.

Ashworth, A.R.S. and Dror, I. E. (2000). Object Identification as a Function of Discriminability and Learning Presentations: The Effect of Stimulus Similarity and Canonical Frame Alignment on Aircraft Identification. Journal of Experimental Psychology: Applied, 6 (2), 148-157.

Ask K and Granhag P A. (2003) Individual Determinants of Deception Detection Performance: Need for Closure, Attributional Complexity and Absorption. Goteborg Psychological Reports 33[1].

Ask K and Granhag P A. (2005) Motivational Sources of Confirmation Bias in CriminalInvestigations: The Need for Cognitive Closure. Journal of Investigative Psychology and Offender Profiling 2, 43-63.

Ask K and Granhag P A. (2006) Hot Cognition in Investigative Judgments: The Differential Influence of Anger and Sadness. Law and Human Behaviour 3. American Psychology-Law Society.

Ask K and Granhag P A. (2007) Motivational Bias in Criminal Investigators' Judgments of Witness Reliability. Journal of applied social psychology 37[3], 561-91. 2007.

Babler W J. (1977) The prenatal origins of population differences in human dermatoglyphics. Dissertation to University of Michigan , 1-184.

Babler W J. (1978) Prenatal Selection and dermatoglyhic patterns. American Journal of Physical Anthropology 48[1], 21-27.

Babler W J. (1987) Prenatal Development of Dermatoglyptic Patterns : Associations with Epidermal Ridge, Volar Pad and Bone Morphology. Collegium Anthropologium 11[2], 297-304.

Back K. (1951) Influence through social communication. Journal of Abnormal and Social Psychology 46, 9-23.

Backan J. (2004) The Corporation: The Pathological Pursuit of Profit and Power. Simon and Schuster, New York.

Baddeley, A.D. and Hitch, G.J.L (1974). Working memory, In G.A. Bower (Ed.), The psychology of learning and motivation: advances in research and theory (Vol. 8, pp. 47-89). New York: Academic Press.

Baddeley, A.D. (1992). Working memory. Science, 255, 556-9.

Bailey C, Daily C, and Phillips T. (2006) A Study of Kruglanski's Need for Closure Construct and Its Implications for Judgment and Decision Making in Accounting and Auditing. Social Science Research Network .

Bain A. (1985) The importance of fingerprints. Fingerprint Whorld 10[40].

Balcetis E and Dunning D. (2006) See what you want to see: Motivational influences on visual perception. Journal of Personality and Social Psychology 91[4], 612-25.

Bamber R: (2005) Fingerprints fail reliability test, The Times, Public Agenda 20 September 2005.

Beavan C. (2001) Fingerprints: Murder and the race to uncover the science of identity. London, Harper Collins.

Begley S. (2005) Fingerprint Matches come under more fire as potentially fallible. The Wall Street Journal 176.

Bektas S, Bahadir B, Kandemir N O, Barut F, Gul A E, Ozdamar S O (2009) Intraobserver and interobserver variability of fuhrman and modified fuhrman grading systems for conventional renal cell carcinoma, Kaohsiung J. Med. Sci. 25 596–600.

Berry J. (1976) Fingerprints in Antiquity. Fingerprint Whorld 2[6], 23-24.

Berry J. (1977) The Pseudo Expert. Fingerprint Whorld 3[9].

Berry J and Leadbetter M. (1978) Fingerprint Transgressions. Fingerprint Whorld 4[13].

Berry J. (1991) In two minds. Fingerprint Whorld 16[64].

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. Psychological Review, 94, 115-147.

Binet, A. (1905). New methods for the diagnosis of the intellectual level of subnormals. L'Année Psychologique, 12, 191-244. (reprinted in E. S. Kite, The development of intelligence in children (1916). NJ: Vineland).

Blank, H., Musch,j., & Pohl, R., F. (2007) Hindsight Bias: On Being Wise After the Event. Social Cognition, 25(1), 1-9.

Block E. (1970) Fingerprinting. London, Franklin Watts Ltd.

Bodner, G. M. and Guay, R. B. (1997). The Purdue visualization of rotations test. The Chemical Educator, 2 (4), 1-17.

Boehm-Davis, D. A., Holt, R. W., and Hansberger, J. (1997). Pilot abilities and performance. Proceedings of the Ninth International Symposium on Aviation Psychology. Columbus, OH: The Ohio State University Press.

Boiney L G, Kennedy J, and Nye P. (1997) Instrumental bias in motivated reasoning: More when more is needed. Organisational behaviour and human decision processes 72[1], 1-24.

Borman, W. C., Hanson, M., and Hedge, J. W. (1997). Personnel selection. Annual Review of Psychology, 48, 299-337.

Borman, W. C., Hedge, J. W., Ferstl, K.L., Kaufman, J.D., Farmer, W. L, and Bearden, R.M. (2003). Current directions and issues in personnel selection and classification. Research in Personnel and Human Resources Management, 22, 287-355.

Boudreau L A, Baron R, and Oliver P V. (1992) Effects of Expected Communication Target Expertise and Timing of Set On Trait Use in Person Description. Personality and Social Psychology Bulletin 18, 447-52.

Bower, J. M. (1990). Reverse engineering the nervous system: An anatomical, physiological, and computer-based approach. In S.F. Zornetzer, J. L. Davis, and C. Lau (Eds), An Introduction to Neural and Electronic Networks (pp. 3-24). San Diego: Academic Press.

Bregman, N. J., & McAllister, H. A. (1987). Perceived innocence of guilt: role of eyewitness identi cation and fingerprints. Southern Psychologist, 3(2), 49–52.

Bright D. A., & Goodman-Delahunty, J. (2004). The influence of gruesome verbal evidence on mock juror verdicts. Psychiatry, Psychology & Law, 11(1),154–166.

Britten N. (1995) Qualitative Research: Qualitative interviews in medical research. British Medical Journal 311, 251-53.

Brown D. (2005) Human Sight and the Identification Process "Focussed attention on an object does not guarantee accurate perception". University of Durham Diploma in Fingerprint Examination. Centrex/NTC.

Busemeyer J R and Townsend T T. (1993) Decision Field Theory: A Dynamic-Cognitive Approach to Decision Making in an Uncertain Environment. Psychological Review 100[3], 432-59.

Busey, T. A. and Vanderkolk, J. R. (2005). Behavioral and electrophysiological evidence for configural processing in fingerprint experts. Vision Research, 45, 431–448.

Busey T, Champod C, and Langenburg G. (2008) Amicus Brief State of Maryland v Bryan Keith Rose. Case No. 03K06000545.

Busey T, Dror I E (2010) Special Abilities and Vulnerabilities in Forensic Expertise in Fingerprint Sourcebook, NIJ Press, Washington, DC, USA, (Chapter 15).

Byrd, J. S. and Bertram, D. (2003). Form-blindness. Journal of Forensic Identification, 53 (3), 315-331.

Byrne, A., & Eysenck, M. W. (1993). Individual differences in positive and negative interpretive biases. Personality and Individual Differences, 14, 849–851.

Calvo, M. G., & Eysenck, M. W. (1998). Cognitive bias to internal sources of information in anxiety. International Journal of Psychology, 33, 287–299.

Caminiti, M.F. (2000). Psychomotor testing as predictors of surgical competence: The predictive value of aptitude tests. Focus Surgical Education, 2 (17), 20–22.

Carter, C. S., LaRussa, M. A., and Bodner, G. M. (1987). A study of two measures of spatial ability as predictors of success in different levels of general chemistry. Journal of Research in Science Teaching, 24, 645.

Chaiken S, Liberman A, and Eagly A H. (1989) Heuristic and systematic information processing within and beyond the persuasion context. Ed Uleman J S and Ed Bargh J A. Unintended Thought. 212-52. New York, NY, The Guilford Press.

Charlton D. (2001) Comment. Fingerprint Whorld 27 (106), 133-135.

Charlton D. (2002) Animal Farm. Fingerprint Whorld 28 (107), 5-6.

Charlton D (2002) Kobayashi Maru. Fingerprint Whorld 28 (108), 65-67.

Charlton D. (2002) The Bunker Mentaility. Fingerprint Whorld 28 (110), 193-194.

Charlton D. (2002) Report Into the Chicago Fingerprint Forum (Post Daubert). Fingerprint Whorld 28 (110), 215-219.

Charlton D. (2003) Comment. Fingerprint Whorld 29 (112), 65.

Charlton D. (2003) To Swim the English Channel Only to Drown in the Harbour. Fingerprint Whorld 29 (113),125-126.

Charlton D. (2003) Closing the Criminal Justice Gap. Fingerprint Whorld 29 (114), 201.

Charlton D. (2004) Counting the Cost. Fingerprint Whorld 30 (115), 5.

Charlton D. (2004) 2000 Years of History in a Soundbite!! Fingerprint Whorld 30 (115) 6-8.

Charlton D. (2004) A Biometric Future. Fingerprint Whorld 30 (116), 41.

Charlton D. (2004) Keeping it Simple. Fingerprint Whorld 30 (117), 101.

Charlton D. (2004) Bete Noire!! Fingerprint Whorld 30 (118), 161-162.

Charlton D. (2005) The Greenhouse Effect. Fingerprint Whorld 31 (119), 5-6.

Charlton D. (2005) Devolution or Dissolution? Fingerprint Whorld 31 (120), 73-74.

Charlton D. (2005) The Forensic Family. Fingerprint Whorld 31 (121), 129-130.

Charlton D. (2005) Evolution or Revolution? Fingerprint Whorld 31 (122), 205.

Charlton D. (2006) Two for the Price of One. Fingerprint Whorld 32 (123), 5.

Charlton D. (2006) Six of One. Fingerprint Whorld 32 (125), 145.

Charlton D. (2006), Forensic Expertise and Self Trust, Forensic Update (Division of Forensic Psychology), the British Psychological Society, Number 87 P14-19.

Charlton D, Del-Manso H, and Dror I E. (2007) Expert error: the mind trap. Fingerprint Whorld 33 (128), 151-55.

Charlton D. (2007) Group Networks and Dynamics Associated with Forensic Fingerprint Analysis and Decision-Making. Canadian Identification Society. Identification Canada 30, (4) 126-129.

Charlton D and Galloway V. (2007) Forensic Human Identification: An Introduction (Chapter 3. Latent Print Identification). Edited by T Thompson and S Black. 57-72. CRC Press Taylor and Francis Group.

Charlton D (2008), Chapter…..Finding latent prints at crime scenes, Text Book: Crime Scene Management…Wiley Publishing Group.

Charlton D. Fraser-McKenzie P, Dror I E. (2010) Emotional Experiences and Motivating Factors Associated with Fingerprint Analysis. Journal of Forensic Sciences. 10.1111/j.1556-4029.2009.01295.

Chiu M and Khoo L. (2003) Rudeness and Status Effects During Group Problem Solving: Do They Bias Evaluations and Reduce the Likelihood of Correct Solutions? Journal of Educational Psychology 95[3], 506-23.

Coghlan A. (2005) How far should prints be trusted? New Scientist .

Cohen J A. (1960) Coefficient of agreement for nominal scales. Educational and psychological measurement 20, 37-46.

Cole, S. (1999). What counts for identity? The historical origins of the methodology of latent fingerprint identi cation. Science in Context, 12(1), 139–172.

Cole S. (2002) Latent print evidence and expert knowledge. Fingerprint Whorld 28[107], 23-46.

Cole S A (2004) A Grandfathering of Evidence: Fingerprint Admissibility Rulings From Jennings To Llera Plaza And Back Again. American Criminal Law Review, 41 Am. Crim. L. Rev. 1189

Cole S. (2005) More than zero: accounting for error in latent fingerprint examination. The journal of criminal law and criminology. 95[3], 985-1078.

Cole S. (2006) Is Fingerprint Identification Valid? Rhetorics of Reliability in Fingerprint Proponents' Discourse. Law and Policy 28[1], 109-35.

Cole S. (2006) The prevalence and potential causes of wrongful conviction by fingerprint evidence. Golden Gate University Law Review 37, 39-105.

Cole S. (2008) The 'Opinionization' of Fingerprint Evidence. Biosocieties 3, 105-13.

Cook, N. D., Früh, H., and Landis, T. (1995). The cerebral hemispheres and neural network simulations: Design considerations. Journal of Experimental Psychology: Human Perception and Performance, 21, 410-422.

Cooke T G. (1932) Fingerprints, Secret Service, Crime Detection. Chicago, US, Finger Print Publishing Association.

Cooper, L. A., and Podgorny, P. (1976). Mental transformations and visual comparison processes: Effects of complexity and similarity. Journal of Experimental Psychology: Human Perception and Performance, 2, 503-514.

Corbetta, M., Miezen, F. M., Schulman, G. L., and Petersen, S. E. (1993). A PET study of visuospatial attention. Journal of Neuroscience, 13, 1202-1226.

Coren, S., Ward, L.M., Enns, J. T. (1994) Sensation and Perception (4th Ed.) Toronto: Harcourt Brace.

Crick, F. (1989). The recent excitement about neural networks. Nature, 337, 129-132.

Cummins H. (1929) The topographic history of the volar pads in the human embryo.

Cummins H and Kennedy R. (1940) Purkinjes' Observations (1823) On Fingeprints and Other Skin Features. American Journal of Police Science 31[3].

Cusa N de. The Layman: About Mind. Miller C L. (1979). New York, US, Abaris Books.

Cuschieri, A, Francis, N., Crosby, J., and Hanna, G.B. (2001). What do master surgeons think of surgical competence and revalidation? American Journal of Surgery, 182, 110–116.

Czerwinski, M., Lightfoot, N., and Shiffrin, R. M. (1992). Automatization and training in visual search. American Journal of Psychology, 105 (22), 271-315.

Darley, J. M., & Gross, P. H. (1983). A hypothesis-confirming bias in labeling effects. Journal of Personality & Social Psychology, 44, 20–33.

Dascal, M. & Dror, I. E. (2005). Cognitive Technologies. Pragmatics & Cognition, 13 (3).

Dawes J. (2008) Do Data Characteristics Change According to the number of scale points used? An experiment using 5-point, 7-point and 10-point scales. International Journal of Market Research 50[1], 61-77.

Dechesne M, Janssen J, and Knippenberg A (2000). Derogation and distancing as terror management strategies: the moderating role of need for closure and permeability of group boundaries. Journal of Personality and Social Psychology 79[6], 923-32.

Deutsch M and Gerard H. (1955) A study of normative and informational social influences upon individual judgment. Journal of Abnormal and Social Psychology 51, 629-36.

DiFonzo J H and Stern R C. (2007) Devil in a white coat: the temptation of forensic evidence in the age of CSI. Heinonline 41, 503-32.

Ditto, P H; Lopez, D F. (1992) Journal of Personality and Social Psychology. Volume 63(4) 568-584

Dolgin, D.L. and Nontasak, T. (1990). Initial validation of a personnel selection system for landing craft air cushion (LCAC) vehicle operations. Proceedings of the Psychology in the Department of Defense 12th Symposium, USAFA-TR-90-1, 245-249, (AD A221 877).

Dosher, B. A. and Lu, Z. (1998). Perceptual learning reflects external noise filtering and internal noise reduction through channel reweighting. Proceedings of the National Academy of Science, 95, 13988-13993.

Dreyfus, H.; Dreyfus, S. (2005)."Expertise in real world contexts". Organization Studies 26 (5): 779 792. Doi: 10.1177/0170840605053102
Dror, I. E., Kosslyn, S. M., and Waag, W. (1993). Visual-spatial abilities of pilots. Journal of Applied Psychology, 78 (5), 763-773.

Dror, I. E. and Kosslyn, S. M. (1994). Mental imagery and aging. Psychology and Aging, 9, 90-102.

Dror, I. E., Ivey, C., and Rogus, C. (1997). Visual mental rotation of possible and impossible objects. Psychonomic Bulletin and Review, 4, 242-247.

Dror, I. E., & Kosslyn, S. M. (1998). Age degradation in top-down processing: identifying objects from canonical and noncanonical viewpoints. Experimental Aging Research, 24(3), 203–216.

Dror, I. E., Busemeyer, J. R., & Basola, B. (1999). Decision making under time pressure: an independent test of sequential sampling models. Memory and Cognition, 27(4), 713–725.

Dror, I. E. Charlton, D. & Peron A. E. (2004). Evaluating 'scientific' evidence for the court: What contributing factors are really involved in fingerprint identification? Conference of International Centre for Advanced Research in Identification Science Meeting, Sheffield, UK.

Dror I E (2005) Fingerprint doubt, Forensic News, Police Professional, 3rd November 2005.

Dror I E, (2005) Perception is far from perfection: the role of the brain and mind in constructing realities, Brain Behav. Sci. 28 (6).

Dror I E, Thomas R D, (2005) The cognitive neuroscience laboratory: a framework for the science of mind, in: C. Erneling, D. Johnson (Eds.), The Mind as a Scientific Object: Between Brain and Culture, Oxford University Press, pp. 283–292.

Dror, I.E., Peron, A., Hind, S., & Charlton, D. (2005). When emotions get the better of us: The effect of contextual top-down processing on matching fingerprints. Applied Cognitive Psychology, 19(6), 799-809.

Dror I E, (2005) Technology and human expertise: Some do's and don'ts, Biometric Technology. Today 13 (9) 7–9.

Dror I E. (2006) Don't forget us humans. The Times . 31-7-2006.

Dror, I.E., Charlton, D., & Peron A. (2006). Contextual information renders experts vulnerable to making erroneous identifications Forensic Science International, 156 (1), 74-78.

Dror, I.E. & Charlton, D. (2006). Why experts make errors. Journal of Forensic Identification, 56 (4), 600-616.

Dror I E. (2006) Cognitive Science Serving Security: assuring useable and efficient biometric and technological solutions. Aviation Security International 12[3].

Dror I E (2006) Cognitive Science Serving Security: assuring useable and efficient biometric and technological solutions. Aviation Security International: The Journal of Airport & Airline Security. Volume 12, Issue 3.

Dror I E, Charlton D, (2007) Why are Experts Prone to Error? Centre of Visual Cognition Technical Report, University of Southampton, UK.

Dror I E and Charlton D. (2007) Improving perception and judgment: An examination of expert performance. Fingerprint Whorld 33[129], 232-34.

Dror I E and Charlton D. (2007) Methodological and conceptual problems in measuring and enhancing expert performance. Fingerprint Whorld 33[129], 234-38.

Dror I E and Charlton D. (2007) The Psychology of Police Performance and Decision Making. Police Review.

Dror, I. E., Stevenage, S. V., and Ashworth, A. (2008). Helping the cognitive system learn: Exaggerating distinctiveness and uniqueness. Applied Cognitive Psychology, 22 (4), 573-584.

Dror I E. (2008) Biased Brains. Police Review 20, 21-23. 2008.

Dror I E, Rosenthal B, (2008) Meta-analytically quantifying the reliability and biasability of fingerprint experts' decision making, J. Forensic Sci. 53 (4) 900–903.

Dror I E and Fraser-Mackenzie P. (2008) Cognitive Biases in Human Perception, Judgment, and Decision Making: Bridging Theory and the Real World . K Rossmo (Ed). Criminal Investigative Failures. Taylor & Francis Publishing.

Dror, I. E. (2009) On Proper Research and Understanding of the Interplay between Bias and Decision Outcomes. Forensic Sci. Int., 191, 17–18.

Dror I E, (2009) How can Francis Bacon help forensic science? The four idols of human biases, Jurimetrics: J. Law Sci. Techn. 50 93–110.

Dror I E and Harnad S. (2009) Offloading cognition onto cognitive technology. Cognition distributed: How cognitive technology extends our minds. Amsterdam, John Benjamins.

Dror I E, Cole S, (2010) The vision in 'blind' justice: expert perception, judgment and visual cognition in forensic pattern recognition, Psychol. Bull. Rev. 17 (2) 161–167.

Dror I E, J.L. Mnookin, (2010) The use of technology in human expert domains: challenges and risks arising from the use of automated fingerprint identification systems in forensics, Law Prob. Risk 9 (1) 47–67.

Dror, I.E. (2010) Perceptual, Cognitive and Psychological Elements Involved in Expert Identification. Friction Ridge Sourcebook.

Dror I E, Wertheim K, Fraser-Mackenzie P and Walajtys J (2012). The Impact of Human-Technology Cooperation and Distributed Cognition in Forensic Science: Biasing Effects of AFIS Contextual Information on Human Experts. Journal of Forensic Sciences**.**

Dror, I. E. (in press). Paradoxical functional degradation in human expertise. In N. Kapur, Pascual-Leone, & V. S. Ramachandran (Eds.), The Paradoxical Brain. Cambridge, UK: Cambridge University Press.

Dror I E. (In Press) The paradox of human expertise: Why experts can get it wrong. In: Kapur N, editor. The Paradoxical Brain. Cambridge: Cambridge University Press.

Duncan J H. (1942) An Introduction to Fingerprints. London, Butterworth and Co (Publishers) Ltd.

Eberhardt, J. L., Goff, P. A., Purdie, V. J., & Davis, P. G. (2004). Seeing black: race, crime, and visual processing. Journal of Personality and Social Psychology, 87(6), 876–893.

Eggen D (2006) US settles suit filed by Ore. Lawyer: $2 million will be paid for wrongful arrest after Madrid attacks, Washington Post, 30th November, pA03.

Ellad E. (2003) Effects of feedback on the overestimated capacity to detect lies and the underestimated ability to tell lies. Applied Cognitive Psychology 17, 349-63.

Englich B and Mussweiler T. (2001) Sentencing Under Uncertainty: Anchoring Effects in the Courtroom. Journal of applied social psychology . VH Winston & Sons Inc.

Ericsson K A, N. Charness, P.J. Feltovich, in: R.R. Hoffman (Ed.) (2006) The Cambridge Handbook of Expertise and Expert Performance, Cambridge University Press, New York.

Ericsson K A (Ed.), (2009) Development of Professional Expertise, Cambridge University Press, New York.

Evans J. (1990) Bias in human reasoning: Causes and Consequences. Hove, UK, Lawrence Erlbaum Associates Ltd.

Evett I and Williams R L (1995) A review of the 16 point fingerprint standard in England and Wales. Fingerprint Whorld 21[82].

Evett I W (1996) Expert evidence and forensic misconceptions of the nature of exact science, Sci. Justice 36 118–122.

Eysenck, M. W., Macleod, C., & Mathews, A. (1987). Cognitive functioning and anxiety. Psychological Research, 49, 189–195.

Eysenck M W. (1997) Anxiety and cognition: a unified theory. Psychology Press . Hove, England.

Fairhurst D. (2003) Is fingerprint examination a science? Fingerprint Whorld 29[113], 139-43.

Fairhurst D. (2005) The Psychology of Latent Print Examination – A UK Perstective (Lecture). The Fingerprint Society Annual Educational Conference, Brighton .

Faulds H. (1912) Dactylography: Or the study of Fingerprints. Halifax, Raglan Works, Milner and Company.

Faulkner. H., Regehr, G., Martin, J., and Reznick, R. (1996). Validation of an objective structured assessment of technical skill for surgical residents. Academic Medicine, 71, 1363–1365.

Festinger L. (1950) Informal social communication. Psychological Review 57, 271-82. 1950.

Festinger L. (1954) A theory of social comparison procesess. Human Relations 7, 117-40.

Fleiss J L. (1981) Statistical methods for rates and proportions. (chapter 13: The measurement of inter-rater agreement) , 212-36. New York, US, 2nd ed. John Wiley & Sons.

Frieden T and Schuster H. (2006) Sloppy FBI work led to wrong man. CNN.

Fu J H, Morris M W, Lee S, Chao M, Chiu C, and Hong Y. (2007) Epistemic motives and cultural conformity: need for closure, culture, and context as determinants of conflict judgments. Journal of Personality and Social Psychology 92[2].

Galton F. (1883) Inquiries Into Human Faculty and Its Development. London, UK, Dent and Sons Ltd.

Galton F. (1892) Fingerprints. New York, MacMillan and Co.

Garn S M, Burdi A R, and Babler W J. (1975) Early Prenatal Attainment of Adult Metacarpal-Phalangeal Rankings and Proportions. American Journal of Physical Anthropology 43[3], 327-32.

Garvie M. (1999) Beyond all doubt a truly great man..the life and times of Sir Edward Henry. Fingerprint Whorld 25[95], 183-202.

Gazzaniga, M. S. (2005). The cognitive neurosciences. Cambridge: The MIT Press.

Gerber S. (1983) Chemistry and Crime: From Sherlock Holmes to Today's Courtroom. Washington, US, American Chemical Society.

Gluck, M. A. and Thompson, R. F. (1987). Modeling the neural substrates of associative learning and memory: A computational approach. Psychological Review, 94, 176-191.

Gobet, F.; Lane, P.C.R.; Croker, S.; Cheng, P.C.H.; Jones, G.; Oliver, I.; & Pine, J.M. (2001). Chunking mechanisms in human learning. Trends in Cognitive Sciences, 5, 236-243.

Gobet, F.; de Voogt, A. J.; Retschitzki, J. (2004). Moves in mind: The psychology of board games. Hove, UK: Psychology Press. ISBN 1841693367.

Gold J. (2004). Review of the use of NAFIS by Fingerprint Bureaux – Update. Home Office, Police Standards Unit.

Gray, R. (2009). A Model of Motor Inhibition for a Complex Skill: Baseball Batting. Journal of Experimental Psychology: Applied, 15 (2), 91–105.

Grew N. (1684) The description and use of the pores in the skin of the hands and feet. Philosophical Transactions of the Royal Society of London 14, 566-67.

Grieve D. (1996) Possession of Truth. Journal of Forensic Identification 46[5].

Grieve D. (2000) Built by many hands. Fingerprint Whorld 26[100], 51-60.

Grossberg, S. (1987). Competitive learning: From interactive activation to adaptive resonance. Cognitive Science, 11, 23-63.

Guilford, J. P., Fruchter, B., and Zimmerman, W. S. (1952). Factor analysis of the Army Air Forces, Sheppard Field Battery of Experimental Aptitude Tests. Psychometrika, 16, 45.

Halberstadt, J. B., Niedenthal, P. M., & Kushner, J. (1995). Resolution of lexical ambiguity by emotional state. Psychological Science, 6, 278–282.

Hall L J and Player E. (2008) Will the introduction of an emotional context affect fingerprint analysis and decision-making? Forensic Science International 181, 36-39.

Halpern, D. F., Loehlin, J. C., Perloff, R., Sternberg, R. J., and Urbina, S. (1996). Intelligence: Knowns and unknowns. American Psychologist, 51, 77-101.

Hama, I.J.M., Raemaekersb, M., Wezelc, R.J.A., Oleksiakc, A., and Postma, A. (2009) Categorical and coordinate spatial relations in working memory: An fMRI study. Brain Research, 1297, 70-79.

Hardin C D and Higgins E T. (1996) Shared reality: How social verification makes the subjective objective. Handbook of motivation and cognition 3, 28-84. New York, Guildford Press.

Harling M (1996) Origins of the New York State Bureau of Identification. NYS, Division of Criminal Justice Services.

Harmon, L.D. and Julesz, B. (1973). Masking in visual recognition: effects of two-dimensional filtered noise. Science, 180, 1194 - 1197.

Harnad S (Ed.) (1987) Categorical Perception: The Groundwork of Cognition, Cambridge University Press, New York.

Henry E. (1900) Classification and Uses of Fingerprints. Classification and Uses of Fingerprints. London, George Routledge and Sons Ltd.

Hermsen M and Have H. (2005) Decision making in palliative care practice and the need for moral deliberation: a qualitative study. Patient Education and Counseling 56, 268-75.

Hirsch, C., & Mathews, A. (1997). Interpretive inferences when reading about emotional events. Behaviour Research and Therapy, 35, 1123–1132.

Hofstadter, D. R. Gödel, Escher, Bach (1979) An Eternal Golden Braid; Basic Books: New York, 1979.

Hogarth R M. (1981) Beyond discrete biases: functional and dysfunctional aspects of judgmental heuristics. Psychological Bulletin 90[2], 197-217.

Hogg M A. (2000) Subjective uncertainty reduction through self categorisation: A motivational theory of social identity processes. European review of social psychology 11, 223-55.

Hollan, J., Hutchins, E., Kirsh, D. (2000) ‘Distributed cognition: toward a new foundation for human-computer interaction research’. ACM transactions on computer-human interaction, 7(2), 174-196.

Holt J. (1936) Fingerprints Simplified. Chicago, Frederick J Drake and Co.

Howel D C. (2002) Statistical Methods for psychology. 5th Edition. Pacific Grove, California, USA, Wadsworth.

Huber R. (1959) Expert witnesses. The Criminal Law Quarterly 2, 276-95.

Hughes G. (1994) Editorial. Fingerprint Whorld 20[77], 81.

Hughes G. (1998) Comment. Fingerprint Whorld 24[93], 93-95.

Hutchins, E., and Klausen, T. (2000) 'Distributed cognition in an airline cockpit. Cognition and communication at work'. In Cognition and communication at work, Y. Engström and D. Middleton, Eds. Cambridge University Press, New York, NY, 15-34.

Jackson J. (2002) Intergroup Attitudes as a Function of Different Dimensions of Group Identification and Perceived Intergroup Conflict. Self and Identity. Volume 1, Issue 1, Pages 11 - 33. DOI: 10.1080/152988602317232777

Jacobs, R.A., Jordan, M.I., and Barto, A.G. (1991). Task decomposition through competition in modular connectionist architecture. Cognitive Science, 15, 219-250.

Jager, G. and Postma, A. (2003). On the hemispheric specialization for categorical and coordinate spatial relations: a review of the current evidence. Neuropsychologia, 41 (4), 504-515.

Joffe H, Yardley L, Marks D, and Yardley L (Ed). (2004) Content and thematic analysis. Research Methods for Clinical and Health Psychology. London, UK, Sage.

Johnson S. (2007) Criminal Records Chief Criticised by McKie Report. Daily Mail , 31. 23-24.

Kamin, L.J. (1995). The pioneers of IQ testing. In R. Jacoby and N. Glauberman (Eds.), The Bell Curve debate: History, documents, opinions. New York: Times Books.

Kassin S. (1997) The Psychology of Confession Evidence. American Psychologist 52[3], 221-33.

Kavaklia, M. and Gerob, J. S. (2002). The structure of concurrent cognitive actions: a case study on novice and expert designers. Design Studies, 23 (1), 25-40.

Kent T. (1993) Fingerprint Development Handbook. Manual of Fingerprint Development Techniques . UK Home Office. 2nd Edition.

Klayman J and Ha Y W. (1987) Confirmation, disconfirmation and information in hypothesis testing. Psychological Review 94, 211-28.

Klayman J. (1995) Varieties of Confirmation Bias. The psychology of learning and motivation 32, 385-419.

Klug D, Petersen J, Stoney D. (1992) Automated fingerprint identification systems: Their acquisition management performance and organizational impact. Washington DC: National Institute of Justice, Report No.: NCI 13749.

Knudsen, E. (2007). Fundamental Components of Attention. Annual Review of Neuroscience, 30 (1), 57–78.

Koppl R. (2005) Diversity and Forensics: Diversity in Hiring is Not Enough. Medicine Science and the Law.

Kosnik L D. (2005) Refusing to Budge: A confirmatory bias in decision making. US, University of Missouri.

Kosslyn, S. M. (1987). Seeing and imagining in the cerebral hemispheres: A computational approach. Psychological Review, 94, 148-175.

Kosslyn, S. M., Chabris, C. F., Marsolek, C. J., and Koenig, O. (1992). Categorical versus coordinate spatial relations: Computational analyses and computer simulations. Journal of Experimental Psychology: Human Perception and Performance, 18, 562-577.

Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Chabris, C. F., Rauch, S. L., and Anderson, A. K. (1994). Identifying objects seen from different viewpoints: A PET investigation. Brain, 117, 1055-1071.

Kosslyn, S. M., Chabris, C. F., Marsolek, C. J., and Koenig, O. (1995). On computational evidence for different types of spatial relations encoding: Reply to Cook et al. Journal of Experimental Psychology: Human Perception and Performance, 21, 423-431.

Kosslyn, S. M., Thompson, W. L., Gitelman, D. R., and Alpert, N. M. (1998). Neural systems that encode categorical versus coordinate spatial relations: PET investigations. Psychobiology, 26 (4), 333-347.

Krane D E, et al., (2008) Sequential unmasking: a means of minimizing observer effects in forensic DNA interpretation, J. Forensic Sci. 56 1006.

Krippendorff K. (1980) Content analysis: an introduction to its methodology. Content analysis: an introduction to its methodology. Beverley Hills, CA, USA, Sage.

Kruglanski A. (1980) Lay epistemic logic, process and contents: another look at attribution theory. Psychological Review 87[1], 70-87.

Kruglanski A and Freund T. (1983) The freezing and unfreezing of lay-inferences: Effects on impressionai primacy, ethnic stereotyping, and numerical anchoring. Journal of Experimental Social Psychology, 19, 448-68.

Kruglanski A and Ajzen I. (1983) Bias and error in human judgment. European Journal of Social Psychology 13, 1-44.

Kruglanski A and Mayseless O. (1987) Motivational effects in the social comparison of opinions. Journal of Personality and Social Psychology 53[5], 834-42.

Kruglanski A. (1989) Lay epistemics and human knowledge: Cognitive and motivational bases. New York, Plenum Press.

Kruglanski A. (1989) The psychology of being right: the problem of accuracy in social perception and cognition. Psychological Bulletin 106[3], 395-409.

Kruglanski A, Webster D M, and Klem A. (1993) Motivated resistance and openness to persuasion in the presence or absence of prior information. Journal of Personality and Social Psychology 65, 861-76.

Kruglanski A and Webster D M. (1996) Motivated Closing of the Mind: Seizing and Freezing. 2. Psychological Review 103, 263-83.

Kruglanski A, Shah J, Pierro A, and Mannetti L. (2002) When similarity breeds content: Need for closure and the allure of homogeneous and self resembling groups. Journal of Personality and Social Psychology 83[3], 648-62.

Kruglanski A, Pierro A, Mannetti L, and de Grada E. (2006) Groups as epistemic providers: need for closure and the unfolding of group-centrism. Psychological Review 113[1], 84-100.

Kundel, H.L. and Nodine, C. F. (1983). A visual concept shapes image perception. Radiology, 146, 363–368.

Laeng, B. (1994). Lateralization of categorical and coordinate spatial functions: A study of unilateral stroke patients. Journal of Cognitive Neuroscience, 6, 189-203

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (1997). International Affective Picture System (IAPS): technical manual and affective ratings. Gainesville, FL: University of Florida, NIMH Center for the Study of Emotion and Attention.

Langenburg G (2004) Pilot study: a statistical analysis of the ACE-V methodology – analysis stage, J. Forensic Ident. 54 (1) 64–79.

Langenburg G, Champod C, Wertheim P (2009) Testing for potential contextual bias effects during the verification stage of the ace-v methodology when conducting fingerprint comparisons, J. Forensic Sci. 54 (3) 571–582.

Langenburg G (2009) A method performance pilot study: testing the accuracy, precision, repeatability, reproducibility, and biasability of the ACE-V process, J. Forensic Ident. 59 (2) 219–257.

Leadbetter M. (1984) Letter to Editor. Fingerprint Whorld 9[35], 64.

Leadbetter M. (2007) Letter to Editor. Fingerprint Whorld 33[129], 231.

Leveson Rt Hon Lord Justice (2010). Expert Evidence in Criminal Courts – The Problem. Keynote Speech Forensic Science Society, Kings College, University of London.

Levy, B., Ashman, O., & Dror, I. E. (2000). To be or not to be: the effects of age stereotypes on the will to live. Omega: Journal of Death and Dying, 40(3), 409–420.

Lewin K. (1947) Frontiers in group dynamics/Resolving Social Conflicts: Selected Papers on Group Dynamics. 5-41. New York, US, Harper & Row.

Leyden J. (2002) Gummi bears defeat fingerprint sensors. Security Magazine 23rd May.

Likert R. (1932) A Technique for the Measurement of Attitudes. Archives of Psychology 140, 1-55.

Linden, K. W. and Linden, J. D. (1968). Modern Mental Measurement: A Historical Perspective. Boston: Houghton Mifflin.

Liversedge, S. P., & Findlay, J. M. (2000). Saccadic eye movements and cognition. Trends in Cognitive Sciences, 4, 6-14.

Livi S, Mannetti L, Pierro A, Kruglanski A, and Kenny D A. (2006) Epistemic motivation and groups' reaction to change: Effects of need for cognitive closure on norms transissions in a laboratory micro culture.

Lu, Z., Lesmes L. A., and Dosher, B. A. (2002). Spatial attention excludes noise at the target location. Journal of Vision, 2, 312-323.

Lubin, A. (1961) "The Interpretation of Significant Interaction", Educational and Psychological Measurement, Vol.21, No.4, Winter 1961, 807–817.

Mace C A. (1932) The Psychology of Study. Middlesex, UK, Penguin.

Maceo A V. (2005) The basis for the uniqueness and persistence of scars in the friction ridge skin. Fingerprint Whorld.

Madden D. (2008) Aging and Visual Attention. Current directions in psychological science. 16 (2), 70-74.

Maltoni, D., Maio, D., Jain, A. K., & Prabhakar, S. (2003). Handbook of fingerprint recognition. New York: Springer.

Marr, D. (1982). Vision: A computational investigation into the human representation and processing of visual information. New York: W. H. Freeman.

Marteniuk, R.G. (1976). Information Processing in Motor Skills. New York: Holt, Rinehart, & Winston.

Martin, D. W. (1996). Doing psychology experiments. (4th ed). Pacific Grove, CA: Brooks/Cole.

Martin, J.A., Regehr, G., Reznick. R, MacRae, H., Murnaghan, J., abd Hutchison C. et al. (1997). Objective structured assessment of technical skill (OSATS) for surgical residents. British Journal of Surgery, 85, 273-278.

Mayo K. (2006) Fingerprints and the Human Mind. Evidence Technology Magazine , 12-40.

Mayseless O and Kruglanski A. (1987)What makes you so sure? Effects of epistemic motivations on judgmental confidence. Organisational behaviour and human decision processes 39, 162-83.

McCartney C. (2006) Forensic Identification and Criminal Justice: Forensic Science, Justice and Risk. Devon, UK, Willan Publishing.

McCartney C. (2007) Nuffield Council on Bioethics. The forensic use of bioinformation: Ethical issues. Nuffield Council on Bioethics , 15. 2007. Nuffield Council on Bioethics.

McClusky, D.A., Ritter, E.M., Lederman, A.B., Gallagher, A.G., and Smith, C.D. (2005). Correlation between perceptual, visuo-spatial, and psychomotor aptitude to duration of training required to reach performance goals on the MIST-VR surgical simulator. American Surgery, 71, 13–20.

McGuire, F. (1994). Army alpha and beta tests of intelligence. In R. J. Sternberg (Ed.), Encyclopedia of human intelligence (pp. 125-129.) New York: Macmillan

McKay R, Langdon R, and Coltheart M. (2006) Need for Closure, Jumping to Conclusions, and Decisiveness in Delusion-Prone Individuals. The Journal of Nervous and Mental Diseases 194[6], 422-26. Lippincott Williams and Wilkins.

McKie I and Russell M. (2007) The Price of Innocence. Edinburgh, Scotland, Birlinn Ltd.

McNichol, D. (1972) A Primer of Signal Detection Theory. London: George Allen & Unwin

Meadmore, K.L., Dror, I.E., Bucks. R.S., and Liversedge, S.P. (2011). Eye movements during visuospatial judgments. European Journal of Cognitive Psychology.

Menchaca-Brandan, A., Liu, A.M., Oman, C.M., and Natapoff A. (2007). Influence of perspective-taking and mental rotation abilities in space teleoperation. Proceedings of the ACM/IEEE international conference on Human-robot interaction, 271-278.

Merlo, L.J. and Matveevskii, A.S. (2009). Personality testing may improve resident selection in anaesthesiology programs. Medical Teach, 31 (12), e551-e554.

Mesulam, M.-M. (1981). A cortical network for directed attention and unilateral neglect. Annals of Neurology, 10, 309-325.

Millard, K. (1975) In An Approach to the Automatic Retrieval of Latent Fingerprints, Proceedings of Carnahan Conference on Electronic Crime Countermeasures, Lexington, KY; pp 45–51.

Millard, K. (1983) In Development on Automatic Fingerprint Recognition, Proceedings of the Carnahan Conference on Security Technology, Zurich, Switzerland; pp 173–178.

Miller, R. B. (1962). Task descriptions and analysis. In R. M. Gagne (Ed.), Psychological Principles in System Development (pp. 353-380). New York: Hold, Rinehart and Winston.

Mikulincer M. (1997) Adult attachment style and information processing: Individual differences in curiosity and cognitive closure. Journal of Personality and Social Psychology 72[5], 1217-30.

Moenssens, A. A. (2003). Fingerprint identification: A reliable ‘forensic science’? Criminal Justice ABA, 18(2), 31–37.

Moennsens A and Meagher S.(2008) Fingerprints and the Law. Chapter 15.

Moore, R. T. (1991) Automatic Fingerprint Identification Systems. In Advances in Fingerprint Technology, 1st ed.; Lee, H. C.; Gaensslen, R. E., Eds.; Elsevier, NY; pp 163–191.

Moran D A and Gross S R. (2008) Put scientists, not cops, in crime labs. University of Michigan Law School.

Murphy, S. T., & Zajonc, R. B. (1993). Affect, cognition, and awareness: affective priming with optimal and suboptimal stimulus exposures. Journal of Personality and Social Psychology, 64, 723–739.

Myles-Worsley, M., Johnston, W. A., & Simons, M. A. (1988). The influence of expertise on X-ray image processing. Journal of Experimental Psychology: Learning, Memory, and Cognition, 14(3), 553-557.

National Academia of Sciences. (2008) Identifying the Needs of the Forensic Sciences Community.

National Academy of Sciences (2009) Strengthening Forensic Science in the United States: A Path Forward. National Academy of Sciences, Washington, DC.

Neisser, U., Boodoo, G., Bouchard, T. J., Boykin, A. W., Brody, N., Ceci, S. J.,
Nettle D. (2005) Happiness. New York, US, Oxford University Press.

Nickerson Raymond S. (1998) Confirmation Bias: A Ubiquitous Phenomenon in Many Guises. Review of General Psychology 2[2], 175-220.

Niedenthal, P. M., Halberstadt, J. B., & Setterlund, M. B. (1997). Being happy and seeing 'happy'. Emotional state mediates visual word recognition. Cognition and Emotion, 11(4), 403–432.

Niedenthal, P. M., Halberstadt, J. B., Margolin, J., & Innes-Ker, A. H. (2000). Emotional state and the detection of change in facial expression of emotion. European Journal of Social Psychology, 30, 211–222.

Nobre, A. C., Sebestyen, G. N., Gitelman, D. R., Frith, C. D., and Mesulam, M. M. (2002). Filtering of distractors during visual search studied by positron emission tomography. Neuro Image, 16, 968–976.

Office of the Inspector General. (2006) A Review of the FBI's Handling of the Brandon Mayfield Case. US Dept of Justice.

O'Neill E (2006) Mark of Innocence, the Guardian Newspaper (P6-9) 18 April.

Osborn, A. S. (1939). Form Blindness and Proof: Sight Defects in Relation to the Administration of Justice. Journal of Criminal Law and Criminology, 30 (2), 243-249.

Pankanti S, Prabakhar S, and Jain A. (2001) On the Individuality of Fingerprints. Proceedings of the IEEE Computer Society Conferenceon Computer Vision and Pattern Recognition (CVPR) , 805-12.

Payne B. (2006) Weapon Bias Split-Second Decisions and Unintended Stereotyping. Current directions in psychological science. 15[6], 287-91.

Payne K B, Shimizu Y, and Jacoby L L. (2004) Mental control and visual illusions: Toward explaining race-biased weapon misidentifications. Journal of experimental Social Psychology 41, 36-47.

Penrose L S and Plomley N J B.(1980) Structure of Interstitial Epidermal Ridges. Fingerprint Whorld 5[20], 104-06.

People v. Jennings (1911), 96 NE 1077.

Petersen JL, Moore J. (1996) The status of AFIS systems worldwide: Issues of organization, performance and impact. In: Almog J, Spinger E, editors. Proceedings of the International Symposium on Fingerprint Detection and Identification.

Petersilia, J. (1975) The Collection and Processing of Physical Evi- dence; WN-9062-DOJ; RAND Corporation: Santa Monica, CA.

Phillips V L, M.J. Saks, J.L. Peterson, (2001) The application of signal detection theory to decision-making in forensic science, J. Forensic Sci. 46 (2) 294–308.

Pierce D S Mechanics of Impression Evidence. (2011) Auerbach Publications. Pages 133–140. Print ISBN: 978-1-4398-1370-6, DOI: 10.1201/b10612-9.

Pincus, T., Pearce, S., & Perrott, A. (1996). Pain patients' bias in the interpretation of ambiguous homophones. British Journal of Medical Psychology, 69, 259–266.

Posner, M. I., Snyder, C.R.R., and Davidson, D.J. (1980). Attention and the detection of signals. Journal of Experimental Psychology: General, 109, 160-174.

Posner, M.I. and Petersen, S.E. (1990). The attention system of the human brain. Annual review of neuroscience, 13, 25-42

Posner M.I. (Ed) (2004). Cognitive Neuroscience of Attention. New York: Guilford.

Potchen E J, T.G. Cooper, A.E. Sierra, G.R. Aben, M.J. Potchen, M.G. Potter, J.E. Siebert (2000) Measuring performance in chest radiography, Radiology 217, 456– 459.

Powis, D. (2009). Personality testing in the context of selecting health professionals. Medical Teacher, 31, 1045–1046.

Pribyl, J. R. and Bodner, G. M. (1987). Spatial ability and its role in organic chemistry: a study of four organic courses. Journal of Research in Science Teaching, 24, 229.

Pronin E. (2006) Perception and misperception of bias in human judgment. Trends in cognitive sciences 11[1].

Pylyshyn, Z. W. (1984). Computation and Cognition: Towards a foundation for cognitive science. Cambridge, MA: MIT Press.

Raghunathan R and Pham M T. (1999) All negative moods are not equal: Motivational Influences of anxiety and sadness on decision making. Organisational behaviour and human decision processes 79[1], 56-77.

N. Ratha and R. Bolle (Eds.) (2003) Automatic Fingerprint Recognition Systems, ISBN: 0387955933, Springer.

Haber L and Haber N. (2004) Error Rates for Human Latent Fingerprint Examiners. Chapter 17. New York, US, Springer-Verlag.

Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. Psychological Bulletin, 124, 372-422.

Richards, A., Reynolds, A., & French, C. C. (1993). Anxiety and the spelling and use in sentences of threat/neutral homophones. Current Psychology: Developmental, Learning, Personality, Social, 12, 18–25.

Richards, A., French, C. C., Calder, A. J., Webb, B., Fox, R., & Young, A. W. (2002). Anxiety-related bias in the classification of emotionally ambiguous facial expressions. Emotion, 2, 273–287.

Risinger D M, M.J. Saks, W.C. Thompson, R. Rosenthal (2002) The Daubert/Kumho implications of observer effects in forensic science: hidden problems of expectation and suggestion, Calif. Law Rev. 90 (1)1–56.

Robertson, L. C. and Delis, D. C. (1986). “Part–whole” processing in unilateral brain damaged patients: Dysfunction of hierarchical organization. Neuropsychologia, 24, 363-370.

Rogers, Y. and Ellis, J. (1994) 'Distributed Cognition: an alternative framework for analyzing and explaining collaborative working'. Journal of Information Technology, 9 (2), 119-128.

Rose. A (2007) Lethal Witness: Sir Bernard Spilsbury, the Honorary Pathologist. The History Press Ltd; 07 edition.

Rosenthal R (1978) How often are our numbers wrong? Am. Psychol. 33 (11) 1005 -1008.

Rosenthal R, R.L. Rosnow (2007) Essentials of Behavioral Research: Methods and Data Analysis, third ed., McGraw-Hill Press.

Rudin, N. Inman K. (2005) Fingerprints in Print, the Sequel.The CAC News, 2nd Quarter, P 6 -10.

Rueckl, J.G., Cave, K.R., and Kosslyn, S.M. (1989). Why are “what” and “where” processed by separate cortical visual systems? A computational investigation. Journal of Cognitive Neuroscience, 1, 171-186.

Ruff, R., Light, R., and Parker, S. (1996). Visuospatial learning: Ruff light trail learning test. Archives of Clinical Neuropsychology, 11 (4), 313-327.

Rullen, R., Reddy, L., and Koch, C. (2004) Visual search and dual tasks reveal two distinct attentional resources. Journal of Cognitive Neuroscience, 16, 4-14.

Russo, R., Patterson, N., Roberson, D., Stevenson, N., & Upward, J. (1996). Emotional value of information and its relevance in the interpretation of homophones in anxiety. Cognition and Emotion, 10, 213–220.

S. and Marper v. The United Kingdom (December 4, 2008) (Application nos. 30562/04 and 30566/04).

Saks M, Risinger D M, Rosenthal R, and Thompson W C. (2003) Context effects in forensic science: A review and application of the science of science to crime laboratory practice in the US. Science and Justice 43[2], 77-90.

Saks M and Koehler J. (2005) The coming paradigm shift in forensic identification. Science 309, 825-92.

Saks M and Koehler J. (2008) The Individualization Fallacy in Forensic Science Evidence. Vanderbilt Law Review 61[199], 199-219.

Saks M, L.J. (2009) Concerning, E. Hall, Player ‘Will the introduction of an emotional context affect fingerprint analysis and decision-making?’ Forensic Sci. Intern. 191 e19.

Samischenko S S. (2001) Atlas of the unusual papilla patterns. Russia, Mockba.

Sanders J. (2007) Expert Witness Ethics. Fordham Law review 76, 1539-84.

Schiffer B, C. Champod (2007) The potential (negative) influence of observational biases at the analysis stage of finger mark individualization, Forensic Sci. Intern. 167, 116–120.

Schiffer B (2009) The relationship between forensic science and judicial error: a study covering error sources bias and remedies, Ph.D. Thesis, Universite ´ de Lausanne.

Schneider-Kolsky, M., Wright, C., and Baird, M (2006). Evaluation of selection criteria for graduate students in radiation therapy. Medical Teacher, 28 (8), e214–e219.

Schneier B. Beyond Fear. (2003). New York, US, Copernicus Books.

Schwaninger A, Hardmeier D, Hofer F. (2004) Measuring visual abilities and visual knowledge of aviation security screeners. IEEE ICCST Proceedings

Scott H. (1954) Scotland Yard. The Forensic Science Laboratory. New York, US, Random House.

Selfridge, O. G. (1955). Pattern recognition and modern computers. Proceedings of the Western Joint Computer Conference. New York: Institute of Electrical and Electronics Engineers.

Sengoopta C. (2003) Imprint of the Raj: How fingerprinting was born in colonial India. London, MacMillan.

Sergent, J., Zuck, E., Lévesque, M. and MacDonald, B. (1992). Positron emission tomography study of letter and object processing: Empirical findings and methodological considerations. Cerebral Cortex, 2, 68-80.

Sergent, J., Ohta, S., and MacDonald, B. (1992). Functional neuroanatomy of face and object processing: A positron emission tomography study. Brain, 115, 15-36.

Shah J, Kruglanski A, and Thompson E. (1998) Membership has its (epistemic) rewards: Need for closure effects on in-group bias. Journal of Personality and Social Psychology 75[2], 383-93.

Shepard, R. N. and Metzler, J. (1971). Mental rotation of three-dimensional objects. Science, 171, 701-703.

Shepard R. N and Feng C (1972). A chronometric study of mental paper folding. Cognitive Psychology, 3 (2), 228-243.

Shepard, R. N. and Cooper, L. A. (1982). Mental Images and Their Transformation. Cambridge, MA: MIT Press.

Simon H. A & Gilmartin, K. J. (1973). A simulation of memory for chess positions. Cognitive Psychology, 5, 29-46.

Sodhi G S, Kaur J, and Colony J. (2002) On Henry's Fingerprint Classification System. Fingerprint Whorld 28[110], 197-206. 2002.

Solomon, J.A. and Pelli, D. G (1994). The visual filter mediating letter identification. Nature 369, 395 - 397.

Sommers S R and Douglass A B. (2007) Context Matters: Alibi strength varies according to evaluator perspective. Legal and Criminal Psychology 12, 41-54.

Sperling, G. (1960). The information in brief visual presentations. Psychological Monographs, 74.

Stacey R B. (2004) Report on the Erroneous Fingerprint Individualization in the Madrid train Bombing Case. Journal of Forensic Identification 54[6], 706-18.

Starkes, J.L., & Allard, F (1983). Perception in Volleyball: The effects of competitive stress. Journal of Sport Psychology, 5(2), 189-196.

Starkes, J.L. (1987). Skill in field hockey: The nature of the cognitive advantage. Journal of Sport Psychology, 9, 146-160.

State of Maryland v. Bryan Rose (2007) in the Circuit Court for Baltimore County Case No. K06-545.

Staw B M: (2004) "Knee-deep in the Big Muddy: A Study of Escalating Commitment to a Chosen Course of Action". Organizational Behavior and Human Performance 16(1):27-44

Steele L. (2004) The Defense Challenge to Fingerprints. Criminal Law Bulletin 40[3].

Stibel J M, Dror I E, and Ben-Zeev T. (2008) The Collapsing Choice Theory: Dissociating Choice and Judgment in Decision Making. Theory and Decision.

Stocker T. (2007) Forensic Training Handbook. Sussex Police. 1st Edition.

Street, D.R., Chapman, A.E., and Helton, K.T. (1993). The future of naval aviation selection:

Broad-spectrum computer-based testing. In Proceedings of the 35th Annual Military Testing Association. Williamsburg, VA.
Summerfield C T. Egner (2009) Expectation (and attention) in visual cognition, Trends Cogn. Sci. 13 (9) 403–409.

Swets, J.A (1964) Signal detection and recognition by human observers. New York: Wiley

Tansley, P., Kakar, S., Withey, S., and Butler, P. (2007). Visuospatial and technical ability in the selection and assessment of higher surgical trainees in the London deanery. Annual Royal College of Surgery England, 89, 591–595.

Thompson W C (1995) Subjective interpretation, laboratory error and the value of DNA evidence: three case studies, Genetica 96,153–168.

Thompson W and Cole S. (2005) Lessons from the Brandon Mayfield case. The Champion 29[3], 42-44.

Thompson W C (2009) Interpretation: Observer effects. A. Moenssens, A. Jamieson (Eds.), Encyclopaedia of Forensic Sciences, John Wiley & Sons, London, pp. 1575–1579.

Tiller C. (1986) Stop and think about errors. Fingerprint Whorld 11[43], 75-76.

Triplett M and Cooney L. (2006) The etiology of ACE-V and its Proper Use: An exploration of the relationship between ACE-V and the scientific method of hypothesis testing. Journal of Forensic Identification 56[3], 345-53.

Tuckman, B. (1965) Developmental sequence in small groups. Psychological bulletin, 63, 384-399

Tversky A and Kahneman D (1974) Judgment under uncertainty: Heuristics and biases. Science 185, 1124-31.

Valk J.P.J and Eijkman E.G.J (1984). Analysis of eye fixations during the diagnostic interpretation of chest radiographs. Medical and Biological Engineering and Computing, 22, 353–360.

Van Hiel A and Mervielde I. (2002) Effects of ambiguity and need for closure on the acquisition of information. Social Cognition 20[5], 380-408.

Vernon M D. (1962) The Psychology of Perception: Chapter 11: The relation to perception of motivation and emotion. Middlesex, London, Penquin.

Vincent P. (1985) Skin, a brief look under the surface. Fingerprint Whorld 11[41], 9-12.

Vincent P. (1985) Skin.Pt 2. Fingerprint Whorld 11[42], 28-33.

Vokey J R, Tangen J M, and Cole S A. (2008) On the preliminary psychophysics of fingerprint identification. Journal of Experimental Psychology 0000[0], 1-18.

Walker R. (1985)Applied Qualitative Research. Aldershot, UK, Glower.

Warren, R. M. (1970). Perceptual restorations of missing speech sounds. Science, 167, 392–393.

Waywell, L. and Bogg, J. (1999). Spatial ability assessment: an aid to student selection for therapy radiography training. Journal of Radiotherapy in Practice, 1 (2), 89-95.

Weber P R. (1985) Basic Content Analysis. Beverley Hills, CA, USA, Sage.

Webster D M and Kruglanski A W. (1994) Individual Differences in Need for Cognitive Closure. Journal of Personality and Social Psychology Bulletin 67[6], 1049-62.

Webster D M, Kruglanski A W, and Pattison D.(1997) Motivated language use in intergroup contexts: need for closure effects on the linguistic intergroup bias. Journal of Personality and Social Psychology 72[5], 1122-31.

Wei L, Li X, and Soman D. (2005) Getting it Off the Desk: The Motivation to Attain Closure on Tasks. University of Toronto.

Wertheim K, G. Langenburg, A. Moenssens. (2006) A report of latent print examiner accuracy during comparison training exercises, J. Forensic Ident. 56 (1) 55–93.

Westman, B., Ritter, M., Kjellin, A., Torkuist, L., Wredmark, T., Fellander, L., and Enochsson, L. (2006).Visuospatial abilities correlate with performance of senior endoscopy specialists in simulated colonoscopy. Journal of Gastrointestinal Surgery, 10, 593–599.

Wickens, Thomas D., (2002) Elementary Signal Detection Theory. New York: Oxford University Press. (ISBN 0-19-509250-3)

Wilkinson S, Joffe H, Yardley L(Ed), and Marks D. (2004) Qualitative data collection. Research Methods for Clinical and Health Psychology . London, UK, Sage.

Wolfe, J.M. (1994). Guided Search 2.0: A revised model of visual search. Psychonomic Bulletin and Review, 1 (2), 202-238.

Wood, B. P. (1999). Visual Expertise. Radiology, 211, 1–3.

Wright, R.D and Ward, L.M. (2008). Orienting of Attention. Oxford University Press.

Yang, H., & Zelinsky, G.J. (2009). Visual search is guided to categorically-defined targets. Vision Research, 49, 2095-2103.

Yerkes, R. M. (1919). Report of the psychology committee of the national research council. The Psychological Review, 26 (2), 83-149.

Yerkes, R. M. (1921). Psychological examining in the United States Army. Memoirs of the National Academy of Sciences, Vol. XV. Washington, DC: U.S. Government Printing Office.

Yue, J (2007). Spatial visualization by isometric view. Engineering Design Graphics Journal, 71(2), 5-19.

Zabell S. (2005) Fingerprint Evidence. Journal of Law and Policy , 143-79.

Zamvar, V (2004). No time to train the surgeons: choosing trainees carefully may be the solution. British Medical Journal, 328, 1134–1125.

Zeelenburg A. The McKie case in short. http://www.clpex.com/McKie.htm Accessed 2008.

Published Book Chapters, Papers and Presentations

Subject to Peer Review Contributing to the Writing of this Thesis

Published Book Chapters

Charlton D, Galloway V (2006), Chapter…. Identification of latent fingerprints, Text Book: Forensic Human Identification, An Introduction. BAHID, CRC Press.

Charlton D (2008), Chapter…..Finding latent prints at crime scenes, Text Book: Crime Scene Management…Wiley Publishing Group.

Published Papers:

Charlton D. (2006), Forensic Expertise and Self Trust, Forensic Update (Division of Forensic Psychology), the British Psychological Society, Number 87 P14-19.

Charlton, D., Del Manso, H., and Dror, I.E. (2007). Expert error: The mind trap. Fingerprint Whorld, 33, 151-155.

Charlton D. (2007) Group Networks and Dynamics Associated with Forensic Fingerprint Analysis and Decision-Making. Canadian Identification Society. Identification Canada 30, (4) 126-129.

Charlton, D., Fraser-Mackenzie, P. and Dror, I. E. (2010). Emotional Experiences and Motivating Factors Associated with Fingerprint Analysis. Journal of Forensics Sciences, 55 (3).

Dror, I.E., Peron, A., Hind, S., & Charlton, D. (2005). When emotions get the better of us: The effect of contextual top-down processing on matching fingerprints. Applied Cognitive Psychology, 19 (6), 799-809.

Dror.I.E, Charlton.D (2006) Why Experts Make Errors
Journal of Forensic Identification, 56, (4), 600-616

Dror, I.E., Charlton, D., & Peron A. (2006). Contextual information renders experts vulnerable to making erroneous identifications. Forensic Science International, 156 (1), 74-78.

Dror, I.E. & Charlton, D. (2007), "Improving perception and judgment: 'An examination of expert performance" Fingerprint Whorld Vol 33 No 129

Dror, I.E. & Charlton, D., 2007, "Methodological and conceptual problems in measuring and enhancing expert performance". Fingerprint Whorld Vol 33 No 129

Dror, I. E., Champod, C., Langenburg, G., Charlton, D., Hunt, H., & Rosenthal R. (2011). Cognitive issues in fingerprint analysis: Inter-and intra-expert consistency and the effect of a 'target' comparison. Forensic Science International, 208, 10-17.

Dror I. E., Charlton. D, Fraser-Mackenzie. P, Bucht. R, Taylor. M, Rosenthal. R and Boehm Davis. D (In Submission NIJ) The Industrial Revolution in Forensic Decision Making: Scientifically Based Management and Standards.

Fraser-Mackenzie P, Dror I E, Charlton D, Sung M, Johnson J E V (Submitted) Bounded Quantitative Judgments: Cognitive Distortions in the Evaluation of Forensic Evidence, Investment Performance and Alternative Gambles. Organisational Behaviour and Human Decision Processes. Elsevier.

Published Articles

Charlton D. (2001) Comment. Fingerprint Whorld 27 (106), 133-135.

Charlton D. (2002) Animal Farm. Fingerprint Whorld 28 (107), 5-6.

Charlton D (2002) Kobayashi Maru. Fingerprint Whorld 28 (108), 65-67.

Charlton D. (2002) The Bunker Mentaility. Fingerprint Whorld 28 (110), 193-194.

Charlton D. (2002) Report Into the Chicago Fingerprint Forum (Post Daubert). Fingerprint Whorld 28 (110), 215-219.

Charlton D. (2003) Comment. Fingerprint Whorld 29 (112), 65.

Charlton D. (2003) To Swim the English Channel Only to Drown in the Harbour. Fingerprint Whorld 29 (113),125-126.

Charlton D. (2003) Closing the Criminal Justice Gap. Fingerprint Whorld 29 (114), 201.

Charlton D. (2004) Counting the Cost. Fingerprint Whorld 30 (115), 5.

Charlton D. (2004) 2000 Years of History in a Soundbite!! Fingerprint Whorld 30 (115) 6-8.

Charlton D. (2004) A Biometric Future. Fingerprint Whorld 30 (116), 41.

Charlton D. (2004) Keeping it Simple. Fingerprint Whorld 30 (117), 101.

Charlton D. (2004) Bete Noire!! Fingerprint Whorld 30 (118), 161-162.

Charlton D. (2005) The Greenhouse Effect. Fingerprint Whorld 31 (119), 5-6.

Charlton D. (2005) Devolution or Dissolution? Fingerprint Whorld 31 (120), 73-74.

Charlton D. (2005) The Forensic Family. Fingerprint Whorld 31 (121), 129-130.

Charlton D. (2005) Evolution or Revolution? Fingerprint Whorld 31 (122), 205.

Dror, I E, Charlton D. (2006): The Psychology of Police Performance and Decision Making: Police Professional August 2006, P37-39.

Charlton D. (2006) Two for the Price of One. Fingerprint Whorld 32 (123), 5.

Charlton D. (2006) Six of One. Fingerprint Whorld 32 (125), 145.

Dror I E, Charlton D, (2007) Why are Experts Prone to Error? Centre of Visual Cognition Technical Report, University of Southampton, UK.

Presentations:

Charlton D, (2002) Ethics in Forensics (Summing up Speech), International Association for Identification Conference, Miami, Florida.

Charlton D. (2003) Terrorism through the ages. Rocky Mountain Division of the IAI, Breckenridge, Colorado 2003:

Charlton D, (2003) Ethics and morals in biometric science. Association for Biometrics AGM, London.

Charlton, D., Peron, A.E., & Dror, I. E., (2004). The interplay of perceptual and cognitive elements in fingerprint identification: When higher-level cognition can facilitate or hinder fingerprint matching. International Biometric Society, British Region Annual Meeting. Royal Statistical Society, London.

Charlton D, (2004). The Reliability of Fingerprint Evidence, The Royal Society, Edinburgh.

Charlton D, The History of Fingerprints (2004), NEC Corporate event. The launch of the new laptop fingerprints reader, London.

Charlton D, (2005), Anatomy of a Tsunami and the DVI response. Cogent User Group Conference, Washington DC, US.

Charlton D, (2005). The history of Fingerprint Science and its effect on the future of biometrics. Royal Society Medicine, London.

Charlton D, (2006), Logistical and technical issues when dealing with mass disaster. International Symposium on Advances in Fingerprint Technology. University of Delhi, India.

Charlton D, (2006), McKie to Mayfield…a journey to extinction for latent print examination as a reliable science? Or an opportunity to reinforce its validity? International Symposium on Fingerprint science, De-Paul University of Law, Chicago, USA.

Charlton D, (2007), Is the Number Up for Fingerprint Science? (The Motivations of a Fingerprint Examiner), CIS Educational Conference Montreal, Canada.

Charlton D, (2008), The 'Wahey' Moment (The Motivations of Forensic Examiners), the 2008 National Laboratories Conference, Manchester, UK.

Charlton D, (2009), How to Lose Friends and Influence People: The Motivations of a Fingerprint Examiner and the Importance of Speaking Your Mind, Fingerprint Society Lectures, Manchester, UK.

Charlton D, (2009), The Motivations of a Fingerprint Examiner and the Importance of Speaking Your Mind, Scientific Support Manager's Conference, Tewksbury, UK.

Charlton D, Dror I E, (2010), Competency of Fingerprint Examiners: How Do We Assess Competence? First We Have To Know We Are Asking The Right Questions! Interpol 6th International Symposium, Lyon.

Charlton D, (2010), The Standards of Expediency, Presentation as part of a panel discussion on forensic standards. Fingerprint Society Lectures, University College London, UK.

Charlton D, (2010), Trust Me, I'm an Expert, Presentation to MSc Forensic Science Group, Bournemouth University, UK.

Charlton D, (2010), Beware Sherlock Holmes, Seeing is not always believing, Presentation to MSc Forensic Science Group, Bournemouth University, UK.

Charlton D, (2011), The Asian Tsunami, Presentation to MSc Forensic Science Group, Bournemouth University, UK.

Charlton D, (2011), Forensic Science From Clay to Silicon: 2000 Years of Biometric Identification, Presentation to MSc Forensic Science Group, Bournemouth University, UK.

Charlton D, (2011), Fingerprint Age Determination: Is there any hope of reaching this Holy Grail? Presentation to MSc Forensic Science Group, Bournemouth University, UK.

Charlton D, (2011), An Industrial Revolution for Fingerprint Science?
The Impact of Cognition and Human Factors on Fingerprint Examiners: Implications for the Use of Fingerprint Examiner Expertise and Administration within Law Enforcement. Presentation to Postgraduate Research Conference, Bournemouth University, UK.

Charlton D, (invited to speak October 2011), Cognitive Technology: A Theory of Phased Cognitive Engagement in the Fingerprint Domain. Cogent User Conference, Connecticut, US.

Dror, I. E., Charlton, D., & Peron, A.E. (2004). Evaluating ‘scientific’ evidence for the court: What contributing factors are really involved in fingerprint identification? International Centre for Advanced Research in Identification Science (ICARIS) Sheffield.

Dror, I. E., Peron, A.E., & Charlton, D. (2004). Psychological factors involved in fingerprint identification, Autumn Conference of the Forensic Science Society, Bedford, UK.

Dror, I. E., Charlton, D. (2005). The vulnerability of fingerprint science and potential pitfalls in the identification process. How can they be addressed and overcome? Fingerprint Society Educational Conference, Brighton, UK.

Peron, A.E., Dror, I. E., Hind, S.L, & Charlton, D. (2004). Decision making processes involved in fingerprint identification: The influence of emotional context on finding a match, 14th International Forensic Science Symposium, Granada, Spain.