COLLABORATIVE PRACTISE BETWEEN UNITED KINGDOM FINGERPRINT BUREAUX AND ACADEMIA TO ESTABLISH EFFECTIVE DOCUMENTATION STRATEGIES FOR FINGERPRINT EXAMINATIONS

Megan Needham

MSci, PgCHPE

A thesis submitted in partial fulfilment of the requirement of Staffordshire University for the degree of Doctor of Philosophy

Contents

[Acknowledgements v](#_Toc146200958)

[Abstract vi](#_Toc146200959)

[Glossary viii](#_Toc146200960)

[List of Figures x](#_Toc146200961)

[List of Tables xi](#_Toc146200962)

[Chapter 1: Introduction 1](#_Toc146200963)

[1.1 - Introduction to Friction Ridge Detail 2](#_Toc146200964)

[1.2 - Introduction to Fingerprint Examinations (ACE-V process) 4](#_Toc146200965)

[1.2.1 – Analysis stage of ACE-V process 4](#_Toc146200966)

[1.2.2 - Comparison stage of ACE-V process 5](#_Toc146200967)

[1.2.3 - Evaluation stage of ACE-V process 6](#_Toc146200968)

[1.2.4 - Verification stage of ACE-V process 7](#_Toc146200969)

[1.2.5 – The reliability of the ACE-V process 7](#_Toc146200970)

[1.3 – Literature Review 10](#_Toc146200971)

[1.3.1 – The documentation procedures in relation to Fingerprint Examinations 10](#_Toc146200972)

[1.3.2 - Streamlined Forensic Reporting: Legal requirement for all forensic evidence types. 12](#_Toc146200973)

[1.3.3 – Quality Standards for Fingerprint Examinations 14](#_Toc146200974)

[1.3.4 - Guidance for documentation of fingerprint examinations 20](#_Toc146200975)

[1.4 - Aims and objectives 34](#_Toc146200976)

[Chapter 2: Methodologies 36](#_Toc146200977)

[2.1 – Friction ridge detail selection for the purposes of this research 37](#_Toc146200978)

[2.1.1 – The creation of Staffordshire’s Ground Truth Database 38](#_Toc146200979)

[2.1.2 – A comparison between a systematic scoring system and an exemplar police grading system 38](#_Toc146200980)

[2.2 – The establishment of pre-accreditation documentation strategies for Fingerprint Analysis; study one 39](#_Toc146200981)

[2.2.1 - Design 39](#_Toc146200982)

[2.2.2 - Participants 41](#_Toc146200983)

[2.2.3 - Materials 41](#_Toc146200984)

[2.2.4 - Procedure 42](#_Toc146200985)

[2.3 – The establishment of post-accreditation documentation strategies for Fingerprint Examinations; study two 42](#_Toc146200986)

[2.3.1 - Design 42](#_Toc146200987)

[2.3.2 – Participants 43](#_Toc146200988)

[2.3.3 – Materials 43](#_Toc146200989)

[2.3.4 - Procedure 43](#_Toc146200990)

[2.4 – The determination of effective documentation strategies for Fingerprint Examinations 44](#_Toc146200991)

[2.4.1 - Design 44](#_Toc146200992)

[2.4.2 – Participants 44](#_Toc146200993)

[2.4.3 – Materials 44](#_Toc146200994)

[2.4.4 - Procedure 45](#_Toc146200995)

[2.5 – The assessment of inter and intra-examiner variation of reporting outcomes at Analysis and Comparison/Evaluation stage 45](#_Toc146200996)

[2.5.1 – Design 45](#_Toc146200997)

[2.5.2 – Participants 45](#_Toc146200998)

[2.5.3 - Materials 45](#_Toc146200999)

[2.5.4 - Procedure 46](#_Toc146201000)

[Chapter 3: Results and Discussion 46](#_Toc146201001)

[3.1 – The selected friction ridge impressions for study one and study two 46](#_Toc146201002)

[3.2 – Objective 1 to 4: The establishment of pre-accreditation documentation strategies for Fingerprint Analysis; study one 50](#_Toc146201003)

[3.2.1 – Consensus of reporting outcomes at Analysis stage 57](#_Toc146201004)

[3.2.2 – Initial observations and chi-square tests for association for “Generic information” recorded from the ACE-V checklist. 59](#_Toc146201005)

[3.2.3 – Initial observations and chi-square tests for association for “Factors affecting the mark” recorded from the ACE-V checklist 64](#_Toc146201006)

[3.2.4 – Initial observations and chi-square tests for association for “Level of detail and annotations” from the ACE-V checklist 72](#_Toc146201007)

[3.2.5 – Creation of a new documentation form “Mark Analysis Form” 79](#_Toc146201008)

[3.3 – Objective 5 to 7: The establishment of post-accreditation documentation strategies for Fingerprint Examinations; study two 84](#_Toc146201009)

[3.3.1 – An overview of a selection of UK post-accreditation documentation strategies 85](#_Toc146201010)

[3.3.2 – Consensus of reporting outcomes during ACE-V process 90](#_Toc146201011)

[3.3.3 – Initial observations and chi-square tests for association for “Generic information” from the ACE-V checklist 94](#_Toc146201012)

[3.3.4 – Initial observations and chi-square tests for association for “Factors affecting the mark” from the ACE-V checklist 98](#_Toc146201013)

[3.3.5 – Initial observations and chi-square tests for association for “Level of detail and annotations” from the ACE-V checklist 104](#_Toc146201014)

[3.3.6 – Initial observations and chi-square tests for association for “Comparison and evaluation notes” from the ACE-V checklist 109](#_Toc146201015)

[3.4 – Objective 8 to 11: The determination of effective documentation strategies for Fingerprint Examinations; study three 112](#_Toc146201016)

[3.4.1 – Exemplar post-accreditation documentation strategies for the effectiveness study 115](#_Toc146201017)

[3.4.2 – Initial observations and chi-square tests for associations between a correct or incorrect outcome and expertise group 128](#_Toc146201018)

[3.4.3 – Initial observations and chi-square tests for associations between a correct or incorrect outcome and documentation strategy 133](#_Toc146201019)

[3.4.4 – Initial observations and chi-square tests for associations between a correct or incorrect outcome and grade of the mark 137](#_Toc146201020)

[3.5 – Objective 12: The assessment of inter and intra-examiner variation of reporting outcomes at Analysis and Comparison/Evaluation stage 139](#_Toc146201021)

[3.5.1 - Analysis and Evaluation outcomes for Grade 1 impressions 142](#_Toc146201022)

[3.5.2 - Analysis and Evaluation outcomes for Grade 2 impressions 148](#_Toc146201023)

[3.5.3 - Analysis and Evaluation outcomes for Grade 3 impressions 151](#_Toc146201024)

[3.5.4 - Analysis and Evaluation outcomes for Grade 4 impressions 153](#_Toc146201025)

[Chapter 4: Conclusions 161](#_Toc146201026)

[4.1 – The establishment of pre-accreditation documentation strategies; study one 161](#_Toc146201027)

[4.2 – The establishment of post-accreditation documentation strategies; study two 161](#_Toc146201028)

[4.3 – The determination of effective documentation strategies for Fingerprint Examinations 162](#_Toc146201029)

[4.4 – The assessment of inter and intra-examiner variation of reporting outcomes at Analysis and Comparison/Evaluation stage 164](#_Toc146201030)

[4.5 – Further work suggestions based on the research findings 165](#_Toc146201031)

[Chapter 5: Key findings from the research and Recommendations for Practitioners 167](#_Toc146201032)

[5.1 – The establishment of pre accreditation documentation strategies; study one 167](#_Toc146201033)

[5.2 – Post accreditation documentation strategies; study two 170](#_Toc146201034)

[5.3 – The effectiveness of documentation strategies for Fingerprint Examinations 173](#_Toc146201035)

[5.4 – Inter and Intra-examiner variance for reporting outcomes 175](#_Toc146201036)

[Chapter 6: References 178](#_Toc146201037)

[Appendices 185](#_Toc146201038)

[Appendix 1 – Scores 6 to 15 fingermarks from Ground Truth Database 185](#_Toc146201039)

[Appendix 2 – Example of the pre-accreditation documentation strategy (analysis form) 187](#_Toc146201040)

[Appendix 3 – ACE-V checklist 188](#_Toc146201041)

[Appendix 4 – An example of a post-accreditation documentation – strategy 2 (tick sheet form) 189](#_Toc146201042)

[Appendix 5 – An example of a post-accreditation documentation – strategy 6 (tick sheet form) 189](#_Toc146201043)

[Appendix 6 – An example of a post-accreditation documentation – strategy 4 (tick sheet form) 189](#_Toc146201044)

[Appendix 7 – An example of a post-accreditation documentation – strategy 7 (free text / back of a photograph) 189](#_Toc146201045)

[Appendix 8 – An example of a post-accreditation documentation – strategy 1 (free text / back of a photograph) 190](#_Toc146201046)

[Appendix 9 – An example of a post-accreditation documentation – strategy 3 (free text / word document) 190](#_Toc146201047)

[Appendix 10 – An example of a post-accreditation – strategy 5 (free text / word document) 190](#_Toc146201048)

[Appendix 11 – Raw data of response rate for documentary suggestions and individual strategies 191](#_Toc146201049)

# Acknowledgements

I would like to take this opportunity to acknowledge and thank every single person that has helped me along the way of completing this research. It would not have been possible without the support from my supervisory team, friends and family and mostly the participants in all three studies.

Firstly, I would like to thank my supervisory team, Dr Sarah Fieldhouse, Dr John Wheeler, Mr William Morris for their continued support through regular meetings and guidance on the research over the years. Also, an extended appreciation to Sarah, as my principal supervisor and my critical friend, I could not have achieved this without you. Sarah has been my academic mentor since I started at Staffordshire University in 2012 and has supported me continuously throughout. Your expertise and professional advice but also a shoulder to cry on during tough times have been invaluable to me. I am profoundly grateful for the support over the years.

This research has been heavily reliant on participants and therefore, I would like to express my gratitude to all participants especially those which completed all parts of the studies. Together I believe we have achieved novel research to understand what is deemed as effective documentation for fingerprint examinations. Although, there is further work to consider, I hope we can work together in the future.

I would like to thank my fellow PhD researchers. Being there for each other in the highs and lows, sharing amazing experiences with you all through international conferences, sightseeing and networking during these conferences has made great memories.

My family have been a huge part of this journey, and I would like to acknowledge the support I have received from you all. Especially my mum and fiancé, Will. Mum, you are my best friend and seeing your dedication to provide the best for Amy and me have given me inspiration to be the best for my children. To achieve this, you have helped so much with childcare, looking after Oliver and then the twins, Esmé and Isla, when they made their arrival halfway through the research. Will, you have been my rock, constantly putting my needs before your own and being by my side throughout the difficulties. To both of you, I will be forever grateful. Finally, a special mention to my Nan, who sadly passed away in 2016, she always encouraged me to complete my doctorate. Well, 6 years later, during a pandemic, having my beautiful twin girls. I did it Nan!

# Abstract

The documentation of forensic evidence is an essential and routine part of the examination process. Issues relating to the documentation of fingerprint evidence were highlighted by the Fingerprint Inquiry (2011) and the case of R-v-Smith (2011). The mandatory introduction of ISO 17025 accreditation has forced changes to the documentation of fingerprint examination practice in the UK. The requirement for effective documentation to provide a sufficient auditable trail is consistent, yet despite numerous sources of academic and grey literature reporting on the information to be documented from fingerprint examinations, there appears to be relative dissimilarity in the guidance offered and subjectivity with its interpretation. Most of the guidance was published before the announcement of accreditation, therefore the documentary suggestions could already be in use. As there is no evidence of pre-accreditation documentation strategies in literature, only the criticisms in court, this research facilitated a gap analysis through professional conversations discussing what was recorded and the documentary suggestions from literature and policies. Subsequently, a confirmed list of documentary suggestions referred to as the ACE-V checklist in this document was created. The aim of this research was to provide evidence-based recommendations for the content of contemporaneous notes that are recorded by identification experts as part of fingerprint examinations in casework.

Study one involved 31 UK fingerprint experts who analysed eight impressions of varied quality. In the first part, the analysis was documented using the pre-accreditation documentation strategy with an additional free text box to gather the thought processes. In the second part, the analysis was documented using an online software. All documentation was cross-compared to documentary suggestions (also referred to as ACE-V checklist in this study). The results found more information was recorded consistently between experts when prompted using the online software. In addition, there were significant associations between documentary suggestions i.e., generic information and more factors and grade of the mark. These results indicate what could be within effective documentation. Taking this into consideration, a new documentation strategy was created.

Study two involved 33 UK fingerprint experts of these 27 experts were from study one. Due to workload, 4 experts could not continue participating in the study however, 6 new fingerprint experts from another UK bureau joined. Each expert completed a full analysis, comparison and evaluation (ACE) of the same eight impressions using the new documentation post-accreditation strategy. The documentation produced was cross-compared to the ACE-V checklist. Due to excessive workload, 11 of the 33 experts from part one of the study completed another full examination (ACE process) using the online software. The results have shown varied documentation strategies not only in approach but also the extent of notes within. Despite the variations, all the selected UK fingerprint documentation strategies have been accepted by United Kingdom Accreditation Service. There are significant associations between documentary suggestions i.e., level of detail and annotations, and comparison/evaluation notes and the grade of the mark. These findings show the variance of notes over the different quality of marks and contribute to what could be within effective documentation.

Although all have achieved accreditation, this is not the only measurement of effectiveness. As there are clear differences between strategies another means of measurement should be conducted to provide evidence-based research on an effective documentation strategy including the approach and extent of notes within. Within study three, a matching exercise was conducted to determine if participants could correctly identify the impressions matched the documentation presented side by side. A correct outcome will demonstrate if an individual, expert or novice, can successfully follow the notes of the original expert and therefore be deemed as effective. There are significant associations between obtaining a correct or incorrect outcome and strategy and grade of the mark. The most effective documentation strategies and their inclusion of documentary suggestions from the ACE-V checklist is used to suggest new minimum requirements for fingerprint documentation.

As experts completed multiple examinations of the same impressions, the inter and intra variance of reporting outcomes were investigated, which contributes to existing research. However, the novel aspect falls within the use of documentation to understand the variance of opinion. There were a number of occasions whereby experts changed their decision and using the documentation can allow an understanding of the thought process and if an error has occurred, then training can be provided where necessary. In some post-accreditation documentation strategies, no notes were provided for insufficient or inconclusive outcomes. As a result, the thought process could not be determined and provides additional support to include reasons for conclusions, not only the outcome.

# Glossary

**ACE** – An acronym which stands for Analysis, Comparison, Evaluation. It refers to the process used by forensic scientists when assessing the value and suitability of an impression for comparison purposes.

**ACE-V** – An acronym which stands for Analysis, Comparison, Evaluation and Verification. ACE process with an additional stage which is verification.

**ACE-V checklist** – This includes the documentary suggestions from literature and policies.

**Blind verification** – An independent verification done by a second practitioner without knowing the original outcome.

**Comparison and evaluation notes** – A sub-category of the ACE-V checklist. It refers to documentary suggestions to explain a comparison between a crime scene impression to a known print.

**Documentation** – Refers to day to day case notes, contemporaneous notes, working notes, examination sheets and/or the findings from the examinations.

**Generic information** – A subcategory of the ACE-V checklist. It refers to documentary suggestions standard information such as, unique reference number, date of examination, materials used in the examination. In addition, impression information such as anatomical source, single or multiple marks and relationship between marks.

**Grade** – A particular level of quality for impressions i.e., grade 1, 2, 3 and 4.

**Factors affecting the mark** – A sub-category of the ACE-V checklist. It refers to the documentary suggestions regarding quality of the impression and any potential distortion factors including location of the impression, orientation of the impression, the quality and quantity of impression, the matrix the impression is left in and a list of distortion factors (movement, substrate, overlapping).

**Level of detail and annotations** – A subcategory of the ACE-V checklist. It refers to documentary suggestions regarding level one, two and three detail and the presence of annotating impressions.

**Pre-accreditation** – Refers to any documentation created using strategies from before ISO 17025 was implemented.

**Post-accreditation** – Refers to any documentation created using strategies from after ISO 17025 was implemented.

**Open verification** – The opposite to a blind verification. A second practitioner conducts an examination knowing the original outcome.

**Semi-structured** – Refers to any documentation strategy includes some direct questions to answer and a free text box to explain their thought process.

**Structured** – Refers to any documentation strategy which includes prompts for users on what to record within the notes.

**Repeatability** – Refers to intra-examiner agreement, whether one expert consistently reaches the same decision.

**Reproducibility** – Refers to inter-examiner agreement, whether two experts reach the same decision.

# List of Figures

Figure 1 - A fingermark showing the core (circle) and delta (triangle) (Nawelmando & Marwala, 2011)

Figure 2 - Fingerprint pattern types (Champod et al, 2004)

Figure 3 - Examples of individual ridge characteristics (minutiae) (Champod et al, 2004)

Figure 4 - Fingermarks selected from Ground Truth Database for test.

Figure 5 - Participant response rate for the generic information against the grade of the mark   
(before accreditation)

Figure 6 - Participant response rate for the factors affecting the mark against the grade of the mark (before accreditation)

Figure 7 - Participant response rate for the level of detail/annotations against the grade of the mark (before accreditation)

Figure 8 - Participant response rate for the level of detail/annotations against the grade of the mark (using the online software)

Figure 9 - Mark Analysis Form

Figure 10 - Participant response rate for the generic information against the grade of the mark (after accreditation)

Figure 11 - Standardised residual counts highlighted to show significant association.

Figure 12 - Participant response rate for the level of detail and annotations against the grade of the mark (after accreditation)

Figure 13 - Bar chart showing the rate of outcomes to expertise group.

Figure 14 - Analysis outcomes for Grade 1 impression (image 1)

Figure 15 - Analysis outcomes for Grade 1 impression (image 4)

Figure 16 - Analysis outcomes for Grade 1 impression (image 8)

Figure 17 - Evaluation outcomes for Grade 1 impression (image 1)

Figure 18 - Evaluation outcomes for Grade 1 impression (image 4)

Figure 19 – Evaluation outcomes for Grade 1 impression (image 8)

Figure 20 - Analysis outcomes for Grade 2 impression (image 2)

Figure 21 - Analysis outcomes for Grade 2 impression (image 7)

Figure 22 - Evaluation outcomes for Grade 2 impression (image 2)

Figure 23 - Evaluation outcomes for Grade 2 impression (image 7)

Figure 24 - Analysis outcomes for Grade 3 impression (image 6)

Figure 25 – Evaluation outcomes for Grade 3 impression (image 6)

Figure 26 – Analysis outcomes for Grade 4 impression (image 3)

Figure 27 – Analysis outcomes for Grade 4 impression (image 5)

Figure 28 - Evaluation outcomes for Grade 4 impression (image 3)

# List of Tables

Table 1 - A summary of guidance relating to information to be recorded during fingerprint examinations according to published sources of literature.

Table 2 - The consensus grade and scores given to the marks examined.

Table 3 – Response rate for reporting outcome at Analysis stage

Table 4 - Response rate for generic information in both approaches

Table 5 – Response rate for factors affecting the mark in both approaches.

Table 6 - Response rate for level of detail and annotations in both approaches

Table 7 - Different types of post-accreditation documentation strategies

Table 8 - Analysis outcomes in the post-accreditation documentation strategy

Table 9 - Evaluation outcomes in the post-accreditation documentation strategies

Table 10 - Analysis outcomes using the structured approach.

Table 11 - Evaluation outcomes using the structured approach.

Table 12 - Response rate for generic information in both approaches

Table 13 - Response rate for factors affecting the mark in both approaches.

Table 14 - Response rate for level of detail and annotations in both approaches

Table 15 - Response rate for comparison and evaluation notes in both approaches

Table 16 – ACE-V checklist inclusion criteria for image 1 / grade 1

Table 17 – ACE-V checklist inclusion criteria for image 2 / grade 2

Table 18 – ACE-V checklist inclusion criteria for image 6 / grade 3

Table 19 – ACE-V checklist inclusion criteria for image 3 / grade 4

Table 20 - ACE-V checklist inclusion criteria for strategy 1 for all grades

Table 21 – ACE-V checklist inclusion criteria for strategy 2 for all grades

Table 22 – ACE-V checklist inclusion criteria for strategy 3 for all grades

Table 23 - ACE-V checklist inclusion criteria for strategy 4 for all grades

Table 24 - ACE-V checklist inclusion criteria for strategy 5 for all grades

Table 25 - ACE-V checklist inclusion criteria for strategy 6 for all grades

Table 26 - ACE-V checklist inclusion criteria for strategy 7 for all grades

Table 27 – Counts for incorrect and correct outcomes from expert and novices.

Table 28 - Counts for correct and incorrect outcomes for all the documentation strategies

Table 29 - ACE-V checklist inclusion for most effective strategies  
Table 30 – Counts for incorrect and correct outcomes for each grade.

Table 31 - The eight impressions used within this study.

Table 32 - Key to interpret the graphs.

Table 33 - Repeatability and reproducibility of experts reporting outcomes on a Grade 4 impression (image 3)

# Chapter 1: Introduction

This chapter introduces the background and theory of friction ridge detail and the fingerprint examination process. Followed by a literature review on the documentation procedures for fingerprint examinations. This includes the expected quality standards adhered to within fingerprint examinations and the guidance within the published literature and policies on what to record within the documentation.

Section 1.1 provides a basic introduction to friction ridge detail, explaining the development of friction ridge skin and the level of detail which can be found in friction ridges. It is important to provide the fundamentals of friction ridge detail to ensure a basic understanding of the topic, but it is unnecessary to include an in-depth review of the development and detail within friction ridges as the focus is on documentation strategies. Similarly, section 1.2 introduces the widely accepted fingerprint examination process (ACE-V) but with an emphasis on documenting this process. Previous literature (Ashbaugh, 1999; Cowger, 1993; Champod *et al,* 2004) provide an in-depth review of the background and theory of friction ridge detail and fingerprint examination process. Section 1.3 reviews the current landscape of the requirement and guidance on fingerprint comparison contemporaneous notes. An overview of documentation procedures in relation to fingerprint examinations is provided and divided into separate sections. The first explains a reporting process which was introduced in 2013 in the UK known as streamlined forensic reporting (SFR). This was for all forensic disciplines, not only fingerprint examinations (section 1.3.2). In 2015 The Forensic Science Regulator (FSR) announced that all UK fingerprint bureaux were to become ISO accredited by October 2018. In part, ISO 17025 refers to the production of documentation to abide by the accreditation requirements. This standard coincides with ILAC G-19 an international standard for forensic processes. These quality standards are detailed in section 1.3.3. The SFR process will remain in place as well as bureaux abiding by ISO standards, therefore it is important to understand both processes. There are many documentary suggestions in literature and policies worldwide which could aid with the adoption of daily documentation for fingerprint experts, as this is new to all bureaux in the UK. Therefore, sub-section 1.3.4 critically analyses the literature of worldwide policies and academic research on the inclusion criteria for the documentation of fingerprint examinations and anecdotal reports of practice. The differences between literature highlight the necessity of establishing effective documentation strategies. This chapter will end with the research aim and the objectives to meet this aim in section 1.4.

## 1.1 - Introduction to Friction Ridge Detail

The skin surface, which is found under the fingers, palms of the hands and soles of the feet, is different to skin on any other part of the body. It is made up of a series of lines known as ridges and furrows, which are created during foetal development in the womb, and this is called friction ridge detail (Forensic Science Regulator, 2017). Champod *et al* (2004) explains the morphogenesis of friction ridge skin, which starts during the very first weeks of foetal development. The hand starts to develop between 5 to 6 weeks, then the fingers develop a couple of weeks later. During this time volar pads appear on the palm and start to appear on each finger between 7 and 8 weeks. Volar pads are formed due to the swelling of the tissue under the epidermis layer of the skin, which are the fundamental aspects of friction ridge skin development. The primary ridges are developed first, and then the general pattern. The pattern is dependent on the shape and size of the volar pads, the onset of primary ridge formation and bone morphology. Symmetrical volar pads usually develop into whorls. Whereas displaced pads tend to develop into loops and the type of loop is dependent on the placement of the pad, for example a pad more to the right will create a loop to the right and a pad more to the left will create a loop to the left. Very low volar pads will develop arches (Champod *et al,* 2004).

The features of friction ridges are categorised into level 1, 2 and 3 detail (Ashbaugh, 1999; Cowger, 1993; Langenburg, 2012; SWGFAST, 2013). Level 1 detail refers to the “pattern type” of the fingermark. The ridges flow to form patterns (Forensic Science Regulator, 2017) and there are three main types of patterns which are loop, whorl and arch (Langenburg, 2012). There are certain characteristics that help to identify patterns such as cores and deltas. A core is located near the centre of the fingermark and deltas are similar shape to a triangle with two of the branches opening out towards the core (figure 1).

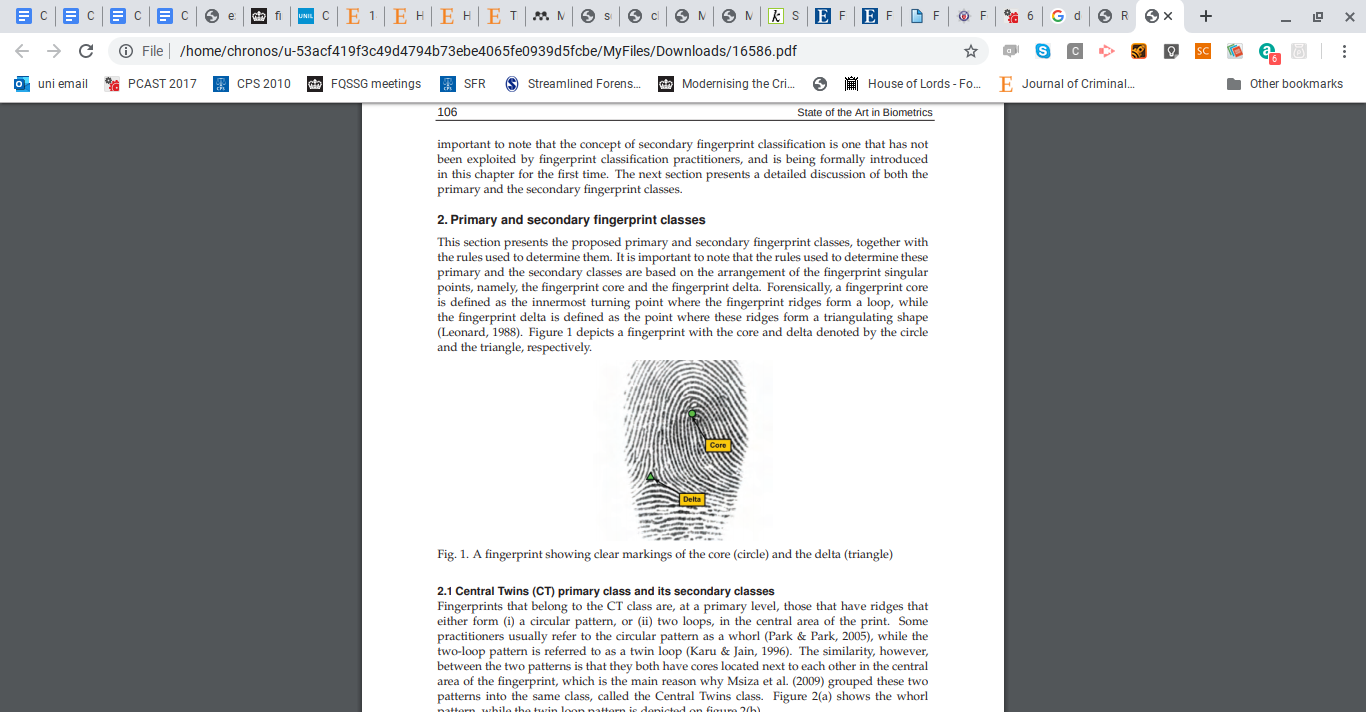


Figure 1: A fingermark showing the core (circle) and delta (triangle) (Nawelmando & Marwala, 2011)

Loops are the most common type of patterns (Yager & Amin, 2004; Bhavana *et al*, 2013). A loop must contain one delta and one core, and the ridges enter one side of the finger and recurve to exit the same side. A whorl contains at least two deltas and a core, and the ridges must complete a 360° formation, whether this is one continuous ridge or short ridges. Arches are the least common pattern. There are no cores or deltas within this pattern, only the ridges flowing from one side of the finger and exit the opposite side. Figure 2 shows examples of arch, loop and whorl patterns.

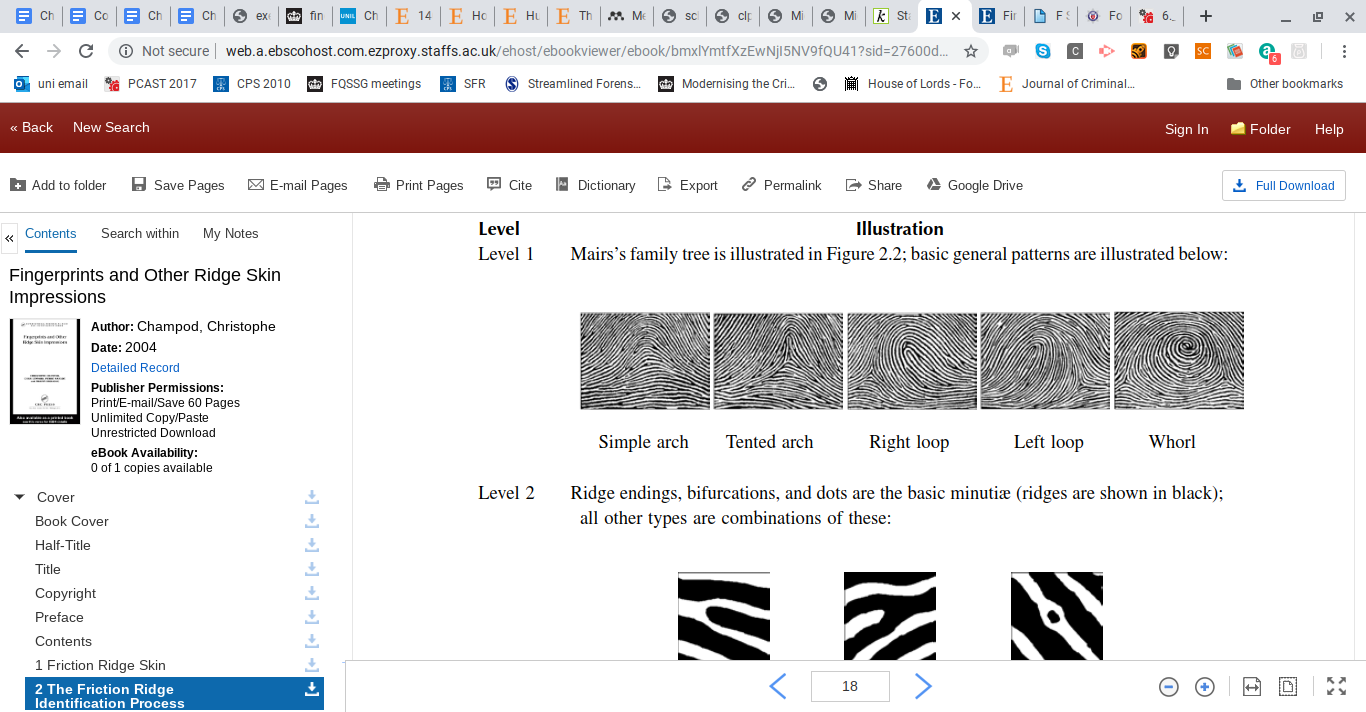


Figure 2: Fingerprint pattern types (Champod *et al*, 2004).

Level 1 detail can be used for excluding individuals from a case but is not sufficient for an identification due to the commonality of the patterns (Ashbaugh, 1999). The uniqueness of friction ridge impressions relates to the level 2 detail, which are the individual ridge characteristics also known as minutiae. The ridge path deviation forms basic minutiae such as bifurcations and ridge endings. A bifurcation is where one ridge splits into two, and a ridge ending is where a ridge stops abruptly (Cowger, 1993; Champod *et al*, 2004; Langenburg, 2012). These are the most defining features, see figure 3 for examples of the minutia types.

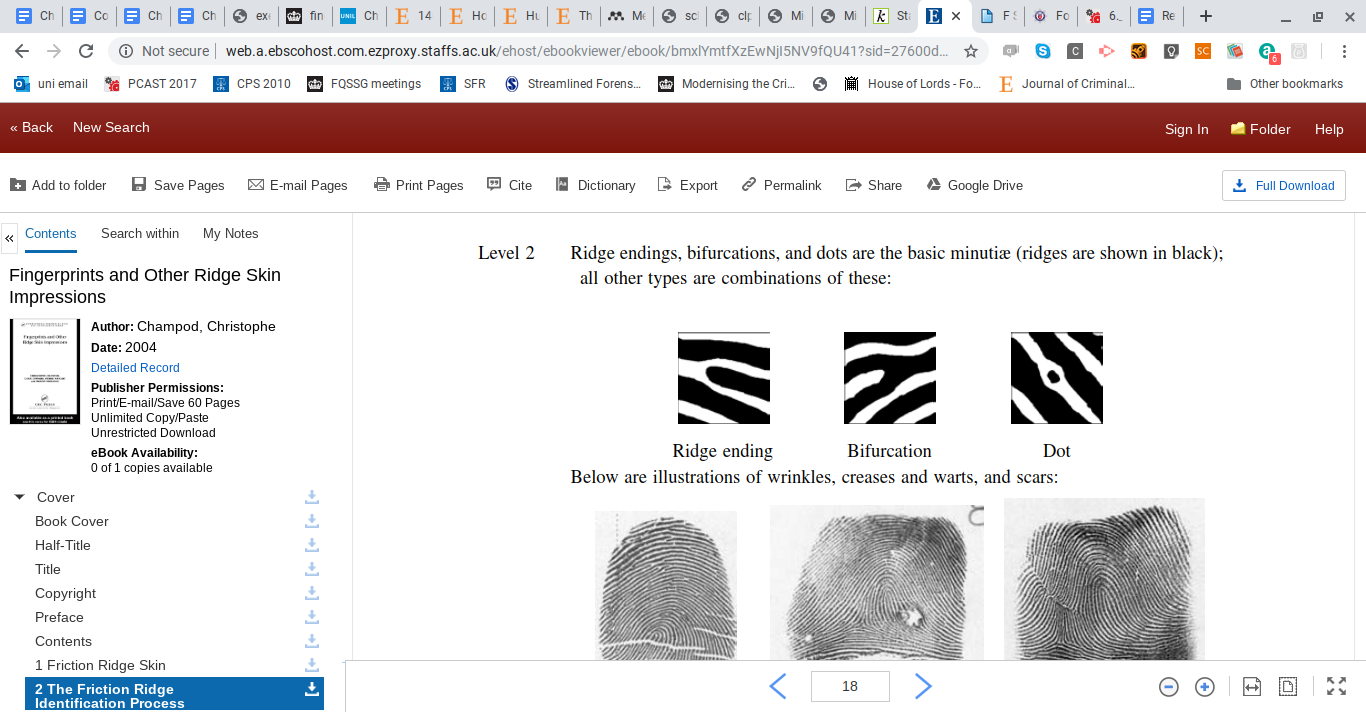


Figure 3: Examples of individual ridge characteristics (minutiae) (Champod *et al,* 2004)

Additionally, level 3 detail can be used as unique individual characteristics. This detail includes ridge shape and pores (Langenburg, 2012). When these characteristics are present, it can be highly useful however, it is rare for this level of detail to be present in actual case work (Langenburg, 2012). SWGFAST (2018) explains that certain features such as scars, creases and warts may be prevalent in all three levels of detail.

Monson *et al* (2019) studied the permanence of friction ridge skin and the persistence of friction ridge impressions and as a result, it appeared the permanence and persistence varied for all levels of detail. Level 1 detail was permanent in the skin and persistent in the impressions, in other terms the pattern within a fingerprint or fingermark will always remain, it cannot be altered. Whereas level 2 detail was persistent in the skin and impressions and permanent in short term, not long term. Level 3 detail was neither persistent nor permanent as it is likely the quality of a fingermark will not be sufficient quality to present level 3 detail such as pores.

## 1.2 - Introduction to Fingerprint Examinations (ACE-V process)

The purpose of fingerprint examination is to compare two friction ridge impressions to determine whether they were made by the same individual or not (Forensic Science Regulator, 2017). The examination of fingerprints is a cognitive process whereby examiners are required to use their expertise to perform four stages: Analysis, Comparison, Evaluation and Verification (Forensic Science Regulator, 2016), commonly known as ACE-V. Ashbaugh (1999) was the first to describe how ACE could be used in fingerprint examinations, with the inclusion of a final stage “verification”. It was noted that the final stage was not a part of the identification process, but a- quality control step. The methodology utilizes a qualitative and quantitative assessment of level 1, level 2, and level 3 details, as previously explained (SWGFAST, 2013).

### 1.2.1 – Analysis stage of ACE-V process

This is the first stage of the ACE-V process. Ashbaugh (1999) compares the analysis stage to intelligence gathering during which as much information as possible is obtained and subsequently used for the comparison stage. Information such as the substrate which the mark has been deposited on, the matrix, the development technique and pressure applied during deposition are recorded, as these factors have the potential to distort the mark and affect the suitability for comparison; this has also been addressed by other authors.

The Forensic Science Regulator published the Codes of Practice and Conduct for Fingerprint Comparisons in 2017 for practice in the UK which outlined the objective of each phase. It states during the Analysis stage, a practitioner examines an unknown friction ridge impression considering all variables influencing the impression to determine its suitability for comparison. Several factors must be considered, such as the clarity of the impression, the substrate on which the impression has been left on and any apparent distortion.

If there is sufficient information available in the mark, the expert will proceed to the next phase (Forensic Science Regulator, 2017). In addition, the expert will take into consideration the anatomical aspects for example, digit or palmar determination (Ashbaugh, 1999, Champod et al, 2004, ENFSI, 2015). Determining the potential digit or palmar area is important to narrow down the database search.

Other authors have discussed the Analysis phase and although in agreement with the statement in the Codes of Practice, there can be additional suggestions. Langenburg (2012) describes the Analysis stage as a pre-assessment stage where the decisions are made prior to the Comparison stage. The key elements are determining the origin and orientation of the mark, assessing the quality and quantity of the mark and type of distortion. The immediate difference between the UK practice and this author, which should be noted this is American practice, is the selection of features and assigning confidence levels these features. Such systems as the “GYRO system” can be used to indicate confidence levels. G = green which is high confidence is observing the feature and using it during the Comparison stage, Y = yellow which is medium confidence, R = red, which is low confidence, O = orange which are features observed in the unknown impression during the Comparison stage (Champod & Langenburg, 2011). This system is deemed as a documenting procedure which will be further discussed in section 1.3. The use of the GYRO system has been supported by the European Network of Forensic Science Institutes (2015). ENFSI (2015) and SWGFAST (2013) states that the Analysis phase determines its suitability dependent on level 1, 2 and 3 detail. Similarly, to Langenburg (2012), it implies examiners following the SWGFAST standard assign confidence levels to features observed. Despite the differences in explanations, UK fingerprint examiners will follow the Forensic Science Regulator’s definition.

### 1.2.2 - Comparison stage of ACE-V process

After the expert has deemed the unknown fingermark as sufficient for comparison, the mark is compared to known fingerprints, either from suspects or an automated fingerprint identification system. This phase involves a side-by-side comparison to determine whether the features on the unknown mark observed in the Analysis phase are in agreement with features in the known print (The Forensic Science Regulator, 2017). SWGFAST (2013) goes into further explanation about the Comparison phase. It begins with determining the level 1 detail between two impressions. If the same pattern type is present in both impressions, then, a target group of minutiae is selected in the unknown mark and searched within the known fingerprint. Langenburg (2012) also explains the use of a target group which refers to selecting one area of the mark and cross checking this area on the known print, if available.   
In some cases, the area which has been targeted can be distorted in the known fingerprint therefore, another target group in the unknown fingermark would have to be used when comparing to the known fingerprint. This process has been seen by the research team in UK practice. Side by side comparison could potentially cause bias as the practitioner has both prints in front of them and if characteristics are seen in one then it can influence the examiner to see the characteristics in the other mark.

Within the UK, practitioners look for characteristics in a coincident sequence. This is when the same characteristics are in the same position with the same ridge count between them. The ability or inability to establish this sequence will aid the practitioner with their conclusion (The Forensic Science Regulator, 2013).

### 1.2.3 - Evaluation stage of ACE-V process

The practitioner reviews their observations and determines the outcome of the examination (The Forensic Science Regulator, 2017; SWGFAST, 2013; Langenburg, 2012). In the UK, the terms used are

* Identification - There is sufficient detail in agreement with no unexplainable differences to suggest the two impressions originated from the same source.
* Exclusion - There is sufficient amount of detail in disagreement to suggest the two impressions did not originate from the same source.
* Inconclusive - There are points in agreement and disagreement therefore not possible to determine if the impressions originated from the same source or not.
* Insufficient - A poor quality mark which cannot be examined.  
  (The Forensic Science Regulator, 2017).

ENFSI (2015) explains the different approaches adopted worldwide which are holistic, probabilistic and numerical. Fingerprint examiners working to the holistic and the numerical approach, determine their conclusions based on corresponding configuration, or in UK terms “coincident sequence” as previously explained. Numerical is slightly different as some countries which work to this approach must reach the threshold to aid with their conclusion. This is a pre-determined required number of minutiae and does not include other features considered in the holistic approach such as creases, pores. This approach was used in the UK, all fingerprint experts were required to identify 16 points or more for an identification but was abolished in 2001. In replacement of the numeric standard, experts were required to have their work triple checked before it is presented in court (Mackay, 2001). According to the probabilistic approach, the outcome of the evaluation phase is the estimation of the evidential strength of the findings under two competing hypotheses. The evidential strength is reported as a likelihood ratio, Langenburg (2012) supports this approach.

### 1.2.4 - Verification stage of ACE-V process

An independent analysis, comparison and evaluation of a fingermark is completed by a second practitioner to demonstrate whether the same outcome is made to the original practitioner (The Forensic Science Regulator, 2017). This is also known as a quality control heck (Langenburg, 2012). The verification phase can be either blind or open. Blind is an independent verification done by a second practitioner without knowing the original outcome. Open verification is the opposite, whereby the verifier knows the original outcome (The Forensic Science Regulator, 2017; ENFSI, 2015). A blind verification is preferred to minimise the risk of errors due to bias, specifically due to contextual information (ENFSI, 2015). If there is a dispute between two practitioners, then there should be a policy to resolve the conflict.

### 1.2.5 – The reliability of the ACE-V process

Despite the ACE-V process being an accepted approach within fingerprint examinations, concerns of the reliability were raised in the US National Research Council (NRC) of the National Academies of Science report (National Research Council, 2009). Primarily due to a prominent misidentification of a crime scene mark recovered from the 2004 bombing of the Madrid commuter train system (National Research Council, 2009). In the UK, fingerprint evidence has been scrutinised within the infamous cases of Her Majesty’s Advocate (HMA)-vs-McKie (1997) and R-vs-Smith (2011). In brief, the England and Wales Court of Appeal (EWCA) in R-v-Smith (2011) quashed a homicide conviction due to issues with the fingerprint examination process. A fingermark originally found to have ‘insufficient detail for a meaningful comparison’ was re-examined by the same examiner who then reported 12 corresponding ridge characteristics with the accused’s fingerprints. An absence of working notes within the original and twice verified examinations contributed to the Court of Appeal judgement. It was stated that, “no competent forensic scientist in other areas of forensic science these days would conduct a forensic examination without keeping detailed notes of his examination and the reasons for his conclusions” and suggested that quality standards should be enforced through a “robust and accountable system”. In addition, The Fingerprint Inquiry (2011) was the result of HMA-v-McKie. Shirley McKie, a detective constable within the murder investigation team, attended a potential murder scene, where examiners working at the Scottish Criminal Record Office claimed to have found a fingermark from McKie on a doorframe inside the property. McKie denied having entered the property and was prosecuted for perjury. At trial, it was decided that McKie was not guilty and that a misidentification of the mark had been made. This led to the Fingerprint Inquiry, a public inquiry which resulted in 86 recommendations for the discipline, some of which related to the documentation of fingerprint evidence.

Recommendations 50 to 53, page 747 of The Fingerprint Inquiry report (Campbell, 2011).

* “Recommendation 50 - Examiners should always take notes when they are examining marks that they consider to be complex.”
* “Recommendation 51 – Notes should be taken in any case in which a fresh comparison is made in response to a request from the Crown for a report.”
* “Recommendation 52 – Where notes are required as a result of the preceding recommendations, the notes should be taken at each stage of ACE-V by every examiner involved in the process at that stage and should cover the following matters: (i) the assessment of the quality of the mark at the analysis stage and any sign of distortion; (ii) the characteristics identified at analysis including their type and the sequence of them; (iii) the characteristics taken into account at the comparison stage including their types and sequence in mark and print; (iv) any revision to the initial analysis made at the comparison stage; (v) any differences observed at the comparison stage; (vi) the explanation for any differences; (vii) any third level detail relied upon in arriving at the conclusion; (viii) the reasons for the conclusion at the evaluation stage; and (ix) any consultation with any other examiner during the ACE-V process.”
* “Recommendation 53 - Subject to any requirement under ISO 17025 and recommendations 50 and 51, note-taking as to the detail found on analysis and the process of comparison, though not mandatory, should become the general practice for all fingerprint comparison work.”

These recommendations are discussed further and compared to other documentary suggestions by CPS (2010), Champod and Langenburg (2011), SWGFAST (2013), ENFSI (2015) and The Forensic Science Regulator (2017) in section 1.3.4.

As a result of the NAS report, the reliability and reproducibility of fingerprint experts’ conclusions have been investigated using a black box study approach (Ulery *et al,* 2011, Ulery *et al,* 2012, Tangen, *et al,* 2011, Pacheco *et al,* 2014*,* Eldridge *et al,* 2021*)*. A black box study only tests the fundamental aspects of that process and has no or little relevance with the internal logic and structure. For example, the variance within experts’ reporting outcomes. A very low rate of false positives (matching of fingerprint that did not originate from the same source) was observed by Ulery *et al,* (2011).

A lack of consensus is noticeable among the impressions deemed as identification; 48% on mated pairs and 33% on non-mated pairs. The lack of reproducibility seen here was then compared to repeatability also known as intra-examiner variance. 89.1% of identifications and 90.1% of exclusions were repeated, most changes were to inconclusive decisions (Ulery *et al,* 2011)

Interestingly, none of the false positive rates were repeated (Ulery *et al,* 2011, Ulery *et al,* 2012). In relation to wrong exclusions and wrong associations, Ulery *et al* (2011) reported 7.5% and 0.1%, Tangen *et al* (2011) reported 7.88% and 0.68%. Overall, the lack of repeatability and reproducibility increases with the difficultly of the impressions (Champod, 2015).

In addition, other studies have focused on identifying the sufficiency decisions these are known as white box studies. A white-box study is an investigation of the internal logic and structure of a process. For example, understanding the decision making of an outcome by a fingerprint expert.

Some studies have investigated what fingerprint experts considered as sufficient information for identification decisions (Neumann *et al,* 2013, Ulery *et al,* 2013, Ulery *et al,* 2014). As a collective, the sufficiency decisions relate to the number of minutiae, as well as a quality map, to identify an unknown fingermark to a known fingerprint. When focusing on minutiae count, the range varies between experts. However, this is not new as previously seen in Evett & Williams (1996). Although no efforts have been made to manage this variability (Champod, 2015). That said, the 16-point standard was abolished in the UK in 2001 and since then, the UK uses a holistic approach. Due to this, and the fact fingerprint experts are required to produce documentation to adhere to ISO accreditation, this research focuses on the contemporaneous notes and attempts to understand conclusions using notes rather than number of minutiae identified.

The research work into the reliability and reproducibility has been acknowledged and two appropriately designed black box studies have been applauded by PCAST (2016). The authors find that fingerprint examinations are a foundationally valid subjective method. Although, additional black box studies are needed to clarify the reliability of the method. Also, these black box studies have highlighted important issues that still need to be addressed such as confirmation bias, contextual bias and proficiency testing.

PCAST (2016) reports experts should be required to document their analysis before looking at the known print and separately document any additional information used in the comparison and evaluation.

## 1.3 – Literature Review

By the end of this chapter, there should be a greater understanding of the quality standards underpinning the documentation procedures for fingerprint examinations. Furthermore, an understanding of the published guidance on what to potentially include within documentation when conducting fingerprint examinations. The requirement of quality standards bound by legal statuary powers and the variance within guidance will demonstrate the importance of this research.

### 1.3.1 – The documentation procedures in relation to Fingerprint Examinations

Documenting the examination of forensic evidence is a routine task for forensic scientists across forensic disciplines. Contemporaneous notes are widely used to record the nature of the examinations undertaken, the time and date of examinations and the results and conclusions. In Forensic Science, the use of contemporaneous notes is particularly important to provide transparency to the courts and other sources of external scrutiny about how the evidence has been processed. This is particularly important where evidence is ‘opinion based’, as it is in many fields of forensic science, such as fingerprints. Also, the requirement to maintain a ‘chain of custody’, that records the location of evidence and its transmission within the Criminal Justice System is critical to ensuring its integrity and probative value. In the UK, all expert witness’ are required to document their examinations according to the Code of Practice outlined in the Criminal Procedure and Investigations Act (CPIA) (Crown Prosecution Service, 1996).

In section 1.2.1, the definition of the Analysis stage is expressed by numerous authors. In brief, it is the intelligence gathering stage and the level of detail is assessed to determine its suitability for comparison. Any factors that are affecting the impressions are taken into consideration. Ashbaugh (1999) states that these can be mental observations and the information is not necessarily recorded unless it is deemed as a “complex” mark”. In recent publications (Forensic Science Regulator, 2019) the term “challenging” is also recognised as an alternative for “complex”. However, the term “complex” will be used throughout the thesis. The issue with only producing notes for “complex” marks, is defining what a complex mark is as the definition differs between academics, practitioners and governing bodies. Ashbaugh (1999) defines complex as a mark having more than two distortion issues. These issues can be one or more of the highlighted factors previously mentioned. In addition, SWGFAST (2011) agrees with Ashbaugh’s definition but also includes conflict amongst examiners. In accordance with the standard published by SWGFAST, the extent and detail of documentation will vary based on the complexity of the case. A complex mark will have more detail and documentation. Langenburg (2012) recommends this standard as a means of documenting the ACE process, however, goes further and suggests using a system like GYRO or similar annotation system to declare the expert’s certainty regarding the reliability of the feature identified during the Analysis phase. However, as discussed by Bunter (2016), it is not common practice for fingerprint experts to annotate the impressions.

The Forensic Science Regulator (2017) goes further and includes the following as complex marks too; an insufficient mark subsequently identified, a mark that has been compared to the identified person with a negative result, a mark that is compared even though there are few clear characteristics, a mark in blood and/or the only mark identified in the case. The issue with the determination of a complex mark is that the complexity is not decided until the end of the examination in most situations. Therefore, the documentation would not be produced contemporaneously.

ENFSI (2015) contradicts Ashbaugh’s statement and suggests the information is recorded during the Analysis and used during the Comparison stage, not just for complex marks. The documentary suggestions from ENFSI (2015) and other authors are critically analysed in section 1.3.4.

In instances where ‘differences of opinion’ between examiners of the same evidence arise, the use of contemporaneous notes can help to explain differences and avoid instances such as those encountered in R-v-Smith (2011) and HMA-vs-McKie (1997).

In both cases, a lack of effective documentation ultimately led to the fingerprint evidence being open to interpretation. Following the case of HMA-vs-McKie, The Fingerprint Inquiry was undertaken, from which 86 recommendations were made (Campbell, 2011), some of which related to the documentation of fingerprint evidence. According to The Court of Appeal judgement in the case of R-v-Smith (2011) “no competent forensic scientist in other areas of forensic science these days would conduct a forensic examination without keeping detailed notes of his examination and the reasons for his conclusions” (R-v-Smith, 2011). Bunter (2016) suggested that there had been little change to fingerprint examination documentation strategies since the publication of the Inquiry, although the mandatory introduction of ISO 17025 into fingerprint bureaux in England and Wales by the Forensic Science Regulator, and its implementation to the Scottish Police Services Authority is a likely driver for change. The standard aims to ensure that laboratories are technically competent and includes references to the documentation of examinations. A more in-depth review of quality standards, UK accreditation service and the role of the Forensic Science Regulator is within section 1.3.3.

Prior to the accreditation, there has been no standard within the UK which instructs experts to record in depth their examinations. As evidenced in R-v-Smith (2011), for an identification, a ‘one line’ report was used in court. This documentary procedure remains despite the criticisms, as outlined by Bunter (2016). Another study involving UK examiners evidenced that fingerprint experts consistently defined the ACE-V process, which implies the process remains the same since Ashbaugh’s contribution to fingerprint examinations. There appears to be no discussion of recording this information, despite participants’ emphasis placed on the analysis stage which was often referred to as the information gathering stage (Stevenage, 2016). Mental observations, the production of notes for only complex marks and retrospective notes will not suffice under new accredited guidelines which were implemented in October 2018. Therefore, using pre- and post-accreditation documentary suggestions from literature and policies, new documentary procedures need to be sought.

The deadline for ISO 17025 accreditation was October 2018. However, not all UK fingerprint bureaux met this deadline and were allowed an extension to early 2019. One of the reasons for failing the accreditation was the documentation procedures. UKAS were not prescriptive on what should go into the documentation; therefore, it was open to interpretation from sources of scientific and grey literature already available and for institutions to consider their own ways of working to meet with ISO. There appears to be inconsistencies between these sources which could explain the varied documentation procedures adopted throughout the UK and provide an explanation for the failings of accreditation.

Furthermore, forensic scientists in the UK must comply with Streamlined Forensic Reporting (SFR) (section 1.3.2). In brief, it is a case management process to reduce costs and delays within courts (McCartney, 2019). Practitioners may have been mindful of this process when developing the documentation procedures. There should be a balance between the time taken to produce the documentation and the level of detail within it to align with the SFR process and gaining accreditation.

### 1.3.2 - Streamlined Forensic Reporting: Legal requirement for all forensic evidence types.

Streamlined Forensic Reporting(SFR) was a process that was introduced to England and Wales in 2013 (after The Fingerprint Inquiry) and was designed to enable investigators, scientists and prosecutors to comply with the Criminal Procedure Rules (Richmond, 2018). Part 1 of the Criminal Procedure Rules (Ministry of Justice, 2015) explained that the ‘overriding objective’ was to deal with criminal cases justly, including acquitting the innocent and convicting the guilty, dealing with the prosecution and defence fairly and dealing with cases efficiently. SFR was a revised case management procedure for presenting forensic evidence at court. It reportedly reduced ‘unnecessary cost’ and ‘delays’ in the Criminal Justice System (Crown Prosecution Service, 2015). The SFR procedure was piloted for one year and the results found that there was an increase in guilty pleas and fewer discontinued cases following its use (Richmond, 2018).

There are two types of reports that may be created which are related to the documentation of examinations: SFR1 and SFR2. The SFR1 report consists of a summary of the forensic evidence and is neither a witness statement nor an expert’s report. It requires the defence to comply with their duties and to identify any issues in the case. The form is used to report the results from an initial analysis, comparison, data or database hit. Section 8.1 of the National Streamlined Forensic Reporting guidance states that the content of the SFR1 varies between circumstances, but it may contain the results that were obtained; a comparison of those results; and/or the result of the comparison or analysis (Crown Prosecution Service, 2015). In terms of fingerprint examination, if an expert had made an identification, contemporaneous notes were to include the name of the expert, the date of the examination, the name of the identified person and the digit or palm used in the identification. This knowledge was acquired through the authors observation of bureaux practice. If the reporting outcome did not report an identification, the outcome would simply be recorded on the case management system. There was an absence of reference to the analysis or comparison stages in the SFR1. This could make any subsequent audit of the work difficult to assess and reproduce, which could have significant implications for any instances where a ‘difference of opinion’ exists, such as that witnessed in the cases of HMA-V-McKie (Campbell, 2011) or R-v-Smith (2011). Furthermore, in the case of R-v-Smith, fingermarks taken from the crime scene were re-examined. An absence of working notes failed to provide transparency to the examinations that were undertaken, which contributed to the conviction of a murder being quashed.

If the case is disputed by the defence expert, an extended report is produced, known as an SFR2 report. The SFR2 report aims to answer any questions and/or disagreements that the defence expert has (Crown Prosecution Service, 2015), which forms the content of the documentation. It is presented as a witness statement, with an expert’s declaration under Criminal Procedure Rules 19.4(j) (Crown Prosecution Service, 2015) and the 2015 Criminal Practice Directions 19B, if required (Courts and Tribunals Judiciary, 2019). This means that the witness can be called to trial if their evidence is not agreed. According to Bunter, when a full evidential statement is made for court more detailed analysis notes are made, suggesting that some of the notes prepared for Court may be retrospectively recorded and not contemporaneous (Bunter, 2016).

It is reasonable to accept that in some instances working notes are made at the point at which is practicably possible, such as following a DNA clean environment or filtered light analysis, where the working environment may prevent notes from being recorded. A significant issue and risk with the retrospective process described by Bunter, is the opportunity for bias since the scientist is writing notes to support a known conclusion. Without making notes at the time of the examination there is no evidence to support the opinion of the scientist that the notes concur with the thoughts and opinions that the person held at the time of the original examination. Also, the SFR1 and SFR2 reports are not necessarily completed by the same person, since the SFR1 is not necessarily produced by a person qualified to give evidence or who has undertaken the forensic analysis (Crown Prosecution Service, 2015). If the person writing the report is not the person that carried out the initial examination, then the difficulties associated with the production of retrospective note taking are likely to be magnified, as the reporting examiner has to assume the opinion of the former examiner.

It is possible that evidence used for preparing a statement for court may include materials which were not used or available at the stage of the initial identification being made. For example, if the arrest set of fingerprints disclosed different areas of friction ridge skin to those originally used to identify the mark. This would understandably create differences between the content of the initial examination and any subsequent examination.

### 1.3.3 – Quality Standards for Fingerprint Examinations

The Forensic Science Regulator ensures that the provision of forensic science services across the criminal justice system is subject to an appropriate regime of scientific quality standard. In 2015, the FSR announced all fingerprint bureaux must obtain the appropriate quality standard, ISO 17025, by October 2018. By this date, only three bureaux achieved this non-legally binding deadline. The FSR used this as evidence for the need of statutory enforcement powers. A private members bill to establish statutory powers for the FSR was introduced by Parliament in 2020. The Forensic Science Regulator Act (2021) was passed by Parliament in April 2021. The act introduced powers for the FSR to “intervene where there is reason to believe a person i.e., body of persons corporate or unincorporate may be undertaking Forensic Science activity which this Code applies in a way that creates a substantial risk of adversely affecting any criminal investigation and/or Impeding or prejudicing the course of justice in any proceedings.” (The Forensic Science Regulator Act, 2021). Within a UK fingerprint bureau, it is expected the following the quality standards are met; ILAC G-19, ISO 17025 and the Codes of Practice and Conduct from the Forensic Science Regulator. In part, these refer to the production of documentation during fingerprint examinations. An overview of ILAC G-19 and ISO 17025 accreditation will be provided with a focus on the documentation aspects of the standards due to the research design. The Codes of Practice and Conduct is detailed in section 1.3.4 guidance for documentation for fingerprint examinations due to the documentary suggestions within the Codes.

#### 1.3.3.1 - ILAC G-19

ILAC is a global association for the accreditation of laboratories, inspection bodies, proficiency testing providers and producers of reference material. In 2002, ILAC published guidelines for forensic science laboratories however, this was replaced in 2007 to cover forensic science as a whole process and therefore, provide guidance on both ISO 17020 and ISO 17025. A revised document has recently been published in 2022 (ILAC, 2022).

The document covers general guidance to all activity modules in the forensic science process i.e., competence, records, risks and opportunities, witnessing scene of crime activities, non-conforming examinations, method validation, environmental conditions and equipment traceability. For this research, it is important to discuss the guidance in relation to records, risks and opportunities and non-confirming examinations.

Section 3.5 relates to “Records” (ILAC, 2022). It states there shall be documented procedures to maintain and retain records relating to each case. The records shall be sufficient for an auditable trail and include but not limited to the following: records of any communication with the customers, item receipts, descriptions of items, records of examination results, reference to procedures used with any diagrams, print outs, photographs or digital videos. The standard requirements are very broad to suit all forensic processes therefore it is important to review other literature to gain an understanding of the specifics of what to document within the records of examination results. This review of other literature is within section 1.3.4.

For a true reflection of the thought process of the original expert, records shall be made at the time of the examination. However, the standard allows for records taken as soon as practicable thereafter implying retrospective note taking is permitted.

To cover all forensic science processes, the standard allows for numerous recording methods such as drawing or writing, diagrams, photocopies, computer, sound recording, voice recording, photographs, digital images, video, and 3D laser scanning. Where technical abbreviations are made in records, these abbreviations should be clearly defined and readily understood. This allows flexibility within UK fingerprint bureaux to record as it is deemed appropriate for that force.

Section 3.6 refers to considering the risks and opportunities associated with forensic processes. Risk based thinking is a proactive approach to reduce potential undesired effects through early identification, planning and action. As stated, the guidance is broad to cover all forensic processes therefore most of the examples are unrelated to fingerprint examinations. However, it does mention a risk to consider is quality control or assurance to give confidence in the output for example, peer review or proficiency testing. This is also discussed in ISO 17025 accreditation and the Codes of Practice and Conduct for Fingerprint Examinations in more detail and subsequently falls within those section in the thesis.

Non-conforming examination refers to any aspect of the forensic process, including, examinations, results or expert witness testimony that do not conform to the bureau policies, procedures or the agreed requirements of the customer (ILAC, 2022). This refers to errors, it must not be confused with difference of opinion. There shall be policies and procedures in place to identify non-conforming work and subsequently policies and procedures when non-confirming work is identified. ILAC (2022) suggests ways to monitor this by peer review, proficiency tests and collaborative exercises, auditing and customer complaints.

Initially the significance of a non-conformity in relation to the validity of examination/test results shall be evaluated and its root cause identified. This shall include thoroughly investigating the review of casework already reported. Therefore, it is crucial to have effective documentation procedures to allow for the errors identifiable.

Where it is found that the forensic unit has issued a report containing non-conforming work that is deemed to significantly affect the result, the customer shall be notified immediately, the work or report recalled (where possible) and additional work or report issued by the forensic unit. Where it has been identified that the non-conforming work could recur, appropriate corrective action shall be implemented. This shall include the potential review of casework already reported prior to the non-conforming work being identified and implications for other cases and other sections in the forensic unit as well as implications for the forensic unit’s own internal policies and procedures, for example, through a review of the risks and opportunities. The non-conforming work and all actions taken shall be recorded.

#### 1.3.3.2 – ISO 17025 accreditation

UKAS is the National Accreditation Body for United Kingdom. The organisation is appointed by the government to assess and accredit organisations who provide services including certification, testing, calibration, and inspection. ISO 17025 is an international standard requirement for testing and calibration within laboratories. It enables laboratories to demonstrate their competency and produce reliable, valid results. As fingerprint examinations are opinion based, ISO 17025 standard applies better to opinion and interpretations as opposed to ISO 17020 standards which is used for crime scene investigations.

As an opinion-based process, there can be occasions of difference of opinion between experts and false or missed identifications. ISO 17025 standard allows the laboratory to plan and implement actions to address risks and opportunities. It is the responsibility of the laboratory to decide which risks and opportunities need to be addressed. The uncertainty of measurement shall be periodically reviewed using data from dip sampling, competency and proficiency testing. In addition, to assess the quality of results, documentation procedures shall be in place for verification that will manage the process of checking critical findings for fingerprint examinations. Critical findings refer to an outcome that has a significant impact on the conclusion reached and the interpretation and opinion provided, an outcome that cannot be repeated or checked in the absence of the exhibit or sample and/or, an outcome that could be interpreted differently.

An independent examination conducted by an expert who has no prior knowledge of the findings of another, any information on the previous examination or any further case information is known as blind verification. Blind verifications form part of the risk management approach to mitigate risks associated with cognitive bias. In situations where there is a difference of opinion, the bureau shall have a policy in place to deal with this. However, the codes are not prescriptive on this procedure. When “shall” is used throughout the standard, this infers that it is a requirement, not a recommendation.

Section 7.5 of the standards relates to “technical records”. It is stated that technical records for each laboratory activity contains the results. All issued reports shall be retained as technical records and can be kept as hard copies or electronically. As the standard is for laboratories the guidance is broad to meet all types of laboratories.

For the reports, there are common requirements which can be met by all laboratories.

These requirements include

* Name of laboratory
* Location of laboratory
* Unique identification number
* Identification of the method used.
* Date of examination and date of report created.
* Statement to the effect that the report only relates to the items tested.
* The results
* Identification of person authorising the report

As fingerprint examinations are opinion-based, the standard requires the opinions and interpretations to be expressed within reports, and if there have been any direct communications this must also be recorded. If there are any amendments to the report, this must be clearly identified and the reason for change must be included. UKAS (2019) provides guidance on dealing with opinions and interpretations statements relating to test data. It states opinions and interpretations must be traceable to the raw data and include all personnel involved in the process. Furthermore, any data that is not reported, but is a product of the testing and opinions and interpretations, must still be available and controlled so that it can be retrieved if needed.

The broadness of the ISO standard allows for fingerprint bureaux to interpret themselves how to meet this standard i.e., how to record the examination. The Codes of Practice and Conduct for Fingerprint Examinations (2017) was produced to provide further guidance for fingerprint experts, then amended in 2020. However, this is not the only guidance on what could be documented during examinations. Section 1.3.4 of this chapter critically analyses the guidance available to fingerprint examiners, which are dated pre and post accreditation official deadline.

The Forensic Science Regulator oversees the quality of forensic science within the UK and to do so, there are many groups including leading professionals to aid with decisions. One of the groups is “Fingerprint Quality Specialist Standards Group (FQSSG)” and the minutes of the meeting are publicly published. In October 2017, it was stated that none of the UK fingerprint bureaux held accreditation as of this date. UKAS arranged assessment slots with all UK fingerprint bureaux as this could potentially provide mini deadlines for individual bureaux and time management for UKAS to visit all bureaux to ensure the official deadline is met.

However, in early 2018, ahead of the official October deadline, members of the group heard that since the previous meeting, 5 accreditation visits to UK police force fingerprint bureaux had been completed. Of the 5 visits, only one bureau achieved accreditation, and one further offered accreditation subject to a follow-up visit.

The low success rate was concerning, and the readiness pre-assessment checks suggested that many others were not yet ready for assessment. The NPCC were determining how they can assist fingerprint bureaux in obtaining accreditation, but it was anticipated that only 13 would gain accreditation by the October deadline. It was suggested that a ‘buddy’ system might be implemented, whereby those bureaux which had already received accreditation share lessons-learned with others. Concerns were raised that this may cause resourcing issues for those who had already obtained accreditation.

One issue noted in fingerprint bureaux was a general lack of producing contemporaneous notes, which was felt to impact service delivery by staff (Tully, 2018). As it may be expected due to Bunter (2016) account of fingerprint examinations prior to October 2018 and the varied suggestions for documentary procedures in literature as previously described. To encourage more notetaking, it was suggested that digital systems could be implemented to make the process easier, which the Transforming Forensics programme has been supportive of. UKAS were not prescriptive in the method for note taking and were open to different practices to achieve the same goal (Tully, 2018). Therefore, it is open to interpretation of the bureau managers of the documentary procedures put into place to abide by ISO accreditation. It was suggested that a working group could be created to assess the practices of taking contemporaneous notes, this relates to the research aims of this PhD therefore, the outcomes can be provided to the Forensic Science Regulator to help create a standardised documentation strategy. However, there appears to be no announcements on such working group since the meeting.

At the following meeting, another 10 accreditation visits had been conducted. Of these, only 3 bureaux resulted in offers for accreditation; one granted and the other two in progress depending on actions. There were another 6 visits scheduled before October 2018 and another 5 in October and November 2018. UKAS stated that there did not appear to be any improvements in the number being accredited and recording keeping/notes were raised as an issue again (Tully, 2018).

The struggle was proven when only three fingerprint bureaux gained accreditation by the initial deadline. Bureaux were given an extension; all must have gained accreditation by March 2019. It was not explicit what the consequences were for not gaining the accreditation by the new date however, the FSR continued to provide support to bureaux working towards the accreditation. This support is not detailed in the publications. By the end of December 2018, another 6 bureaux and collaborations were recommended for accreditation and were expected to meet the March deadline. In addition to this, UKAS had assessments pre-booked with 10 more bureaux and collaborations between December 2018 and March 2019. It was estimated that 80% of fingerprint capacity for England and Wales will be accredited (Tully, 2019). As of February 2019, there were 19 accredited bureaux in the UK, 7 working through actions and two further first visits required (Tully, 2019).

As the updates have shown, there were issues in achieving accreditation, one of which was the note-taking aspect. Not necessarily the varied documentary suggestions in literature and policies as seen in section 1.3.4 but this could have also been hugely impacted by financial constraints and limited resources that the police and FSR were aware of (McCartney & Amoako, 2019). The cost of a UKAS visit is at least £8000 and can go up to £30,000. This is a substantial amount, especially for smaller organisations with low-income it could potentially lead to closure. McCartney & Amoako (2019) states that the costs of accreditation are substantial, and the benefits look inadequate. It continues to be the case that the Criminal Procedure Rules do not require forensic evidence to come from an accredited body. Also, the FSR has no powers to penalise non-accredited organisations. Two bureaux were interviewed by McCartney & Amoako (2019) to determine opinions on accreditation. The high costs with little evidence of benefits frustrated one interviewee whereas, the second can see a tangible improvement.

### 1.3.4 - Guidance for documentation of fingerprint examinations

There are many suggestions for information to be documented during fingerprint examinations within published scientific and grey literature, which are summarised in table 1 and then discussed chronologically. For this research, this summary of guidance relating to information to be recorded during fingerprint examinations will be referred to as ACE-V checklist. The dissimilarity in the guidance offered by these sources is surprising. To some extent, differences were attributed to those resulting from academic research, which may have been based upon external guidance from those aligned with operational work, and the development of regulations for working practice. It is important to highlight that some of these publications existed prior to the announcement of the ISO 17025 accreditation requirement for fingerprint examinations.

As highlighted by Tully (2018) the production of contemporaneous notes impacted achieving accreditation by the October 2018 deadline, one reason for this could be the varied documentary suggestions and limited guidance on what to include in the documentation.

The Crown Prosecution Service: Guidance for Experts instructs experts on how to prepare expert evidence and explains the disclosure obligations for expert witnesses instructed by the Prosecution Team (Crown Prosecution Service, 2010), including the three R’s; retain, record and reveal which have been previously explained in this paper. According to this guidance, documentation should be produced contemporaneously, whenever practicable. As this guidance is not specific to fingerprints it suggests best practice and the minimum requirements of what to record, for any discipline. Generic information such as the date, names of experts involved in the examinations and the materials used in the examination. This information is required to locate the original examination and replicate the examination using the same materials. This information would be necessary if there was a dispute or in cases where the case goes to court and it is a different expert re-examining the mark, as described previously. There must be a record kept of all the verbal communications against parties involved for continuity and integrity purposes. This may include meeting notes, telephone conversations, emails and other electronic transmissions and witness accounts (Crown Prosecution Service, 2010). It is important to note that certain information recorded during these verbal communications can be subsequently used later, for example the agreements and disagreements between experts. In terms of note taking, it mentions the requirement for experts to produce ‘sufficient detailed notes’, which can be interpreted by another expert witness with similar expertise to follow the nature of the work undertaken and understand the conclusions drawn (Crown Prosecution Service, 2010).

This statement has also been seen in the Forensic Science Regulator’s Codes of Practice and Conduct (The Forensic Science Regulator, 2017), which provides guidance that is specific to fingermark examinations. A perceived issue with the term ‘sufficient’ is that it is subjective in nature, meaning that there are likely to be differences in the quantity and content of the notes that are produced, which may or may not meet with another examiner’s opinion of sufficiency.

In 2011, a reportedly transparent approach to minutiae annotation was proposed by Langenburg and Champod, referred to as the ‘GYRO’ system. This system was also recommended by (Bunter, 2016), whereby the expert assigns a confidence level to each characteristic observed. By recording this information at the time of examination, it prevented retrospective note taking for court and its associated issues.

This system did not suggest any further content to the documentation of examinations given that the focus of the research referred to minutiae annotation as part of the examination process. It is unknown whether this is used within operational practice or not in England and Wales.

There is evidence from the work of Ulery *et al* to suggest that the number of minutiae recorded may affect an examiners decision of sufficiency. In this study, for an identification, one third of the participant group (n = 170) recorded eight or fewer minutiae, whereas others had a minimum count of 14. When outliers were removed, the results showed identifications were made with as few as six corresponding minutiae. The study showed that there were substantial differences in examiner annotations. In the author’s experience of reviewing annotations of ridge detail, it is not possible to ascertain whether minutiae annotations are exhaustive of all minutiae available (in the opinion of the examiner). If the number of minutiae affects decisions made by examiners, it could be logical to assume that the decision to record minutiae (or not) could equally contribute to differences in reporting outcomes, consistent with the work of Dror, where the written report itself if believed to influence the work that is produced (Dror, 2012). The research team are also interested in the impact of annotating minutiae on the identification process, such as the effects of intra and inter examiner variations on an AFIS comparator list and, how annotations contribute to a ‘sufficient auditable trail’.

Of the 86 recommendations that were made following the Fingerprint Inquiry, a number of these related to notetaking. Recommendations 50 to 53 provided guidance on what, when and by whom they should be taken. For example, recommendation 50 stated that “examiners should always take notes when they are examining marks that they consider to be complex”.

A perceived issue with this recommendation is the difficulty with the subjective definition of the term ‘complex’, which appears to have no universally accepted definition, although the term is recognised by the Forensic Science Regulator (2017). Complexity may refer to fingermarks affected by factors that cause a substantive impact upon the visibility of ridge detail, such as movement or force, which logically may to encounter increased amounts of differing judgements. An explanation for such differences could be provided with effective documentary practice. This has the capability to be used as an effective learning aid, to help examiners to develop their professional practice and to assist with the prevention of missed identification opportunities. According to Ulery, “procedures for detailed documentation of the features used in the analysis or comparison decisions could be used to assist in arbitrating inter-examiner disagreements at the feature level” (Ulery *et al,* 2012).

The later published PCAST report highlighted the need to better understand inter- and intra-examiner variability (Executive Office of the President President’s Council of Advisors on Science and Technology, 2016) providing further support for effective notetaking.

*Table 1 - A summary of guidance relating to information to be recorded during fingerprint examinations according to published sources of literature.*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Sources of literature** | | | | | | |
| **Information to be recorded** | **CPS** | **Champod and Langenburg** | **The Fingerprint Inquiry** | **SWGFAST** | **ENFSI** | **Bunter** | **Forensic Science Regulator** | |
| **Year of Publication** | **2010** | **2011** | **2011** | **2013** | **2015** | **2016** | **2020** | |
| Name of the experts | X |  |  |  |  |  | X | |
| Date/URN | X |  |  |  |  |  | X | |
| Materials used | X |  |  |  |  |  | X | |
| Any communications | X |  | X |  |  |  | X | |
| Assigning confidence levels |  | X |  |  |  | X |  | |
| Assessment of mark quality |  |  | X |  | X |  |  | |
| Deltas and cores present |  |  |  |  | X |  |  | |
| Orientation of the mark |  |  |  | X |  |  |  | |
| What the matrix is |  |  |  | X |  |  |  | |
| Mark origin |  |  |  | X | X |  |  | |
| Presence of level 1 |  |  |  | X | X |  |  | |
| Any signs of distortion |  |  | X | X | X |  |  | |
| Factors affecting the mark, e.g., superimposition, substrate |  |  |  | X | X |  |  | |
| Annotate the images |  | X |  |  |  | X |  | |
| Minutiae at analysis |  |  | X | X | X |  |  | |
| Minutiae at comparison |  |  | X | X | X |  |  | |
| Revisions at comparison |  |  | X |  |  |  |  | |
| Third level detail |  |  | X | X | X |  |  | |
| Explanation of mark differences |  |  | X |  |  |  |  | |
| Reasons for conclusions | X |  | X | X |  |  | X | |
| Sufficient detailed notes | X |  |  |  |  |  | X | |
| Records of examination |  |  |  |  |  |  | X | |
| Sequence of recording contemporaneous notes |  |  |  |  |  |  | X | |
| Reporting outcomes |  |  |  |  |  |  | X | |

In The Codes of Practice and Conduct for Fingerprint Comparison (2017) a report template has been included for documenting complex marks. It includes recording generic information such as.

* Reference number
* Name of panel lead
* Organisation details
* Date received.
* Background information provided by requestor.
* Information provided to panel members.

This information is a record of the case and individuals involved for future investigative purposes. Such as, the case going to court, it is not necessarily the same expert from the original examination who presents the fingerprint evidence in court.

In the Analysis section, the form includes multiple choice questions and free text boxes. It covers some aspects which have been previously discussed such as Q1 - does the mark/s have sufficient detail for comparison, Q2 - is there sufficient detail to exclude, Q3 - is the mark/s complex, Q4 - are there any signs of distortion and Q4 - does this impact the mark. The questions cover what an examiner would mentally be observing during an examination but are broad. Each question is answered with “yes, no or n/a”. There is no capability of explaining each answer, which can become problematic if there is a difference of opinion, as seen in R-v-Smith (2011).

Annotations could be incorporated into question one and two, to allow experts to explain what makes the mark sufficient detail for comparison or exclusion. Bunter (2016) explains the different ways to annotate an image. In addition, experts could potentially assign confidence levels to the features observed to aid with their conclusion (Langenburg, 2012).

As there is no clear definition of a complex mark, this could cause issues for question three. What one expert or bureau believes to be complex, others may not. Due to these differences, the answer should be accompanied with an explanation of why the mark/s have been deemed as complex. The Fingerprint Inquiry (2011) states within the Analysis stage, the expert must assess the quality of the mark. If the expert believes the mark is complex, then this should be recorded and follow the separate process for complex marks as outlined in recommendation 42. During this process, there should be three experts involved, and each must take notes independently at each stage of the ACE process. Recommendation 52 highlights the necessary information to record.

For the Analysis stage, it is recommended the assessment of quality and any signs of distortion are noted. Followed by the characteristics identified including their type and the sequence of them. It is not prescriptive in the manner of which this information should be recorded. This relates to question 4. The signs of distortion assumedly cover substrate interference, movement, development technique, matrix and pressure as these have been drawn out of literature. However, the report template does not explicitly include these factors. Again, there is no capability of recording what is causing the distortion. As the literature agrees as to factors which affect the mark, a suggestion could be tick boxes with each factor and experts tick as appropriate. Using this approach, it shows that the expert has taken into consideration all possible factors and only tick which they believe is causing distortion (The Fingerprint Inquiry, 2011).

A free text box was included to comment on where the panel concur and where they differ. This can be used to identify difference of opinion if it occurs (The Codes of Practice and Conduct, 2017).

The Fingerprint Inquiry (2011) recommends the conclusions reached by each expert cannot be disclosed until all three experts have finished their examination. On completion of the individual examinations, all three experts should meet and review the basis of their conclusions. Despite the outcome, whether it is an identification or not, then the reasons for this conclusion are discussed. If there are differences of opinion, then this should be discussed to determine whether the conclusion is reliable. To decide the outcome, the experts could be using the consensus. This meeting should be kept on record.

Recommendation 51 of the Fingerprint Inquiry (Campbell, 2011) stated that “notes should be taken in any case in which a fresh comparison is made in response to a request from the Crown for a report”. Although not necessarily linked, this recommendation is satisfied by the streamlined forensic reporting process, introduced in 2013, as previously mentioned. There is evidence that additional examinations may lead to inter or intra examiner variability, as highlighted in the work of Ulery *et al* (2012), where decisions were changed to or from inconclusive or value for exclusion by examiners repeating an analysis of 25 image pairs of fingermarks. Given that differences of opinion can exist with repeat comparisons (within and between examiners), effective note taking practice would be essential to establish the basis for any differences in the conclusions drawn. In addition, notes made at a later stage may also be written with the knowledge of other forensic evidence, providing further opportunity for contextual bias. Finally, a retrospective process poses difficulties for any expert who changes their mind, of which there is evidence to suggest can happen given that the process is cognitive (Ulery *et al,* 2012). Dror (2012) explains that there are cognitive effects that result from producing a written report.

In Dror’s opinion, the written report both reflects and influences the work, which are important for obtaining high quality forensic results. His research refers to the work of Klein (1999) and their research regarding the effects of writing-to-learn. It is unknown whether jurors are aware that the notes produced for court may be retrospective and whether it would influence the opinion of the jurors if it was known.

Recommendation 52 of the Fingerprint Inquiry (Campbell, 2011) related to the content of fingerprint examination documentation. It stated that notes should be taken at each stage of the ACE-V process and individual notes by each expert involved in the examination, as shown in table 1. The recommendation relating to revisions of the initial analysis at the comparison stage did not specify whether this referred to a verifying examiner coming to a different analysis than the initial examiner, a court reporting officer coming to a different analysis than the analyst(s) who made the initial identification, or an analyst at any stage making revisions to their own previously recorded analysis.

Whereas recommendations 50 and 51 related to note taking for complex marks or additional examinations, recommendation 53 suggested that notetaking should become general practice for all fingerprint comparison work and that by recording this information, an expert would be explaining how they reached their conclusion for the fingermark. Although guidance for the content of contemporaneous notes is provided in recommendation 52, there are areas highlighted in alternative literature that are not considered here, as demonstrated in table 1.

A significant problem raised by expert witnesses in The Fingerprint Inquiry (Campbell, 2011) was the length of time in which it could take to produce contemporaneous notes for each case. One expert stated that “it would not be practical to take notes for each case”.  According to expert witnesses involved in the Inquiry, note-taking would have a major impact on productivity and efficiency. However, it was agreed that the quantity of the contemporaneous notes made would be influenced by the perceived quality of the fingermark under examination; the poorer the quality of the mark, the more detailed notes should be provided. The author has witnessed variations in working practice with the creation of working notes for fingerprint examinations following ISO 17025. According to the minutes of Fingerprint Quality Standards Specialist Group published in 2018, the UK Accreditation Service was not prescriptive in the methods required to produce contemporaneous notes and were open to different practices to achieve the same goal (The Forensic Science Regulator, 2018). This is perhaps useful given variations that exist in practice and the availability and usage of technology for examination. However, it does little to direct examiners to meet with the needs of accreditation and good scientific practice.

Variations in documentation practice, including the contents of documentation are also yet to be independently compared and evaluated in this field. A study of the effectiveness of documentation strategies is currently part of further research being undertaken by the author.

In 2013, the Scientific Working Group on Friction Ridge Analysis, Study and Technology (SWGFAST, 2013) published guidelines recommending the linear analysis of unknown marks. In agreement with the European Network of Forensic Science Institute (ENFSI, 2015), the group suggested that the level of detail recorded should depend on the quality and quantity of the friction ridge detail, but that as a minimum with many similarities to the ENFSI (2015) recommendations. Also, if known to the expert, the substrate, development technique and preservation method should be recorded and additional factors such as matrix, deposition pressure, movement and level 3 detail, could be recorded if known (SWGFAST, 2013). In these guidelines, the features identified during the Analysis would be relied upon during the Comparison. Very little was suggested to record during the comparison phase when an identification was made, only the name and date of birth of the individual identified and a reference number for the case. However, if the known mark was deemed as insufficient for comparison, then they recommended that this is explained in the notes. It is not clear why the reasons are recorded for one category and not the other. According to SWGFAST (2012), within the Evaluation phase, the expert should record the friction ridge impression examined, known prints used to reach the conclusion, the specific anatomical source and initials and signature of the expert.

There are some similarities between the Best Practice Manual for Fingerprint Examination, published by The European Network of Forensic Science Institutes (ENFSI, 2015) and The Fingerprint Inquiry recommendation 52. The manual also encouraged reference to any distortion or superimposition present, any effects due to development technique and because of the above, whether ridge flow or individual characteristics could be identified, which could be recorded during the analysis of a mark and used towards the comparison, dependent on the case. The manual requests explanations on experts’ thoughts during the original examination. However, it also suggests providing more information regarding the number and type of minutiae present. This is understandable due to the contribution that minutiae make to fingermark identification. The suggestion to record any identified minutiae was consistent between these documents, and further supported by SWGFAST (2013). The author has witnessed differences between fingerprint bureaux in terms of whether they choose to record minutiae during an examination and a variety of approaches to the recording of minutiae, which supports the findings of the Fingerprint Quality Standards Specialist Group in terms of the openness of UKAS to different approaches to note taking (The Forensic Science Regulator, 2018).

UK fingerprint bureaux were instructed by the Forensic Science Regulator 2015 to gain ISO accreditation by October 2018. In part, refers to the production of contemporaneous notes. It is intended that all Bureaux will meet minimum standards of documentation through validation of their processes to ensure that they conform to the standard. Sources of academic and grey literature could be used to direct fingerprint examiners towards the necessary criteria for documentation and which could contribute to effective practice. Scientific working groups such as SWGFAST and ENFSI, listed specific criteria for inclusion, others, including The Forensic Science Regulator, suggested production of notes sufficiently detailed for an auditable trail. The term ‘sufficient’ is subjective and open to interpretation; what one examiner may feel is sufficient, another may not. This leaves documentation open to inconsistencies in approach and quality from the outset. The differences of documentary suggestions between the sources of literature allows each Bureau to develop its own methodology rather than adhering to a prescribed approach and to date there is still considerable variation, which makes evidence-based approaches to documentation difficult. An “auditable trail” should be a record of the work undertaken including justifications and authorisations (Forensic Science Regulator, 2020). Accreditation is an expectation in other countries also for laboratories carrying out tests and therefore there are similarities in requirements to achieve accreditation. Auditable trails are expected in accredited laboratories despite the country. Illinois v Safford (2009) highlights the lack of fingerprint documentation by the expert. The expert explains that the notes do not explain how decisions are made, only if there is an identification or not. When questioned, the expert could discuss broadly the level of detail i.e., 1, 2 and 3 were present however, could not give specifics on which details were present and used to make the final decision. It must be noted that these documentation processes are accepted by an accredited laboratory and demonstrate what could be an example of an auditable trail. It is argued that limited or no notes cannot be cross-examined and leads to the suggestion of infallibility. If a comparison had been performed by a mechanical or electronical device, it’s expected that there is proof the equipment functions correctly. A comparison by a human should be no different. In order to ensure the expert carrying out the comparison is trustworthy and reliable, how the expert reached their decision should be demonstrated. Taking this into consideration, the fingerprint evidence in this case was rejected and the court remanded a new trial.

The benefits of documentation have been discussed by Langenburg (2012). From this research, there were seven recommendations for the ACE-V process. Each were discussed in detail, however the recommendation most relatable to this paper is recommendation 3. The recommendation states “documentation must be done in each case to the extent that is appropriate for the complexity of the case and to the extent that it is sufficiently transparent how the analyst arrived at his/her conclusion”.

The benefits of implementing documentation procedures are explained by Langenburg (2012). Clear and concise procedures for when to document and to what extent. As stated by SWGFAST (2012), the level of documentation will be dependent on the complexity of the case. Documentation should be sufficiently detailed for another examiner to know what tests were conducted, the results of the tests and the reasons for the conclusions. It conveys the validity of the process. Understanding how an expert has reached the conclusion can be more important than the conclusion itself. However, Langenburg does state that this does not mean every thought and observation is required to meet the standard.

As cases in court can last for over a year, the original written analysis could refresh the memory of the expert, or to inform others of the details of the analysis.

Preparing a detailed written analysis prior to the comparison stages promotes objectivity.   
To achieve ISO accreditation, the contemporaneous notes can be recorded electronically, or paper based which provides flexibility for the fingerprint bureaux, as all bureaux will not have the same equipment or financial budget. It was suggested by Ashbaugh (1999) that tape recorders could be used to dictate their analysis.

It must be noted that the literature search and practical work for this thesis occurred between 2016 and 2022 therefore, any necessary changes within grey literature or practices, made by governing bodies or police forces are included from here onwards. There are some minor differences between the 2017 and 2020 versions of Codes of Practice and Conduct. In the most recent version, (2020) it states all records shall as a minimum demonstrate the examination sequence. This could be interpreted as clearly demonstrating the thought process within each stage of the ACE process. For example, explanations of the information gathered in the Analysis stage which would be useful for the Comparison stage and ending with an outcome at Analysis. Within the Comparison stage, reporting the presence of a coincident sequence, either by words or illustrations. Any additional information which may have been used to aid with the decision. Followed by the Evaluation stage, reporting the outcome and identifier if necessary. How to demonstrate the examination process appears to be to the discretion of the fingerprint bureaux managers who are devising the documentation procedures as there are no specific guidance in the Codes of Practice and Conduct or from UKAS as the accreditation body. However, there is a slight amendment to the minimum requirements. There is the addition of recording the examination, the sequence of recording contemporaneous notes and including the reporting outcome.

This can be seen in Section 8.1.4 (The Forensic Science Regulator, 2020). It is likely that the reporting outcomes are for Evaluation stage, rather than the Analysis outcome as there is no definitions provided for the Analysis outcomes. Although, this would be a crucial aspect of the examination, as it is determining the quality of the unknown mark from a scene to compare against known prints. Furthermore, in section 12 it discusses reporting outcomes such as inconclusive, which confirms it is the Evaluation outcomes.

In both 2017 and 2020 versions of Codes of Practice and Conduct, within section 12 it states, when reporting an inconclusive outcome, there should be an explanation as to why the outcome is inconclusive. Also, where relevant, the notes should include the “mark status” for example, searched with a negative result and remains unidentified.

In more recent news, The Transforming Forensics programme, hosted by the Police Digital Service (PDS), have been working on the development of a new national, digital fingerprint capability for policing. The expected completion date for this was December 2022 (Police Digital Service, 2022). The capability is hosted on the PDS Xchange, which is a secure national cloud-based platform. In April 2022, PDS Xchange successfully integrated with Ident1 (the national automated fingerprint system in the UK). The platform provides tools to send fingerprint and crime scene images in real-time to the bureau, compare fingerprints using Ident1 and allow for the recording of analyses of data, it is said to be fully compliant with ISO 17025 standards. Full deployment of this platform was expected early March 2023 to the East Midlands Special Operations Unit which includes: Derbyshire, Lincolnshire, Leicestershire, Northamptonshire and Nottinghamshire. Information surrounding Xchange is very limited in the public domain and additional information regarding the recording element of the platform has tried to be retrieved however, been unsuccessful due to the confidentiality of the work. For this research, it is worth noting the likelihood of a new digital platform for potential documentation procedures.Prior to this release, there were digital methods available for the documentation of fingerprint examinations such as CSIpix, FSH and Forensic Comparison Software. CSIpix and FSH are software which was used by some UK bureaux and independent companies to annotate impressions by marking up minutiae identified but there was no obligation to from the Forensic Science Regulator to use these digital methods.

In these systems, it allows for part of the documentary suggestions for fingerprint examinations to be recorded however does not have the functionality to record generic information, factors affecting the mark and level of detail which all forms part of the decision-making process.

Forensic Comparison Software similarly to the other digital methods was designed to annotate impressions by marking up minutiae with the added feature of searching against own AFIS databases’ and the production of charts and reports for court or case file purposes. There is limited information available publicly on what constitutes the reports. It appears that the focus is marking up minutiae and identifying corresponding characteristics between the unknown mark and known mark (FCS, 2018).

Within the published literature, there is a strong claim that minutiae count is the driving factor behind suitable for comparison decisions (Ulery *et al,* 2013, Ulery *et al,* (2014) and Langenburg (2012) which signifies the inclusion on marking up minutiae in digital platforms. However, as highlighted by Eldridge *et al* (2020) it is not the only factor important within decision making. This is also supported by the documentary suggestions seen in the ACE-V checklist, table 1.

As UKAS were not prescriptive in the method for note taking and were open to different practices to achieve the same goal (Tully, 2018), these digital methods could be approaches witnessed in this research or potentially be adapted using the end results.

#### 1.3.4.1 – Guidance for documentation outside of fingerprint examinations

To aid with determining effective documentation strategies for fingerprint examinations, it is beneficial to understand the documentation practices in other forensic science disciplines to potentially learn from other areas. In addition, the guidance for best practice in other disciplines outside of a forensic field to learn the purpose, the level of detail and importance of documentation.

Codes of Practice and Conduct for Bloodstain Pattern Analysis, Issue 2 (2020) outlines the reporting of results in alignment to ISO 17025, the same standards in which fingerprint experts adhere to. The Codes recommend the following, when necessary; case information, limitations of the blood pattern from digital media, relevant information such as the location where observations are recorded and measurements, such as areas of origin, room size, heights of bloodstains and distribution of a bloodstain pattern. Descriptors of the reported stains and stain patterns which should be the terminology outlined in ASB technical report 033 as good practice. The results of testing conducted to identify blood, the results of any chemical enhancement of bloodstains and conclusions and interpretations. The above can provide a basis upon which the opinion has been made. An immediate difference within fingerprint examinations and blood pattern analysis reports is the inclusion of limitations, this could be a consideration for fingerprint experts to potentially address the limitations of the evidence they have been submitted for comparison. For example, outlining the factors that are impacting the quality of the finger/palm mark which could be having a negative impact on the ability to identifying individual characteristics and conducting a comparison. These factors impacting a finger/palm mark are suggested by SWGFAST (2013) and ENFSI (2015) and arguably contribute to experts’ decision making and therefore could be a part of the “reasons for conclusions” as highlighted by the CPS (2010) and the Forensic Science Regulator (2020).

In 2020, NIST published a report of the expert working group for human factors in handwriting examination. It is stated that after a forensic document examiner has completed an examination and interpretation of evidence, the results must be communicated either in a written report or by testimony in a judicial forum. The report, in part, this discusses the value of the forensic report and the contents of the report. Although this is targeted at US handwriting experts, other forensic disciplines and countries can learn from this. The report becomes a record of the methods, examinations, limitations, and conclusions regarding the submitted evidence. The report could point the investigation in a particular direction, inculpate or exculpate a suspect/defendant, or be neutral in its impact. The report allows civil and criminal litigators to assess the evidentiary value of the examination results and may help guide the disposition of the case. For those reasons, the report must be accurate, clear, and objective. If not in the report, all other relevant information should be documented in the case record and available for the litigants’ review. It is stated that the report may serve as a stand-alone document during court proceedings without expert testimonial which highlights the level of detail that should be within the report i.e., details of the analysis and comparison including the conclusions and limitations. The act of writing the report can have cognitive effects on the writer (NIST, 2020). A report which is accurate, clear and objective also adheres to the ISO 17025 standards. When employed by an accredited laboratory, a written or electronic report is required each time an examination is conducted. ILAC G-19 provides some flexibility for how the required information is conveyed (NIST, 2020). A study in 2013, by Siegal *et al* surveyed approximately 400 forensic science laboratories report from 38 funded crime laboratories. The report contents were compared to report recommendations from 10 forensic science organisations and working groups. The recommended report contents included; demographics, request for examination, inventory of evidence, executive summary, methods and materials, procedures (step by step), results, discussion, limitations and data.

Within the health care profession, the General Medical Council (GMC) role is to protect patients and improve medical education and practice across the UK. As part of the role, the council set the standards which doctors must follow in part this refers to the recording of work carried out by doctors. Ensuring it is clear, accurate, and legible (GMC, 2013). General Medical Council (2013) outlines the requirement for record keeping for doctors. Notes must be taken at the same time or as soon as possible afterwards, similarly to what is required within the forensic profession. To ensure records are clear and accurate, doctors should include any relevant findings, decisions made, and actions agreed, names of who is making the decisions, any information that has been provided to patients, any drugs prescribed, or treatment, name, date and time of the record being made. Clinical records include a wide variety of documents generated on, or on behalf of, all the health professionals involved in patient care. This includes, but not limited to, handwritten or computerised notes, all personal correspondence and laboratory results (Medical Protection, 2013). The purpose of keeping records in a health profession is to provide continuity of care but also medical records are used for administrative and evidence of decision making, available for patients to request, supporting improvements in clinical effectiveness and a factual base for responding to complainants. A doctors’ defence will depend largely on the records. If information is missing, inaccurate or indecipherable then complainant cases could be lost (Medical Protection, 2013). It is well-known within the health profession, “if its not recorded, it did not happen”. An entry in Nursing Times (2013) discusses what makes “good” records. In some cases, nurses will record observations and no conclusions, or conclusions with no observations. For example, a conclusion could be a patient slept well, however there are no observations/evidence to support why the patient slept well. In addition to the observations and the conclusions, actions from this should also be recorded (Nursing Times, 2013).

In other areas i.e., aviation, there are strict rules on how far airliners must be apart. There are systems in place to help separate aircrafts, it scans for nearby aircrafts and alerts the pilot if close by. Despite such systems in place, there can be instances where the minimum separation is not maintained. When this happens, pilots and controllers can file a report for the incident to be investigated. In the UK, an aviation near miss report is called an airprox. Since reporting, there has been a decline in the number of near misses which shows its benefits of recording these near-misses (Civil Aviation Authority, 2023). The purpose of an airprox is to improve safety, it should be reported up to 72 hours after the incident. Macrae (2016) reports what health care could potentially learn from aviation and states some of the most urgent challenges in patient safety concern how to identify, understand, and act on the early signs of emerging problems before those risks cause harm to patients.

Health professionals could learn from the approach within aviation, identifying and investigating from minor errors and near-miss incidents, for example when a doctor records the wrong dose in a prescription but a pharmacist notices and corrects it. These brief encounters with risk are a normal part of all organized human activity and provide valuable opportunities to improve safety. Similarly, in the forensic field, lessons could be learnt from aviation procedures. A near-miss within fingerprint comparisons could be the complex impressions whereby two experts differ on their decision, one may report insufficient for comparison yet one may be confident in characteristics present. There is the potential of a missed identification here. All airprox will be investigated, however within forensics, it would be near impossible to verify all outcomes and decisions but there could be processes in place to check routinely. If notes are made for every case, and routine dip sampling of case notes was mandatory then potential missed identifications could be highlighted. The reasons for this could be investigate and the appropriate training can be given to prevent missed identifications happening again.

## 1.4 - Aims and objectives

The aim of this research is to provide evidence-based recommendations for the content of contemporaneous notes that are recorded by identification experts as part of fingerprint examination in casework. To achieve this aim there were three studies within this research. Study one established pre-ISO 17025 accreditation documentation strategies and an online structured software as a means of documenting fingerprint examinations. Study two established a selection of post ISO 17025 accreditation documentation strategies and the same online software. Study three assessed the effectiveness of these pre and post accreditation documentation strategies and the results were used to recommend effective documentation strategies, both documentary approach and content within the notes.

The originality of this research is the first of its kind to highlight fingerprint documentation strategies pre and post ISO 17025 accreditation. There have been court hearings scrutinising the documentation produced during fingerprint examinations and there is a number of literature and policies suggesting what could be included in the documentation but no literature to imply how much of the suggestions are included in the actual documentation produced. Due to the method design, inter and intra examiner variance was investigated and contributed to existing research (Ulery *et al,* 2011, Ulery *et al,* 2012, Tangen, *et al,* 2011, Neumann *et al*, 2013, Pacheco *et al,* 2014*,* Eldridge *et al,* 2021). However, the contemporaneous notes were used to attempt to understand the thought processes and potential difference of opinion. In turn, this supports the importance of notetaking.

The objectives to meet this aim are

1. To perform a gap analysis on pre-ISO 17025 accreditation documentation strategies for fingerprint examinations to best practice guidance[[1]](#footnote-1) acquired from academic and grey literature.

The results from this will be used to produce an analysis form to replicate pre-ISO 17025 accreditation documentation.

1. To discover a structured online software which prompts experts to consider and record best practice guidance as another means of documenting fingerprint examinations.
2. To perform a white box study[[2]](#footnote-2) involving the semi-structured pre-ISO 17025 accredited analysis form from objective one and the structured online software from objective two. Using the data generated, use descriptive statistics to examine trends in expert decision-making notes for fingerprint analyses using these documentary strategies and compare the data to best practice guidance.
3. Using the data generated from objective three, statistically assess the relationship between the extent of expert decision-making notes and the quality of the friction ridge impression using chi-square tests for association.
4. To review the ISO 17025 accredited documentation strategies for the participating UK fingerprint bureaux and demonstrate the new approaches taken to record contemporaneous notes.
5. To perform a white box study involving the participating UK fingerprint bureaux ISO 17025 accredited documentation strategies and the same structured online software from study one, objective two. Using the data generated, use descriptive statistics to examine trends in expert decision-making notes for fingerprint comparisons using these documentary strategies and compare the data to best practice guidance.
6. Using the data generated from objective 6, statistically assess the relationship between the extent of expert decision-making notes and the quality of the friction ridge impression using chi-square tests for association.
7. To conduct a matching exercise with a group of experts and novices to assess the effectiveness of the documentation strategies from pre- and post ISO 17025 accreditation.
8. To statistically assess the relationship between obtaining a correct outcome[[3]](#footnote-3) and expertise group.
9. To statistically assess the relationship between obtaining a correct outcome and documentation strategy.
10. To statistically assess the relationship between obtaining a correct outcome and grade of the mark.
11. To assess the inter and intra examiner variance of reporting outcomes and use the contemporaneous notes to understand the thought processes of variances observed.

# Chapter 2: Methodologies

In literature, documentary suggestions for fingerprint examination case notes include generic information of the case, factors impacting the quality of the impressions, the level of detail observed and annotations on the impressions and reporting outcomes (CPS, 2010; Champod and Langenburg, 2011; SWGFAST, 2013; ENFSI, 2015; The Forensic Science Regulator, 2020). However, there is no literature that demonstrates the actual documentation procedures within the United Kingdom. Therefore, in this research, it was necessary to establish pre-accreditation and post-accreditation documentation strategies within the UK. The pre- and post-accreditation documentation produced was compared to an online software which includes some documentary suggestions such as factors impacting the quality of the impressions and level of detail observed (as previously mentioned in the aims and objectives, this is the first and second study of the research). The effectiveness of each of these documentation strategies (pre, post and online software) were investigated to provide evidence-based research on recommendations for effective documentation (this is the third and final study of the research).

The study design and background reading has influenced the thesis structure. As the literature supporting this research is relevant for all three studies and therefore, to avoid repetition, the researcher chose to have a larger introduction covering the background of friction ridge detail, examinations and the guidance on documenting fingerprint examinations. Followed by a standalone methodology, results and discussion, conclusion and further work chapter. As opposed to a traditional thesis structure of every study being its own chapter. It is important to read chapter 1 before continuing onto the studies regardless of which element of this research you are most interested in. For an overview of each study, refer to chapter 5 which highlights the key findings from each study and recommendations for practitioners.

The method for each study is outlined in this chapter, the results and discussion for each study is in chapter 3 using the same corresponding sub chapter numbers. For example, the method for study one - establishing pre-accreditation documentation strategies is under sub chapter 2.2, the results and discussion to this study is under sub chapter 3.3.

This methodology chapter outlines the selection process for the eight impressions used in both study one and study two and the comparison between the scoring system and grading system. Study one sought to establish the documentation strategy pre-accreditation in UK fingerprint bureaux with a comparison to an online digital software accessible to fingerprint experts; study two sought to establish the change in documentation strategies post-accreditation with a comparison to the same an online digital software; study three established effective documentation strategies for fingerprint examinations using strategies witnessed in study one and study two. The research involved eight friction ridge impressions which are used throughout all of the studies. Section 2.1 outlines how these impressions were selected. There are some similarities between study one and study two such as the materials used and the procedure but there are a small number of differences therefore the methods are outlined separately in section 2.2 and 2.3. Both studies were essential to establish how the fingerprint examinations were documented pre- and post-accreditation. As a result, there were a total of nine documentation strategies which were investigated to determine the effectiveness and recommend documentation strategies for fingerprint examiners, this is outlined in section 2.4. The final part of this chapter refers to the methodology for the inter and intra-examiner variance in relation to reporting outcomes. Although this was not originally in the project plan, the results from study one and study two were used to determine the variance in reporting outcomes at analysis and comparison stages of the ACE process. When a variance of opinion occurred, the documentation produced was used to understand the thought processes and when successful, this demonstrated the importance of notetaking to understand the basis for the difference of opinion.

## 2.1 – Friction ridge detail selection for the purposes of this research

The following sections explain how the friction ridge impression samples were created. This was not a part of the PhD research, but it is important to note how the samples were created and assessed to imply how the impressions replicate case work in UK fingerprint bureaux. The primary researcher was involved in both studies.

### 2.1.1 – The creation of Staffordshire’s Ground Truth Database

In a previous collaborative research project, a ground truth database was created for Staffordshire Police Fingerprint Bureau. Staffordshire University and the local police force have joint ownership of the database therefore it can be used for other projects.

The database contains approximately 900 friction ridge impressions of varied quality, on a selection of different surface types and developed using the most common development techniques within the laboratory to replicate casework examples. To ensure there were a wide range of quality marks, all impressions were examined and given an overall score, between 6 (good quality) and 15 (poor quality). As the researcher did not have fingerprint identification experience, a scoring system was created for that study which can be used by novices and still evaluate the quality of the impressions. There were six different criteria to reach the overall score of the impression. The criteria involved factors that can affect impressions and subsequently affect the quality of the impressions, these were substrate interference, pressure distortion, deposition pressure, development techniques, class characteristic and ridge detail affected. For each criterion, the impressions were scored 1, 2 or 3 in relation to how much the impression was affected. Score 1 was 0 to 1/3 of the impression affected, score 2 was 1/3 to 2/3 of the impression affected and score 3 was 2/3 to all the impression affected. For class characteristics, impressions were scored 1 or 2 which was dependent on if the class characteristic was present or not.

### 2.1.2 – A comparison between a systematic scoring system and an exemplar police grading system

CAST devised a grading system which is widely used in fingerprint research (Sears, et al., 2012). One of the participant groups designed their own grading system around the CAST’s grading system. Although, the grades are opposites i.e., Sears *et al* (2012) grade 0 relates to no evidence of mark, grade 1 relates to weak development and no ridge details. Whereas the participant group grade 1 relates to clear 1st and 2nd level detail and clearly identifiable. The research team acknowledges this difference but sees this having no impact on the results.

The definitions of each grade are below.

* Grade 1 – Clear 1st and 2nd level detail sufficient to see pattern or area of palm, clearly identifiable.
* Grade 2 – Clear 1st and 2nd level detail visible sufficient to identify, may not be able to establish pattern or area of palm, some areas not clear or distorted.
* Grade 3 - 1st and 2nd unclear, limited quantity, distortion, but sufficient for identification purposes.
* Grade 4 – poor quality 1st and 2nd level detail, quantity insufficient for identification purposes.

To assess whether the researchers’ scoring system in section 2.1.1 was like another grading system created by fingerprint experts, a blind test was conducted. Four fingerprint experts from the same bureau which use the above grading system were recruited to grade ten images from the ground truth database. These experts had never seen the images prior to this test. The experts were not told of the scores given by the researcher and were instructed to complete the assessment alone to qualify as a blind test. The researcher had already previously scored the ten images during the creation of the ground truth database. The database contains over 900 finger and palm marks and a variety of scores from 6 to 15. The database was filtered by scores and two from scores 6, 8, 10, 12 and 15 were randomly selected for the blind test. [See appendix 1 for examples of fingermarks representing each of these scores.](#Appendix_1) Each participant was asked to record what grade they would give the image based on their current grading system. The primary researcher collated the data to determine if the two systems married up and if so, what scores related to what grade.

## 2.2 – The establishment of pre-accreditation documentation strategies for Fingerprint Analysis; study one

### 2.2.1 - Design

A two-part study investigated the use of the pre-accreditation documentation strategy deemed as semi-structured and the use of an online software deemed as a structured approach for the documentation of the ‘Analysis’ phase of the ACE-V process.

For part one of the study, an analysis form was created through a gap analysis between professional conversations with a senior fingerprint expert to establish what was documented in day-to-day casework and what the documentary suggestions in literature. The form was distributed as a hard copy. The template form included suggestions that related to prompts routinely documented on case management systems. This included the analysis outcome, digit or palm selection, pattern type and geographical search area. A free text box was also included to allow the examiner to explain how their decision was reached. [See appendix 2.](#Appendix_2)

Part two used a structured, online software (PiAnoS), which prompted participants to document their examination according to suggestions from literature and policies.

The software has the capability of annotating marks and assigning confidence levels on the overall quality of the mark and the minutiae identified. On the immediate interface there are a number of tools on the left-hand side, Q, M, R and O. These are the specific tools to annotate marks. Q refers to the quality map tool, the user selects between three different colours: green (high), orange (medium) and red (low). The colours indicate the confidence of using the selected areas for analysis. M refers to the minutiae tool, when the user locates a characteristic, this can be annotated by placing an icon indicating whether the characteristic is believed to be a bifurcation, ridge ending or unknown. R refers to the ridge count tool, it demonstrates the number of ridges between the core and the delta. O refers to the “other” features tool, for example if there are additional features that the user deems significant these can be annotated. This may include creases or scars.

In addition, there are a series of multiple-choice questions.

1 - General distortion factors observed on the mark

2 - Palm and finger segment positions

3 - General pattern

4 - Quality of level 1, 2 and 3 details

5 - Suitability of the mark

Questions 1 and 4 prompted justifications on how participants reached their reporting outcome at the Analysis phase. The suggestion of recording the presence of distortion and the quality of the mark comes from The Fingerprint Inquiry (2011), ENFSI (2015), SWGFAST (2012). Participants were familiar with the remaining questions as they are a part of day-to-day practice, as seen in analysis form. Following the questions, further notes could be entered into the text box available, however this was not compulsory.

Initial observations investigated trends in the dataset and determined the consistency between participants. Following this, a series of chi-square tests for association were conducted to determine the association between the grade of the mark and what was recorded. Chi-square test for association is used to find a relationship between two [categorical variables](https://www.statisticshowto.com/what-is-a-categorical-variable/). As well as association, the test can be used to demonstrate non-association as well. The two categorical variables in this study are grade of the mark (1 - 4) and documentary suggestions from the ACE-V checklist. The checklist contains over 30 different documentary suggestions therefore, the checklist was split into categories: generic information, factors affecting the mark and level of detail and annotations. Binary logistic regression was considered first, as a means of predicting obtaining a correctly identifying if an impression matches the documentation, however the data produced are not appropriate for this statistical test.

### 2.2.2 - Participants

A call for participants detailing the outline of the study was published on a social media platform (LinkedIn) and through the Chartered Society of Forensic Science newsletter to reach as many fingerprint experts as possible. After an expression of interest, a total of 31 fingerprint experts were recruited. Of these experts, there were five fingerprint bureaux from across the UK and two different independent companies. Amongst the five bureaux, this captures different areas of the United Kingdom including larger forces with the responsibility of multiple counties and smaller forces with the responsibility of their own county. The varied force sizes within the sample captures rural and urban areas to account for different workloads. Rural areas have a lower crime rate than urban areas according to Department for Environment Food and Rural Affairs (2022). Due to the size of the bureau and workload commitments, the number of participants depended on availability of experts within the force. As a result, there were more experts from some forces than others. The gap analysis demonstrated little notes were produced pre-accreditation and any notes were recorded on the case management system which was the same for all UK fingerprint bureaux therefore, the difference in participant numbers per bureau has no impact on the representative sample size for pre-accreditation documentation strategies. Each participant was tested at their place of work. It is a requirement of Staffordshire University that all research conducted meets the with regulations of the University’s ethics board. Proportionate ethical approval was gained.

### 2.2.3 - Materials

Eight friction ridge impressions were used in this study. The same eight impressions were used in both parts. The impressions were taken from Staffordshire’s ground truth database. The impressions varied in quality and quantity to allow for each element of the checklist to be recorded at least once. To quantify the quality, each mark was graded 1 to 4. To ensure a range of impressions in the test samples, there were single finger impressions, a palm impression and sequence of marks. As well as, impressions affected by substrate, pressure, movement, overlapping and development technique. See chapter 3, sub-chapter 3.1 for the selection process.

Each impression was printed 1:1 using a DCS5 compatible printer. The size replicates what experts are used to receiving in their day-to-day practice.

### 2.2.4 - Procedure

The eight images of friction ridge impressions were provided to all participants. Participants were briefed on the analysis form and given the chance to ask questions. No guidance was given on what to write in the free text box so that participants were not influenced. Participants were given two weeks to complete the analysis of each impression. Two months later, the same eight impressions were preloaded onto PiAnoS and participants were provided training. Training was provided at the participants’ place of work. Test samples were uploaded onto PiAnoS for training purposes and a step-by-step demonstration of the software was provided. Each tool and functionality were explained with the opportunity to trial the software under observation to allow for any questions to be answered in person before completing the study. Participants were given two weeks to complete the analysis of each impression.

## 2.3 – The establishment of post-accreditation documentation strategies for Fingerprint Examinations; study two

### 2.3.1 - Design

A two-part study investigated the use of new documentary practice following the implementation of ISO accreditation (semi-structured and structured approaches) and the original structured approach from study one, for the documentation of the ACE-V process.

Part one of the study involved participants’ recording fingerprint examinations using their new documentary practice. These practices have been approved by UKAS as each participant group has now been granted ISO accreditation. The documentary practice varies between participant groups (UK fingerprint bureaux), some of which were more structured approach than others.

Part two used the same structured, online software (PiAnoS), which was used in study one. It prompted participants to document their examination according to suggestions from literature and policies. The software has the capability of annotating marks and assigning confidence levels on the overall quality of the mark and the minutiae identified. In addition, there are a series of multiple-choice questions, which are explained in section 2.2.1. If the impression is suitable for comparison, the software presents a known print for comparison purposes. The known print is annotated similarly to the unknown mark during the Analysis phase. There is an additional tool whereby the paired minutiae can be highlighted to show the features in agreement and subsequently aid in the decision process. The only multiple-choice questions for this phase are outcome and if third level detail aided with the identification. Following the questions, further notes can be entered into the text box available, however this is not compulsory.

Initial observations investigated trends in the dataset and determined the consistency between participants. Following this, a series of chi-square tests for association were conducted to determine the association between the grade of the mark and what was recorded.

### 2.3.2 – Participants

A total of 33 fingerprint experts were recruited. 27 of the 33 experts participated in study one, the remaining participants could no longer participate due to the workload in the bureau. However, an additional UK fingerprint bureau agreed to participate which consisted of six fingerprint experts. The new participant group was recruited following a conference presentation. The participant group consisted of six UK police forces and one independent company which covers the largest city in Scotland, North East, North West, West Midlands, East Midlands and Northern Ireland. Each participant was tested at their place of work. Proportionate ethical approval was gained.

### 2.3.3 – Materials

The same eight friction ridge impressions were used in this study as were used in study one. The impressions consisted of single impressions, palm impressions and sequence of marks. As well as, impressions affected by substrate, pressure, movement, overlapping and development technique. Each impression was printed 1:1 using a DCS5 compatible printer. The size replicates what experts are used to receiving in their day-to-day practice. The comparison marks consisted of eight ten-print forms which were taken from all the individuals who contributed to the ground truth database.

### 2.3.4 - Procedure

The eight images of friction ridge impressions and eight ten-print forms were provided to all participants. Participants were asked to use their new documentary strategy to record the fingerprint examinations. Participants were given two weeks to complete the analysis of each impression. Two months later, the same eight impressions were preloaded onto PiAnoS and participants were provided with refresher sessions. The correct known print was uploaded for the comparison. It was not important for this study to assess if the participants correctly identified the mark, it was investigating the documentary strategy instead. Participants were given two weeks to complete the analysis of each impression.

## 2.4 – The determination of effective documentation strategies for Fingerprint Examinations

### 2.4.1 - Design

An interactive power-point presentation was used to investigate the effectiveness of documentation strategies for fingerprint examinations. The documentation strategies used in this study were witnessed in study one (pre-accreditation) and study two (post-accreditation). These strategies are a mixture of semi-structured and structured approaches.

### 2.4.2 – Participants

A call for participants was published on a social media platform, LinkedIn, to achieve a new participant group of experts who had not previously taken part in study one and study two. There were two participant groups: 1) a total of 30 fingerprint experts worldwide, 2) a total of 30 individuals with no fingerprint knowledge as the novice group. No fingerprint expert participant participated in both study one and study two. Each participant was tested at their place of work or home. Proportionate ethical approval was gained.

### 2.4.3 – Materials

The interactive power-point was split into four sections. Each section had 12 questions. Each question had an example of a documentation strategy and an impression. The impressions were from study one and study two; one impression from each grade (1-4). There was a total of nine different documentation strategies from either pre- or post-accreditation. In part, study one and study two involved the analysis of all the documentation which was cross compared to documentary suggestions from literature and policies. Any documentary suggestion which had a response rate of 50% or higher was included into the exemplar documentation. To test the effectiveness of the documentation strategy, there must be close matches included, therefore an additional four images were taken from the ground truth database which were the same donor and similar quality i.e., score/grade of the mark.

### 2.4.4 - Procedure

For each question, participants were provided with a friction ridge impression and an example of a documentary strategy. Using the documentation, the participant had to decide whether the notes relate to the friction ridge impression. A tick box was included to highlight whether it was a match, non-match or unsure due to lack of detail. In total there were 48 examinations participants were required to complete.

## 2.5 – The assessment of inter and intra-examiner variation of reporting outcomes at Analysis and Comparison/Evaluation stage

### 2.5.1 – Design

Same research design as study one and study two as the data are collected from those studies. Refer to section 2.2.1 and 2.3.1.

### 2.5.2 – Participants

In study one, 31 fingerprint experts were recruited to record the Analysis phase using the semi-structured approach from pre-accreditation, of these, 27 participants took part in the second part of the study which was recording the Analysis phase using a structured approach: the online software (PiAnoS).

In study two, due to dropouts and additions, there were 33 fingerprint experts who were recruited to record the ACE process using post-accreditation strategies, of these, 13 participants took part in the second part of the study which was recording the ACE process using the online software (PiAnoS).

Of the full participant group, 11 participants completed each stage of study 1 and study 2. In total, these participants reported their conclusions of the Analysis stage four times for the same impression and their conclusions of the Comparison stage twice for the same impression.

### 2.5.3 - Materials

The documentation of 11 experts taken from study one and study two. Specifically, the reporting outcomes from Analysis (four times) and Evaluation (twice) to assess the inter and intra-examiner variation. The documentation of these 11 experts were used to gain insight into the decision-making processes of each outcome.

### 2.5.4 - Procedure

Analysis and Evaluation reporting outcomes were taken from the data collected during study one and study two. Inter refers to the reproducibility of experts reporting outcomes based on how often experts differ between each other in their conclusions. Intra refers to the repeatability of experts reporting outcomes based on how often experts differ within their own conclusions.

# Chapter 3: Results and Discussion

This chapter independently reviews the results for each study of the research and discusses the results in turn. The research involved eight friction ridge impressions which were used throughout all the studies. Section 3.1 defends the selection of impressions used in the study by aligning a systematic scoring system used by the research team and a current UK fingerprint bureau grading system, to obtain a varied selection of impressions to replicate casework. Section 3.2 discusses the results from study one which addresses objectives 1 to 4. Section 3.3 addresses the results from study two which addresses objectives 5 to 7. Over these two sections, pre- and post-accreditation documentation strategies will be discussed, providing readers with an understanding of the documentation strategies by a selection of UK fingerprint bureaux and used in the effectiveness study. Section 3.4 discusses the results from the effectiveness study which led to the recommendations for the content of contemporaneous notes that are recorded by identification experts as part of fingerprint examinations in casework.

## 3.1 – The selected friction ridge impressions for study one and study two

In the PCAST report (2016) it states “casework is not scientifically valid research, and experience alone cannot establish scientific validity. One cannot reliably estimate error rates from casework because one typically does not have independent knowledge of the ‘ground truth’ or ‘right answer’.” Hockey *et al* (2020) explains of a recent case in Canada, which was challenged in court on the reliability of fingerprint evidence. R-v-Bornyk (2017) involved fingerprint evidence whereby an expert was questioned on whether the Royal Canadian Mounted Police Integrated Identification Services used ground truth data for the proficiency tests. The proficiency tests did not include ground truth data, only fingermarks from casework. Therefore, the reliability of the evidence was questioned. If there is no known truth to the test samples, then the accuracy of an expert is uncertain. Although the likelihood of experts reaching an incorrect conclusion is low, it was determined after criticisms from international bodies such as PCAST and discussions in court, to include a more scientifically valid proficiency test which includes the use of ground truth data. Ground truth data contain impressions purposely created to replicate casework samples. It should include a range of quality impressions, developed using techniques frequently used by the bureau. It should also contain fingerprints/palm prints as exemplars, recorded using the appropriate approach i.e., live scan, ten print cards and ink pads (Hockey *et al*, 2020). With the shift in quality standards in the UK, there is the requirement to include proficiency testing and dip sampling to assess individuals’ competencies. During professional conversations with the participant groups, it is apparent that UK bureaux have created their own ground truth databases for proficiency tests to demonstrate to the United Kingdom Accreditation Service the process in place to assess individual competencies. The ground truth database in this study was previously created with the researcher and local police force.

Within the utilised ground truth database, each mark was assigned an overall score between 6 and 15 to infer the quality of the impressions. Whereby a score 6 was a good quality impression and score 15 was a poor, unusable impression. The score was based on six criteria; the substrate interference, pressure distortion, deposition pressure, development techniques effect, presence of class characteristics, and the amount of ridge detail affected overall. Refer to chapter 2, section 2.1.1 for the specific breakdown for the scores. Scores were given based on the amount of area affected by the factors and the presence of class characteristics. This scoring system was created in a previous study with the researcher and local police force when producing a ground truth database. It was created to provide a systematic approach of assessing impressions. Anyone, experts or novices, who uses the system would judge the impressions using the same criteria, breaking it down and helping to defend the score. As this scoring system was created in collaboration with fingerprint experts, its relevance and appropriateness were assessed. The use of a scoring system with separate criteria and breaking down the quality assessment stems from a grading system used by Fieldhouse (2011). Within the Fieldhouse (2011) grading system, there are four criteria;1) the quantity of the surface area available; 2) the proportion of the mark available occupied by usable ridge detail; 3) the amount of contrast between ridges and furrows; and 4) ridge continuity. The difference between this scoring system and Fieldhouse (2011) grading system is the break down in criteria for example, rather than a criterion related to proportion of usable ridge detail, the alternative scoring system breaks down what is causing the distortion within the impression reducing the usable ridge detail. The benefits of having this break down in criteria allows for users to select what impression required for tests such as proficiency tests or research projects. The ground truth database consists of varied quality impressions, all these impressions are stored in a Microsoft Access database. Due to the design of the scoring system and using the filter tools in the Access database, users can retrieve impressions affected by substrate interference, pressure or development technique effects. In addition, users can retrieve impressions with or without level one (pattern type) present.

As previously mentioned, all bureaux will have produced their own ground truth databases and it should replicate the casework samples. Therefore, a methodology for assessing the quality would have been created in each bureau. During initial conversations with participating bureau, it was suggested to assess the alignment between the researcher’s scoring system and the participating bureaux grading system to ensure the varied quality impressions selected for the study replicated casework. The participants’ grading system was based on CAST’s grading system (Sears *et al,* 2012), as explained in section 2.1.2.

To assess the alignment between the two systems, four fingerprint experts graded ten impressions from the ground truth database. Of the ten impressions, there were two from each score 6, 8, 10, 12 and 15. The experts graded the impressions independently. On reflection of the grades and scores, the two systems aligned. Table 2 shows the scores and grades of the marks examined. From the results, it was decided that a mark which scored between 6 and 8 would be referred to as a grade 1 which is an impression with clear 1st and 2nd level detail, sufficient to see a pattern or area of the palm and is clearly identifiable. Scores 9 to 11 would be a grade 2 which is an impression with 1st and 2nd level detail visible, sufficient to identify but may not have a clear pattern or area of palm present due to distortion. Scores 12 to 14 would be a grade 3 which is an impression with 1st and 2nd unclear detail, there is limited quantity of ridge detail due to distortion but sufficient for identification purposes. Scores 15 and above would be a grade 4 which is an impression of poor 1st and 2nd detail and lacks quantity of ridge detail resulting in insufficient detail for identification purposes.

The ground truth database was designed to retrieve marks affected by different factors such as substrate interference, pressure, movement, development technique, therefore the database could be filtered to select marks which contain each of these factors. The test marks ranged from grades 1 to 4, contained single and multiple marks and affected by different factors to allow for experts to record necessary information relating to the mark. Figure 4 shows the different finger or palm marks used in the study. As the truth was known for all the marks, it could be tested whether participants recorded the information i.e., if the mark was deliberately affected by the substrate, was this recorded by the participants. The absence of writing does not mean the expert has incorrectly identified the deliberate factors; it could mean it’s not relevant to the overall decision. These eight impressions were used throughout the research. The first study to determine what was recorded during the Analysis and the second study to determine what was recorded during the full ACE process.

By using the same impressions, inter and intra-examiner variation of documentation and reporting outcomes could also be assessed. This will be discussed in section 3.5. An additional four images were used in the effectiveness study, which were close matches. Close matches refer to impressions that are the same grade taken from the ground truth database. The justification for including additional grades 1, 2, 3 and 4 impressions was to determine whether the original documentation contains sufficient detailed notes to identify a non-match between the documentation and another similar quality impression. This will be discussed further in the results and discussion for the effectiveness study (in section 3.4).

*Table 2 - The consensus grade and scores given to the marks examined.*

|  |  |
| --- | --- |
| **Consensus grade** | **Score from GTD** |
| 1 | 6 |
| 1 | 6 |
| 1 | 8 |
| 2 | 8 |
| 2 | 10 |
| 2 | 10 |
| 2 | 12 |
| 3 | 12 |
| 4 | 15 |
| 4 | 15 |

A picture containing text

Description automatically generated

*Figure 4 - Fingermarks selected from Ground Truth Database for test.*

## 3.2 – Objective 1 to 4: The establishment of pre-accreditation documentation strategies for Fingerprint Analysis; study one

As discussed in section 1.3, there is published literature and policies which suggests the content for fingerprint examination documentation (CPS, 2010, Champod & Langenburg, 2011, The Fingerprint Inquiry, 2011, SWGFAST, 2013, ENFSI, 2015 and Bunter, 2016). Referring to table 1, there are some similarities but also many differences between what is suggested to record. All of these were published before or around the time of the announcement of implementing ISO 17025 accreditation standards in England and Wales.

In summary from section 1.3, CPS (2010) provides guidance across all forensic disciplines which explains the broadness in relation to documentary suggestions.

It includes name of examiner, date of the examination, any materials used, any communications between professionals, reasons for conclusions and sufficient detailed notes. Similarly, the Forensic Science Regulator (2017) provides broad guidance on documentary suggestions for fingerprint examinations. The guidance from CPS (2010), The Forensic Science Regulator (2017) and the non-prescriptive guidance in relation to note-taking from the United Kingdom Accreditation Service, allows for flexibility in content and approach for documenting examinations.

Other authors suggest content for the reasons for conclusions (The Fingerprint Inquiry 2011, Champod & Langenburg, SWGFAST, 2013, Bunter, 2016 and ENFSI, 2015). This includes assessment of mark quality, signs of distortion, minutiae at analysis and comparison, with revisions at comparisons, presence of level three detail, any explanations of mark differences and reasons for conclusions (The Fingerprint Inquiry, 2011). SWGFAST (2013) differs slightly as the assessment of mark quality is not suggested, but signs of distortion, factors causing this distortion and minutiae present at analysis and comparison similarly to The Fingerprint Inquiry (2011). SWGFAST (2013) also suggests recording the orientation of the mark, presence of level one detail (pattern type) and acknowledging the matrix i.e., what the mark is made up of (or left in). The matrix could be used to indicate if the mark is deemed as complex, as the Forensic Science Regulator (2013) states a mark left in blood is a complex mark. ENFSI (2015) documentary suggestions are similar to SWGFAST (2013), however ENFSI like The Fingerprint Inquiry (2011) suggests recording the assessment of mark quality and the presence of cores and deltas. This is the only guidance to suggest presence of cores and deltas could be recorded as well as identifying the level one detail. If an expert records the presence of cores and deltas this explains how the decision was made for the level one detail. As described in section 1.1, the number of deltas differs depending on the pattern type for example, if one delta is present on a full fingermark it would suggest a loop pattern, whereas two deltas on a full fingermark would suggest a whorl pattern. SWGFAST (2013) only suggests recording presence of level one detail and not the cores and deltas, therefore not explaining the decision. Champod & Langenburg (2011) suggest annotating impressions demonstrating the level two detail i.e., individual characteristics observed during Analysis and Comparison stage and assigning confidence levels to these. A colour coding system was created known as the “GYRO” system to show the confidence in usability of the characteristics. This system was supported by Bunter (2016). All these documentary suggestions are included into the ACE-V checklist which is used to cross check against documentation produced throughout this research. The ACE-V checklist was completed through reviewing documentary suggestions in literature and policies (CPS, 2010, Champod & Langenburg, 2011, The Fingerprint Inquiry, 2011, SWGFAST, 2013, ENFSI, 2015 and Bunter, 2016).

Table 1 from section 1.3 forms the basis of the ACE-V checklist, the only differences lie within the distortion factors. In the table, it states any signs of distortion and factors affecting the mark with examples provided. In the ACE-V checklist, each of the suggested distortion factors is its own category, [see appendix 3 for the full checklist.](#_Appendix_3_–) The checklist was divided into subcategories. These are generic information, factors affecting the mark, level of detail and annotations and reporting outcome. An additional category which is irrelevant to this study but is required for study two is comparison and evaluation notes.

Despite the documentary suggestions within the ACE-V checklist published pre-accreditation, there is evidence to suggest not all documentary suggestions are included within notes by Bunter (2016) and within infamous cases such as R-v-Smith (2011) and HMA-vs-McKie (1997). Due to HMA-vs-McKie (1997), an inquiry was conducted and resulted in 86 recommendations for fingerprint examinations. Recommendation 80 stated “The Scottish Police Services Authority should continue to seek to obtain and retain the ISO 17025 external accreditation and such other accreditation as may become relevant in the field of fingerprint identification”. SPSA successfully gained the accreditation in 2015, however there is no evidence to imply the new documentation procedures involving contemporaneous notes. In 2015, The Forensic Science Regulator announced the requirement to achieve ISO 17025 accreditation in England and Wales. Bunter (2016) implies little change has been made to documentation procedures since the announcement of implementing ISO 17025 accreditation. This research will provide further evidence of the extent of note-taking pre accreditation and post accreditation. The first study sought to establish the pre-accreditation documentation strategies, as well as another approach for documenting the analysis stage of ACE and compare to best practice guidance also referred to as the ACE-V checklist. As stated in R-v-Smith (2011), a statement was prepared for court with only one sentence outlining who the impression was identified to. There were no explanations or reasons for conclusions as suggested by CPS (2010) which was published prior to the case in question.

Objective one was to perform a gap analysis on pre-ISO 17025 accreditation documentation strategies for fingerprint examinations to best practice guidance acquired from academic and grey literature. The results from this were used to produce an analysis form to replicate pre-ISO 17025 accreditation documentation and used in study one.

To conduct the gap analysis, there were professional discussions between the research team and a senior fingerprint expert from a collaborating bureau regarding the existing working practice pre-accreditation to the ACE-V checklist. Each section of the checklist was discussed with the senior fingerprint expert in relation to the possibility that a fingerprint expert may observe it in an examination and subsequently record it for future reference. During the gap analysis, it became apparent that most of the ACE-V checklist was not recorded in routine working notes prior to ISO 17025 accreditation and although the content of the ACE-V checklist was supported by the senior expert, it was suggested that fingerprint experts may have cases where they would think about the documentary suggestions seen in the checklist, but there was no requirement to write this down. This provided support to keep all documentary suggestions within the ACE-V checklist.

The only information recorded pre-accreditation was on a case management system called Socrates. This information included the reporting outcome at analysis stage, such as ‘sufficient for comparison on Ident1’, ‘suspect comparison only’ (against a suspect/victim’s ten print card) or ‘insufficient detail for comparison’. In the opinion of an expert, if the mark was ‘sufficient for comparison on Ident1’ the expert would input details to narrow the search for a comparison friction ridge skin impression from the national database. This information included which digit or area of palm it was likely to have originated from, the pattern type visible, and the geographical search location. There were four separate databases: police elimination database; local; regional; and national, and the expert would decide how wide the search would be dependent on the crime. When available, case information is provided to fingerprint experts. For this research, a summary of the fictitious case was provided to align with practice. It contained the offence, location of the mark, development technique and how the mark was lifted from the “scene”. The information can be used towards the experts’ Analysis. It can be used to explain the orientation of the mark or the distortion of the mark. Bunter (2017) shows how important it is to correctly interpret the position and orientation of fingerprints at a crime scene. In one of Bunter’s examples, he reported that a detailed examination showed that it was nearly impossible for the fingerprints to have been placed during the offence. He determined the fingerprints were more likely deposited before the gate was fitted.

Using the information derived from the discussion, an analysis form was created for the white box study to replicate documentary procedures prior to accreditation. This included the start date and end date of the examination, fictional mark information to aid with the examination (as this would be provided in casework), the reporting outcome at analysis (like the case management system) and Ident1 search information such as anatomical source and level one detail i.e., pattern type. As this was the extent of notetaking, the reasons for conclusions are limited and therefore, to gain an understanding of the thought processes and to establish key points to record in future documentation procedures, a free text box was included. [See appendix 2 for the analysis form.](#_Appendix_2_–) Before starting the research, the senior expert from each participating bureaux was asked if the analysis form replicated case notes prior to accreditation to ensure the form represented UK fingerprint examination documentation. All agreed.

No guidance on what should or could be recorded was provided as it was of interest to the research team to see what they thought was important to record and if there was any consistency between experts. This approach was deemed as a semi-structured approach due to prompts in the case management systems and the free text box for additional notes. All documentation produced was cross-checked against the ACE-V checklist to determine the response rate for each documentary suggestions from literature and policies.

The structured online software created for research purposes, which was used in the study included documentary suggestions from literature and policies (CPS, 2010, Champod & Langenburg, 2011, The Fingerprint Inquiry, 2011, SWGFAST, 2013, ENFSI, 2015, Bunter, 2016 and The Forensic Science Regulator, 2020) as seen in the ACE-V checklist. It included multiple choice questions prompting experts to record distortion factors such as substrate interference, overlapping marks and pressure. However, it has additional distortion factors and different terminology. For example, ‘double tap’ within the structured approach would relate to overlapping marks, ‘slippage’ would relate to movement. This also demonstrated the difference in terminology that can be used by experts. The difference in terminology has been witnessed by Earwaker (2017) when fingerprint practitioners were asked to explain their decision-making process when determining sufficiency for submission to the bureau for identification purposes. Within the same research, fingerprint experts were asked to explain their decision making for the same impressions. The terminology within this research in relation to distortion factors included movement in the mark, faint mark, superimposed, dotty, stretched, smudged, blurry, broken up. Within these terms identified, there are three descriptions to indicate marks over marks for example, overlapping, double tap and superimposed. A ‘faint’ mark may be used to suggest pressure applied during deposition and the lack of fingerprint residue to develop. Movement in a mark may also be defined as stretched, smudged, blurry, all these terms would be acceptable but highlights the difference in terminology between practitioners and experts. If an impression is described as ‘dotty’, it can be caused by chemicals used to develop impressions such as Ninhydrin (Laidlow Davis, 2017). All of which are contributing factors to the quality of an impression and were reported more frequently than a quality statement such as ‘clear mark’ or ‘not the best mark’. In addition to this question, the online software prompted experts to record the area of palm or digit identified as suggested by ENFSI (2015) and presence of level one detail i.e., pattern type suggested by SWGFAST (2013) and ENFSI (2015). This information would have been routinely recorded in the pre-accreditation documentation as discovered during the gap analysis. Champod & Langenburg (2011), The Fingerprint Inquiry (2011), SWGFAST (2013), ENFSI (2015) and Bunter (2016) all suggest recording level two i.e., individual characteristics. However, whether this is reported as a numerical value or annotated to show the usable characteristics differs between literature. When using this software, participants are prompted to mark up the finger/palm marks, assign confidence levels to the quality of the mark and the ability to use the ridge characteristics identified as suggested by Champod & Langenburg (2011) and Bunter (2016). However, as seen by Earwaker (2017) when determining sufficiency practitioners are more likely to record ‘not enough characteristics’ or ‘sufficient number of characteristics’ with the absence of a numerical value. Whereas experts were more likely to record a numerical value but not necessarily annotate the impressions. This information is used to determine the suitability of the mark and consequently leads to the conclusion therefore the final question relates to the outcome and is expected to be recorded before moving to the next impression. There was a free text box for further notes and remarks if necessary.

Similarly, to the semi-structured documentation, the structured documentation was cross-checked against the documentary suggestions from literature and policies. The results from this are discussed in sections 3.2.2 to 3.2.5. The results were compared to determine the consistency between examiners when using a semi-structured approach and a structured approach. The level of consistency between examiners using the free text box would imply that the elements recorded more frequently would be useful to record on a day-to-day basis. This may contribute to what should be recorded and potentially included into the effective documentation strategy later devised at the end of the research.

Furthermore, it was of interest to the research team to determine if the level of documentation differed dependent on the quality of the mark, or some may say “complex” marks. A complex mark is defined by the Forensic Science Regulator (2020) which has difficult, unusual or image resolution aspects to it. Examples, but not limited to are, marks in blood, an insufficient mark but later identified, very few clear ridge characteristics or a mark that has been previously compared to an identified person and excluded or negative automated search result.

Such marks could result in an inconclusive decision, and it is stated by Forensic Science Regulator (2020) to record the reasons for an inconclusive mark. This is supported by Langenburg (2011) and SWGFAST (2010) as it is recommended to produce notes based on the complexity of the impression.

To investigate the association between the level of documentation and grade of the mark, a series of Chi-square tests of association were chosen as the appropriate statistical tests. Chi-square test of association are used to find a relationship between two [categorical variables](https://www.statisticshowto.com/what-is-a-categorical-variable/). The assumptions for this test are the variables are measured at an ordinal or nominal level and the variables consist of two or more categorical, independent groups. The categorical variables in this study are documentary suggestions from ACE-V checklist and grade of the mark. As the checklist exceeds 30 independent groups, the tests are split into the different sections of the checklist. These sections are generic information, factors affecting the mark and level of detail and annotations. It must be noted, grades were assigned quality definitions; Grade 1 = excellent, grade 2 = good, grade 3 = poor and grade 4 = unusable.

The *count* value is the actual number of observations in a sample that belong to a category, i.e., if the finger or palm was recorded for an excellent mark (grade 1 mark).

The *expected count* value is the projected frequency that would be expected in a cell if the variables are independent. It is calculated by the row and column totals divided by the sample size. The *standardised residuals* are used to see which variables have the largest difference between the expected counts and actual counts relative to the sample size. Therefore, it is the standardized residuals which are of interest when there is a significant association to pinpoint exactly where this lies in the data (Minitab, 2019). The bar charts from the output of the tests were used to illustrate the counts more easily supported by discussions of the standardised residuals.

To investigate the impact of the significant associations effect size was calculated. For Chi-square this is reported as w. To calculate w, the square root of Chi (ꭓ2) was divided by the number of responses. Cohen (1988) stated a small effect size = ≤0.1, medium effect size = 0.3 and large effect size ≥0.5.

Retrospective power analysis was carried out to determine how much power the study had after completion. This was only required if the results were non-significant. When there is a non-significant association, power analysis is reported to discern the optimal number of participants, based on the results achieved, to obtain a power of 0.8 and increase the likelihood of then getting a significant result in future studies. 0.8 is a common power value used to find significant results. In other terms, 80% chance of detecting a difference or association. To calculate the required number of participants for future research, effect size tables based on degree of freedom were used (Clark-Carter, 2010). However, it must be noted that the optimal sample size may be unrealistic. In relation to this research, recruiting 400 fingerprint experts to get a significant association may be impossible due to workload and other commitments. In addition, even the optimal sample size may result in a non-significant association.

The p-value was reported for the Chi-square test for association, the significant association was explored and discussed using the crosstabulation and bar charts from the output, followed by the effect size and power analysis.

### 3.2.1 – Consensus of reporting outcomes at Analysis stage

During professional discussions with a senior fingerprint expert from a participant group, there were three outcomes at the Analysis stage which were ‘sufficient for comparison’, ‘suspect comparison only’ or ‘insufficient’. ‘Sufficient for comparison’ were marks which had clear ridge characteristics and an identifiable pattern present. In the opinion of an expert, the mark is of good quality which can be put through an AFIS system such as Ident1.

A ‘suspect comparison only’ mark tended to have an identifiable pattern but had unclear ridge characteristics. The mark could not be put through an AFIS system, but the pattern present could be used to exclude an individual. An ‘insufficient’ mark had unclear ridge characteristics or pattern present, and therefore could not be used to compare against known fingermarks, either on a database or manually.

It is evident that for good quality marks (grades 1-3) all participants agreed on the reporting outcome, as shown in table 3. All participants stated the mark was ‘sufficient for an Ident1 search’. For an Ident1 search, the mark had to be of good quality and usable level one and two detail presents. The reporting outcome was expected to be consistent between all participants due to the quality of the mark.

*Table 3 - Response rate for reporting outcome at Analysis stage.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Image** | **Grade** | **Sufficient** | **Suspect Comparison Only** | **Insufficient** |
| 1 | 1 | 31 | - | - |
| 2 | 2 | 31 | - | - |
| 3 | 4 | 9 | 20 | 2 |
| 4 | 1 | 31 | - | - |
| 5 | 4 | - | 1 | 30 |
| 6 | 3 | 31 | - | - |
| 7 | 2 | 31 | - | - |
| 8 | 1 | 31 | - | - |

Whereas, with the poor-quality marks (grade 4), there was a difference of opinion on the reporting outcome. Image 3 had the most variance of opinion. When depositing this mark, force was applied to deliberately affect the mark with pressure distortion. Yet, despite the pressure applied, some ridge characteristics were identifiable to a number of the participants, therefore there was potential to compare to suspects ten prints or run through Ident1. This did not fully fit with the grade 4 definition. There was more consistency between experts’ reporting outcome for image 5. Most likely due to multiple factors affecting this mark (substrate interference, development technique and overlapping marks). However, one participant did report the mark was suitable for suspect comparison only because of the friction ridge detail they believed to be visible.

The reporting outcome from the semi-structured and structured approaches were compared to establish whether the participants had changed their outcome.

The marks which were reported as ‘sufficient for comparison’ during the semi-structured approach were also recorded as ‘sufficient for comparison’ in the structured approach. The intra-variation occurred within the poor-quality marks (grade 4). Within image three and five, there were ten participants, in total, which changed their reporting outcome. There were a number of variations:

* Suspect comparison only to sufficient for comparison
* Sufficient for comparison to suspect comparison only.
* Insufficient for comparison to suspect comparison only.
* Sufficient for comparison to insufficient for comparison

The variances showed the potential for misidentifications. Marks which were deemed as only suitable for comparison against suspect ten prints are restricted by potential comparators. The use of ten prints can aid with eliminating and/or identifying individuals, however the disadvantage is the potential of missing an individual who is on the national fingerprint database. There were instances whereby participants originally reported ‘suspect comparison only’ however, the decision changed when using the online software to ‘sufficient for comparison’. During the original examination, there could be a potential misidentification if the suspect was not a part of the ten prints provided but the suspect was in the fingerprint database. If an expert has deemed the mark as ‘insufficient for comparison’ this mark would get discarded from the case, but the expert stated the mark would be sufficient for suspect comparison only when examined using the structured approach (an online software). In both instances, the outcome changed positively after the use of the online software, whereby an identification could potentially be made. It is possible that the ridge detail was clearer using the online software which may explain the change of opinion. However, there was one participant which changed their reporting outcome from ‘sufficient for comparison’ to ’insufficient for comparison’, after using the online software. There is a research opportunity, extending these results, to establish whether the use of computer software aids with the clarity of ridge detail during the examination of fingermarks.

Despite the general consistency between participants’ reporting outcomes, there are occasions where experts differed in their conclusions, as expected for the poor-quality marks. Prior to accreditation, fingerprint experts produced limited documentation and would create retrospective notes for court purposes following the Streamlined Forensic Reporting procedure. Production of retrospective notes has a disadvantage as there is no original notes to understand the original expert’s decision and as previously seen in these results, experts can differ between and within their decisions. In addition, the original expert may not necessarily be the expert producing the SFR2 report and/or court report and the variance between experts, particularly the complex marks would benefit from having the original examination notes. The following sections explain the results from the study and show there is a variance of opinion on what is recorded within the contemporaneous notes, using both approaches. The importance of each documentary suggestion will be discussed and supports the need for contemporaneous notes rather than retrospective notes.

### 3.2.2 – Initial observations and chi-square tests for association for “Generic information” recorded from the ACE-V checklist.

One of the categories of the ACE-V checklist is “generic information”. This includes the unique reference number which refers to the case (CPS, 2010, Forensic Science Regulator, 2017). The date, start and end time of the examination allow for an auditable trail of the case (CPS, 2010). If the case needed to be reviewed, it can be searched using information such as URN and date. The materials used throughout the examination could be recorded to allow for the verifier to have the option to use the same materials as the original expert (CPS, 2010, Forensic Science Regulator, 2017). Following this, participants may record if the mark originated from the finger or palm and which finger or area of palm is present. This information is required when a mark is ‘sufficient for an ident1 search’ or can aid with the ‘suspect comparison only’ marks. Additional generic information about the mark is suggested by ENFSI (2015) such as if the mark is single or multiple marks present and if there is more than one mark present, participants may record the relationship between the marks present. The spatial relationship may indicate the marks are part of a sequence. Table 4 shows the response rate of recording these documentary suggestions using the semi-structured approach (left figure) and the structured approach (right figure). The grades for each image are also provided to show the relationship between the quality of the mark and what is recorded.

**Initial observations**

When observing the semi-structured documentation, there was high consistency between participants response for the URN, date, time, digit or palm selection. These aspects were expected to be recorded as it forms part of experts’ daily routine which explains the high response rate. The URN, date and time were not recorded using the structured approach as this was not included in the software. There was a low response rate for the materials used being recorded. This could be due to experts not being asked within the semi-structured and structured approach and it is not common practice to record this information. However, it will be expected to be recorded as it forms part of the documentary requirements stated in The Codes of Practice and Conduct for Fingerprint Comparisons (2017). The benefits of recording the materials used in the examination include second experts replicating the original expert’s examination.

Participants recorded if the mark was a finger or palm for the grade 1 to 3 marks. This corresponds to the consensus reporting outcomes as there were enough characteristics visible for an ident1 search or suspect comparison. When a mark is ‘sufficient for an ident1 search’, an examiner will input which finger they believe the mark to be originated from, which explains the high consistency of the response for “which finger or palm”. However, despite the consistency between participants recording which finger or palm, when analysing the results further, not all participants agreed of which finger or palm they believed the mark originated from. There is potential for misidentifications or false identifications, having recorded this aspect, the mistake can be identified.

There is less consistency between examiners response for categories; if the mark is single or multiple and if there is a relationship between the marks. These were not direct questions on the semi structured or structured approach but could have been included in the free text box as they are common mental observations by experts. “If there is a relationship between the marks” was only recorded when there was presence of a sequence of marks or multiple marks. However, there was not a high number of responses for this, this could be due to participants not regularly recording this. Experts would use the spatial relationship, location and orientation of the marks to determine if the marks were related to each other.

Over the eight images within the study, each documentary suggestion within the generic information category was recorded at least once. This implies that these suggestions are relevant and shall be recorded, but not in every case. It is clear that there is more consistency between participants when directed, however this doesn’t necessarily mean that the documentation strategy is effective. The term “effective” refers to successfully producing a desired or intended result. The Forensic Science Regulator (2017) states that documentation should be sufficient for an auditable trail. There could be instances where all participants are consistent in what is recorded within the documentation, but the information is not enough to follow for a sufficient auditable trail as requested by the Forensic Science Regulator (2017).

*Table 4 - Response rate for generic information in both approaches.*

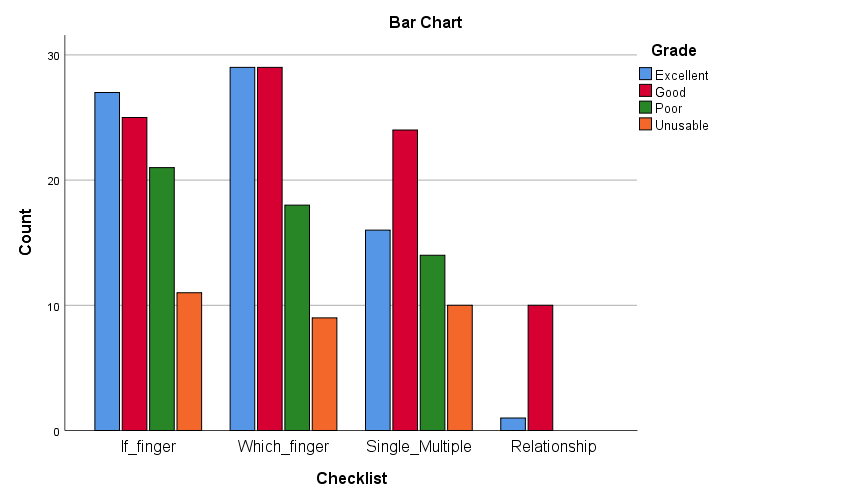
Calendar

Description automatically generated

**Chi square test for association: pre accreditation and online software**

The Chi square test indicated a significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting the generic information during the Analysis using the semi-structured approach (before accreditation documentation strategy), ꭓ2 (9) =17.97, p=0.036, two-tailed test,w=0.76. This implies a lack of consistency between experts’ note taking over varied quality marks, more so in the “Relationship to the marks” category. This has been identified in the initial observations.

The significant association lies within the relationship category. This means experts had recorded the mark had a relationship between other marks presents in the image for example, a sequence of marks or multiple marks. More than, in comparison to the other marks reviewed through the other checklist categories, the participants were significantly higher than expected (this is also shown in the graph, see figure 5, in the large difference between the ‘excellent’ and ‘good’ bars for Relationship). The total count for a good mark (grade 2) was 10 whereas, the expected value is 4. It is significantly higher than expected however, this difference is explainable. A count of 10 was an average taken from two grade 2 marks within the test, one of which was a sequence of marks opposed to a single mark. Using Cohen’s W, the findings have a large effect.If this test was to be carried out again, and the sequence of marks were omitted, this would probably result in no significant association between the level of documentation recorded by experts despite the grade of the mark. This is somewhat seen in the initial observations with the higher consistency between examiners when prompted to record certain aspects i.e., if finger or palm and which finger or palm.



*Figure 5 – Participant response rate for the generic information against the grade of the mark   
(before accreditation*

The Chi square test indicated no significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting the generic information using the online software as a means of documenting the generic information during the Analysis, ꭓ2 (6) =2.487, p=0.870, two-tailed test, w = 0.30. This implies there is more consistency between the experts’ response for that category despite the grade of the mark. This is explainable as all these categories are a requirement to complete in the software and not dependant on the grade of the mark.

The findings have a medium effect and is greatly under powered at 0.19. To achieve a power of 0.80, the study would need 150 experts. However, recruiting this many experts is highly unlikely due to operational workload of completing examinations whilst gaining ISO 17025 accreditation. It must be noted that even if 150 participants were recruited, this would not necessarily produce a significant association between the checklist and grade of the mark if people do not record these categories ever, or very often. It could be suggested not to continue with this study due to no significant associations occurring. However, as previously mentioned the online software asks of these categories to be recorded and in the overall aim of the research to establish effective documentation strategies all this information was recorded despite the grade of the mark may be relevant and useful to include within effective documentation.

### 3.2.3 – Initial observations and chi-square tests for association for “Factors affecting the mark” recorded from the ACE-V checklist

This category relates to the recording of information about the position of the mark, the mark quality and quantity and subsequently what factors may be affecting the quality of the mark. These factors include substrate interference, development technique, pressure applied during deposition, movement and overlapping of marks. Friction ridge detail can be recovered from a variety of surfaces, some more problematic than others. Patterned or textured backgrounds can interfere with the ridge detail which affects the ability to observe continuous ridges. For example, writing on a newspaper could run through the fingermark reducing the clarity of the ridge detail. Some development techniques can affect the ridge detail available. This can be due to the chemicals used such as Ninhydrin producing a dotty appearance (Laidlow Davis, 2017), the application of the user such as over-powdering (Bandey & Gibson, 2006) or the components within the mark not adhering to the treatment (Stubbs-Hayes *et al*, 2015, Sears *et al*, 2012, Croxton *et al*, 2010). Furthermore, fingermarks are left behind on surfaces after a single touch, yet how they are deposited will differ between individuals and circumstances. Due to these differences during deposition, fingermarks could be distorted due to pressure if the individual firmly touched a surface, or a little amount of ridge detail could be present if lightly touched (Fieldhouse, 2011). Similarly, a fingermark may be distorted by movement during deposition. If an individual has touched a surface or picked up an item several times, or it has been handled by different people, fingermarks can be overlapped. This overlapping can interfere with the ridge continuity and therefore affect the clarity of the ridge detail. At a crime scene, it is unlikely to have a perfect print, fingerprint experts will be familiar with the above factors affecting ridge detail. These factors could be the reason why a mark is deemed ‘insufficient for comparison’ or ‘inconclusive’. It appears that this information is not routinely recorded (Bunter, 2016) but is suggested in certain literature and policies (The Fingerprint Inquiry, 2011, SWGFAST, 2012, ENFSI, 2015).

Table 5 shows the response rate of recording these documentary suggestions using the semi-structured approach (left figure) and the structured approach (right figure). The grades for each image are also provided to show the relationship between the quality of the mark and what is recorded. There are two suggestions (location of the mark and what the matrix is) that are not recorded throughout each fingermark examination by all 31 participants, using the semi-structured and structured approach. 31 participants conducting 8 examinations, twice, equates to 496 opportunities of recording these suggestions.

**Initial observations**

The location of the mark was not recorded, potentially because of the location already recorded in the case information. For example, taken from kitchen windowpane, possible point of entry. However, there can be cases whereby the location of the mark aids with the decision. Bunter (2017) discusses how the location and orientation of fingerprints on a security gate showed that it was virtually impossible for the fingerprints to have been placed during the offence. It was suggested that contact was more likely to have occurred before the gate was put in place, ten years earlier.

*Table 5 - Response rate for factors affecting the mark in both approaches.*

Table

Description automatically generated with medium confidence

In addition, the matrix was not recorded. “Matrix” refers to what the mark is made up of, or left in. This is the substance that is deposited by the finger and eventually developed such as sweat or blood (The Forensic Science Regulator, 2013). All test marks were sweat marks, but this was not mentioned by any of the participants. However, if there was a mark in blood, there is potential for this information to be recorded due to The Forensic Science Regulator (2017) stating a mark in blood is deemed as a “complex” mark. As previously mentioned, a complex mark may have a different documentary process, therefore the presence of a mark in blood would be documented.

Despite neither of these documentary suggestions have been used in the test marks, there is evidence to suggest that there will be cases in which they are required, so should not be disregarded.

The remaining documentary suggestions were recorded however, there is inter- and intra-examiner variation. The most recorded within this category is the quality of the mark. The quality of the mark is assessed to determine whether the mark is sufficient for comparison or not, which explains the response rate. However, not all participants recorded this using the semi-structured approach. Whereas, using the structured approach there was 100% consistency between examiners as it was a direct question. The quality of the mark was recorded more frequently for the grade 1 and 2 marks in comparison to grades 3 and 4. For example, participants would state the mark was of good quality (grade 1 and 2) or the mark was affected by a factor rather than stating it was a poor-quality mark (grades 3 and 4). A record of factors affecting the mark are expected to be higher in poor quality marks.

As image one is a grade one mark, there is little interference to the mark from factors such as movement, pressure, substrate, overlapping and development technique. There is a more varied response on what could be affecting the mark in the semi structured in comparison to the structured approach. It appears that when participants are asked direct questions, there is more consistency. As the ground truth for this mark is known, it should be noted that there were no deliberate factors enforced onto this mark during deposition. Despite image four also graded one, more factors affecting the mark were recorded. There was more consistency in the structured approach. The deliberate factor within this mark was movement, the donor placed their finger onto the surface and moved it to the left, causing a swipe. Out of the 27 participants who used the structured approach, 17 recorded this factor. This is a higher response than the semi-structured approach. Image eight also grade one had a mixed response during the semi-structured approach but more consistency in the structured approach. This mark was a palm mark on paper, developed using ninhydrin. This explains the response rate for development technique affecting the ridge detail and the substrate causing background interference.

Image two and image seven were graded two. Image two was a sequence of marks with no deliberate factors however, it was developed using ninhydrin which can cause distortion as previously seen. The development technique affecting ridge detail was discussed by 13 participants out of 27 during the structured approach, this is a higher rate than the semi-structured approach. Image seven was deliberately affected by using a newspaper to cause distortion. The writing on the newspaper runs through the ridge detail. However, the pattern type is still visible therefore the mark is ‘sufficient for comparison’. This is the only mark in which the participants were 100% consistent using the structured approach. Therefore, if there is a clear reason why the ridge detail is affected and the participants are directly asked about this, then the reason is recorded.

Image six was the only grade three in the test images. This mark was affected by development technique. The technique used was cyanoacrylate fuming, and the mark was overdeveloped which reduced the clarity of the mark. This was only discussed by one participant over both approaches. Participants focused on movement and pressure. The response rate for factors affecting the mark is higher within this mark than grades one and two.

Images three and five were graded four. These are the two marks which had the difference of opinion in the reporting outcome, seen in table 3. Despite having the same grade, one mark had more factors affecting the mark than the other which explains the difference of opinion on the reporting outcome. Image three had majority ‘suspect comparison only’ outcome. Therefore, the participants believe they can see the fingerprint pattern and some ridge detail which may help to exclude. The ridge detail is poor subsequently cannot be inputted into ident1. 19 out of the 27 participants, (or 70%) had discussed the substrate was causing distortion. However, the known truth of this mark shows that the mark was affected using heavy pressure. Only 5 participants recorded this factor using both approaches. To establish if the mark is affected by heavy pressure, the ridges may appear closer together which could cause misinterpretation. A heavy pressure mark may be developed darker than a light pressure mark potentially due to the amount of residue left behind after a heavy versus a light deposition. Again, there was more consistency using the structured approach opposed to the semi-structured approach.

30 participants stated image five was ‘insufficient for comparison’, with one exception who stated, ‘suspect comparison only’. The most common factor recorded during the semi-structured was overlapping (n = 16). This was the deliberate factor affecting the mark. However, other factors contributed to the quality as the mark was deposited on newspaper and therefore developed using ninhydrin. These factors were recorded as well as pressure and movement. The results from the structured approach showed the highest response was substrate (n = 25), almost all the participants. Closely followed by substrate (n = 23). Other factors were also recorded. All responses were higher using the structured approach. Having this information can help to show why the expert deemed the mark as ‘insufficient for comparison’ or ‘suspect comparison only’.

Orientation refers to the logical direction of the friction ridge detail if it had been deposited under controlled conditions. The ridge flow, position of the deltas and other detail such as creases are also used to orientate the mark (The Forensic Science Regulator, 2017). This is an important factor to aid with fingerprint examinations however, it is not often recorded using either approach. In a recent case, an expert nearly missed reporting an identification due to looking at the print upside down. This shows the importance and opportunity to potentially miss an identification if the mark has not been orientated correctly. If the expert records the orientation and it is incorrect, this can explain if there is a difference of opinion.

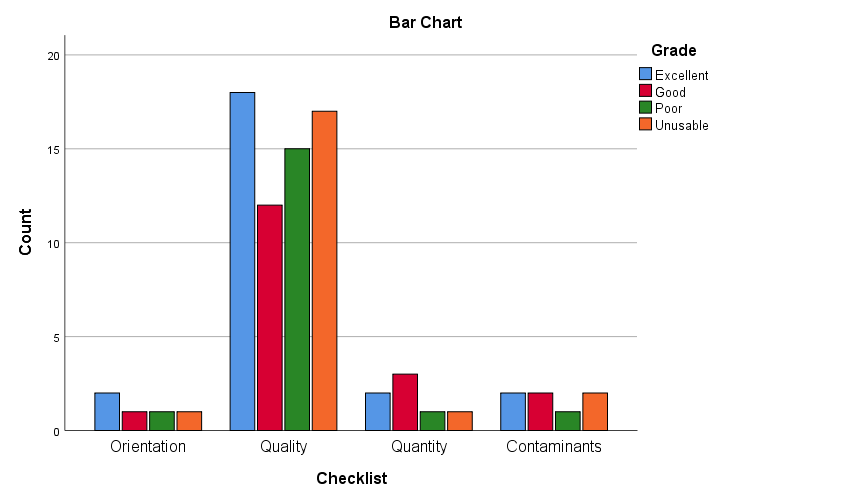
Overall, there is less consistency between experts in this category than generic information category, as expected due to the subjective nature of the factors affecting the mark. However, as explained, these factors explain the thought process by experts therefore are important to record. There appears to be more consistency when prompted using the structured approach, or when the mark is clearly being affected by a particular factor. But again, the question is, is this effective documentation? Does this contribute to a sufficient auditable trail? This will be measured in the effectiveness study. In addition, the information recorded in this category appears to be heavily dependent on the grade of the mark. The quality of the mark is recorded if it’s a good quality mark, whereas if it’s a poor-quality mark the reason is recorded more than stating it’s a “poor quality mark”, this is assessed statistically using chi square tests for association in the section below.

**Chi square test for association: pre accreditation and online software**

Due to the number of elements in this category, for the Chi square tests the category was split into two tests. The first test “factors affecting the mark and the grade of the mark” involved orientation of the mark, quality of the mark, quantity of the mark and the presence of contaminants.

The second test “more factors and the grade of the mark” involved presence of movement, substrate interference, pressure distortion, overlapping marks and development technique affects.

The Chi square test indicated no significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting the factors affecting the mark during the Analysis before accreditation, ꭓ2 (9) =2.925, p=0.967, two-tailed test. This implies there is more consistency between the experts’ response for a particular category despite the grade of the mark. The quality of the mark is recorded consistently despite the grade of the mark, whereas the other categories in this test were recorded very little as seen in the graph (figure 6). This can be explained as the quality of the mark can be recorded despite the grade of the mark because the participants would explain if the mark was of good or poor quality in their notes, potentially. It must be noted that not all the participants did record this expect.



*Figure 6 – Participant response rate for the factors affecting the mark against the grade of the mark   
(before accreditation)*

w = 0.31 is a medium effect size but is under powered at 0.15. To achieve a power of 0.80, the study would need 180 experts. However, recruiting this many experts is highly unlikely due to operational workload of completing examinations whilst gaining ISO 17025 accreditation. It must be noted that even if 180 participants were recruited, this would not necessarily produce a significant association between the checklist and grade of the mark if people do not record these categories ever, or very often.

The Chi square test indicated a significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting the more factors during the Analysis before accreditation, ꭓ2 (12) =29.167, p=0.004, two-tailed test. This implies there is inconsistencies between the experts’ response for that category. This is potentially due to semi structured approach using a free text box and no influence on what to record. The biggest residuals appear in the overlapping category between good and poor (-1.6 and -1.5). For both, there are significantly fewer people recording these than would be expected. Again, this can be seen in the movement category in the good grade (-1.6) where there is also significantly less than expected.

w = 0.97 is a greater than large effect size and is the right level of power as it should be 0.80. It is a robust finding that factors affecting the mark are having an impact on the quality of the mark. It could be suggested that the factors should be recorded to suggest the quality of the mark.

The Chi square test indicated no significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting the factors affecting the mark using the online software as a means of documenting during the Analysis, ꭓ2 (6) =2.532, p=0.865, two-tailed test. This implies there is more consistency between the experts’ response for that category despite the grade of the mark. Participants are likely to respond equally, regarding the quality of the mark despite the grade of the mark, as you might expect as the two are related. Also, this is a question in the online software. Contaminants was removed as there were no responses across all the grades.

w = 0.33 is a medium effect size but is under powered at 0.19, very similar results to before documentation for factors affecting the mark (as seen above). To achieve a power of 0.80, the study would need 150 experts. However, recruiting this many experts is highly unlikely due to operational workload of completing examinations whilst gaining ISO 17025 accreditation. It must be noted that even if this number of participants were recruited, this would not necessarily produce a significant association between the checklist and grade of the mark if people do not record these categories ever, or very often.

The Chi square test indicated a significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting the more factors using the online software as a means of documenting during the Analysis, ꭓ2 (12) =52.023, p<0.001, two-tailed test.

This implies there is less consistency between the experts’ responses for that category, across all grades of the marks. Even though the online software encourages experts to record the different factors affecting a mark, there appears to be inconsistencies between what experts are identifying as affecting the mark across the different grades. This can be expected as a grade 1 mark will have minimal distortion whereas a grade 4 will have excessive amount of distortion to the extent of non-identifiable pattern and ridge characteristics. The highest residual is in the category pressure for a poor grade which was 3.9, a lot more people recorded it than expected. When observing the expected counts, there is a clear increase, it suggests experts are more likely to record factors affecting the mark (movement, pressure, substrate interference, overlapping and development technique) for the poor and unusable marks, this was also seen in the initial observations.

From the results, participants are most likely to record the substrate interference over the other factors, this is potentially due to the ease of identifying this as a factor as opposed to pressure applied during deposition or development technique effects.

w = 1.39 is a greater than large effect size and is the right level of power as it should be 0.80. Again, it is a robust finding that factors affecting the mark are having an impact on the quality of the mark. It could be suggested that the factors should be recorded to suggest the quality of the mark, no matter what structural approach for the documentation strategy. There appears to be more factors recorded using this approach opposed to before accreditation, the explanation for this is the online software prompts the participants to record this expect.

### 3.2.4 – Initial observations and chi-square tests for association for “Level of detail and annotations” from the ACE-V checklist

The final category relates to a significant aspect to fingerprint examinations, identifying fingerprint patterns and ridge characteristics. Literature and policies suggest that the presence of level one (SWGFAST, 2012, ENFSI, 2015), level two and other features (The Fingerprint Inquiry, 2011, SWGFAST, 2012, ENFSI, 2015) could be recorded within contemporaneous notes. This information is ultimately what is used to make a comparison between an unknown mark and a known mark which shows its importance within the original experts’ contemporaneous notes. Level one refers to the fingerprint pattern which could be arch, loop or whorl (Ashbaugh, 1999, Langenburg, 2011, SWGFAST, 2013, ENFSI, 2015). The experts would be required to input this information into ident1 when searching the database. This explains the high consistency between expert’s response for this element during the semi-structured approach.

**Initial observations**

It can be seen in table 6 all participants recorded what level one detail was present for grades 1 to 3. This is expected due to the quality of these marks. It should be noted that all participants recorded and correctly identified the pattern type for each mark. As previously mentioned, the definition of a grade 3 mark is “1st and 2nd unclear, limited quantity, distortion, but sufficient for identification purposes”. The results from this mark coincide with this definition, the reporting outcome was majority ‘suspect comparison only’ (n = 20), followed by ‘sufficient for comparison’ (n = 9), both outcomes imply the pattern is present. 29 participants recorded the pattern present within this mark which equates to the same number of participants recording ‘suspect comparison only’ or ‘sufficient for comparison’. The structured approach included a question relating to the pattern type present within the mark which explains the high response rate for level one detail. It is important to record this during the original examination as if there is a dispute about an identification, having the original notes can be useful to understand what the expert observed at the time of the examination.

If an expert has identified a pattern, a more in-depth examination is carried out, establishing what level two detail (ridge characteristics) is present within the mark such as ridge endings and bifurcations. The formation of these ridge characteristics is unique to individuals. Within the semi-structured approach, the presence of level two detail was rarely recorded by participants, despite being suggested by leading organisations such as SWGFAST and ENFSI. However, the UK are not bound to these guidelines.

In 2017, The Forensic Science Regulator published Codes of Practice and Conduct, the guidance was not as extensive as the other guidelines available. Experts should produce sufficient detailed notes, but what goes into these notes is not explicit within the Codes.

The remaining documentary suggestions within this category are less likely to be recorded within the semi-structured approach. The documentary suggestions include the presence of level two (ridge characteristics) and other features (scars, pores), as well as annotating the mark to show these features. As the analysis stage is the “information gathering” stage, experts would choose the target area and identify the ridge characteristics within this area which can be used within the comparison stage. A colour coding system called “GYRO” was created to assign confidence levels to the features observed within a mark (Langenburg & Champod, 2010). Green equals high confidence, yellow equals medium confidence and red equals low confidence in using the characteristics, within the comparison mark. Orange is used for characteristics observed during the comparison stage. The authors believe that this method will produce more transparent documentation because it demonstrates critical features relied upon to reach a conclusion. It has been a part of training since 2005, but it has yet to be utilised by UK fingerprint bureaux which explains the low numbers. However, it was apparent that one independent company uses a digital software to mark-up friction ridge detail and assign confidence levels. It is argued that this documentary suggestion is too time-consuming and would not be cost-effective as the number of cases which go to court are significantly lower than day to day cases.

The results from the semi-structured approach show it is unlikely that the identified ridge characteristics are permanently recorded, just mentally. As the ISO 17025 accreditation standards require contemporaneous notes to be produced, mental observations would not

suffice under these standards. However, this is not suggesting that the number of ridge characteristics must be noted. Although, it is recommended in other literature such as Langenburg (2012) to annotate impressions during the Analysis stage and use a system like the GYRO system to provide confidence levels in the usability of characteristics. This is justified by the number of characteristics differing between experts but primarily due to some experts only noted the clear characteristics whereas other experts would note a higher number of characteristics but a varied opinion on clarity and usability. For example, one expert stating 8 clear characteristics, one expert stating 8 clear characteristics and 8 unclear characteristics. Here it has shown when providing confidence levels, it gives an indication of the thought process during the Analysis stage. As well as the presence of level two detail and the mark up of these characteristics, it is suggested to record the number of level two detail present. It is not frequent practice for these documentary suggestions to be recorded prior to accreditation which explains the low numbers. Since the numeric standard was abolished, fingerprint experts have not used an official threshold, more so a personal threshold or an operational threshold as seen in recent studies Ulery et al (2012).

The documentary suggestions have been included into the structured approach, which explains the high numbers within the results, particularly the presence of level one, two and three detail and assigning confidence levels to the quality of the mark and characteristics identified. The annotated impressions highlighted the difference in number of minutiae, type of minutiae i.e., ridge ending or bifurcation and the location in which these minutiae were identified. This is like the results seen by Langenburg (2012). The difference in opinion is observed in research and understood due to the subjectivity of the field. However, it is of interest to the research team whether this difference has an impact on the comparator list. This further work idea is explored further in section 4.5.

*Table 6 - Response rate for level of detail and annotations in both approaches.*

Calendar

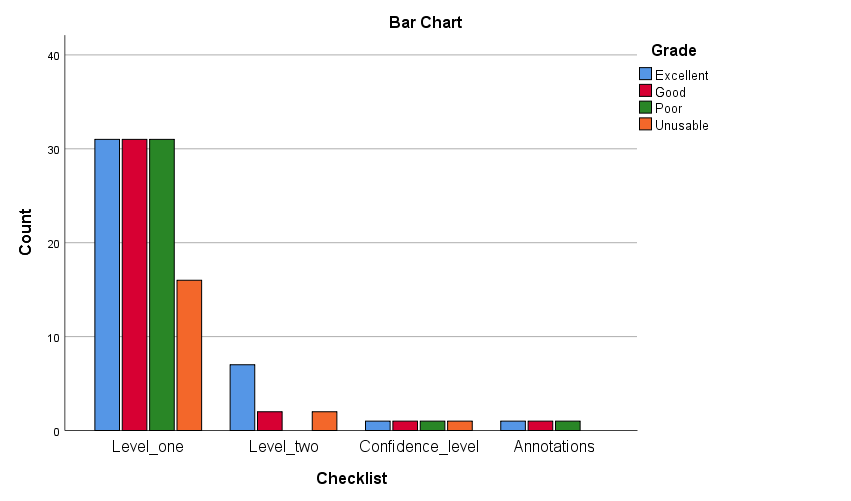
Description automatically generated

Similarly, to the semi-structured, the presence of level one detail was recorded for the grade 1 to 3 marks within the structured approach. There was a low response for the grade 4 mark, as expected due to the quality. There were no annotations for level one detail using the structured approach, however this was not included in the software. To include this, there could be tools to indicate where deltas are within the mark, as this can be used to identify a pattern. This information can be useful in cases where there is a difference of opinion. If an expert has incorrectly identified an individual based on the characteristics present, this can be seen within the contemporaneous notes. Subsequently, the expert can be provided with further training. As the online digital software is purposely built to mark-up friction ridge detail, it was expected to have level two detail annotations. For the good quality marks, grades 1 to 3, all participants marked up the images. The position of the minutiae, the characteristics identified, and the total number of minutiae differed between participants and between marks. The implications of this are suggested as further work in section 4.5.

**Chi square test for association: before accreditation and online software**

The Chi square test indicated no significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting the level of detail during the Analysis before accreditation, ꭓ2 (9) =8.391, p=0.495, two-tailed test. This implies there is consistency between the experts’ response for that category despite the grade of the mark. Looking at the graph (figure 7) it also suggests there is an equal chance of recording the presence of level one detail for each grade of mark. This is explainable as experts were prompted to record the level one detail present before accreditation, but not level two detail, confidence levels associated with the mark up of minutiae and any annotations.

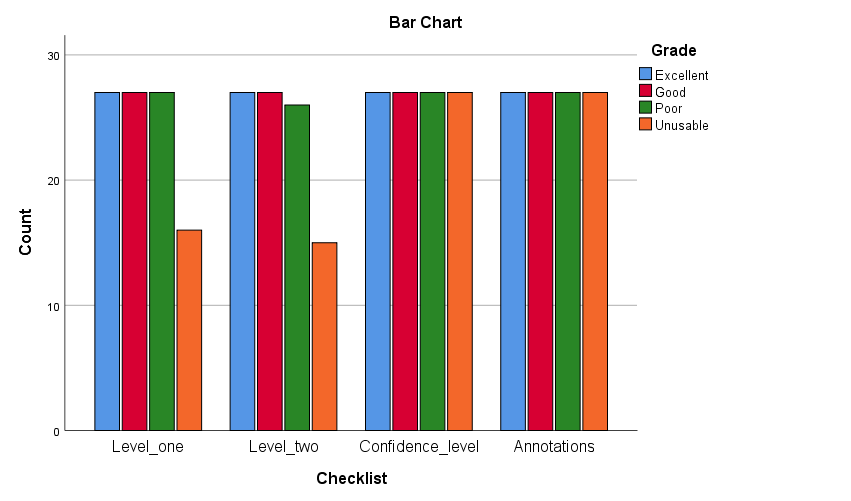
w = 0.52 is a large effect size but is under powered at 0.42. To achieve a power of 0.80, the study would need 65 experts. However, recruiting this many experts is highly unlikely due to operational workload of completing examinations whilst gaining ISO 17025 accreditation. It must be noted that even if this number of participants were recruited, this would not necessarily produce a significant association between the checklist and grade of the mark if people do not record these categories ever, or very often.



*Figure 7 – Participant response rate for the level of detail/annotations against the grade of the mark   
(before accreditation)*

The Chi square test indicated no significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting the levels of detail using the online software as a means of documenting the generic information during the Analysis, ꭓ2 (9) =4.866, p=0.846, two-tailed test. This implies there is more consistency between the experts’ response for that category despite the grade of the mark. The graph in figure 8 shows it more clearly that participants recorded more consistently across all the categories. The response rate is high for this category as the online software prompts experts to record level one and two detail by annotating the mark and using confidence levels (different colours) during the annotations.

w = 0.43 is a medium to large effect size but is under powered at 0.23. To achieve a power of 0.80, the study would need 100 experts. However, recruiting this many experts is highly unlikely due to operational workload of completing examinations whilst gaining ISO 17025 accreditation. It must be noted that even if this number of participants were recruited, this would not necessarily produce a significant association between the checklist and grade of the mark as the online software will always prompt the participants to record each of these categories.



*Figure 8 – Participant response rate for the level of detail/annotations against the grade of the mark   
(using the online software)*

On review of the results, participants are more consistent when using the online software due to the prompts opposed to the before accreditation strategy which was a semi-structured approach. This coincides with the initial observations already reported in the chapter. The tests investigated significant associations between two categories; in this study, documentary suggestions from literature and policies recorded and the grade of the mark (grades 1-4), ranging from excellent to unusable. For most of the tests, there is no significant association between the two categories this suggests grade of the mark is not impacting the extent of notes recorded. There is consistency between experts’ response of the ACE-V checklist and grade of the mark.

The overall robust findings from this statistical analysis are that factors affecting the mark are having an impact on the quality of the mark, as discussed in section 3.2.3. It could be suggested that the factors should be recorded to suggest the quality of the mark, no matter what structural approach for the documentation strategy. There appears to be more factors recorded using this approach opposed to pre-accreditation, the explanation for this is the online software prompts the participants to record this expect. In addition, there is an association between the quality of the mark and the number of participants recording the factors observed for example, the excellent and good quality marks have little responses regarding the different factors affecting the mark and the poor and unusable marks have more responses. This finding is beneficial to the field as it implies what should be suggested in terms of the different quality of marks presented to experts on a day-to-day basis yet does not answer the aim of establishing effective documentation strategies. Therefore, the final stage of the research will use chi square tests for association to determine an association between obtaining a correct or incorrect outcome to a group of individuals (expert or novice), outcome to quality of the mark (grades 1 to 4) and outcome to documentation strategy. A new participant group took part in a study which contained a documentation strategy with an unknown impression. It was important to obtain a new participant group involving other fingerprint experts to assess the effectiveness of each strategy and therefore, not have bias within the results. If a participating bureau took part in the third study, it may have impacted the results as this group would be aware of the strategy. It was also important to involve novices within this study to determine whether the original notes could be understood by novices such as juror members and potentially use these original, contemporaneous notes in court rather than retrospective notes that were currently produced pre-accreditation. The participant was required to read the notes and look at the impression and determine whether the two matched, non-matched or unsure. As the ground truth was known, the accuracy of participants could be determined. A correct outcome will be the indicator of whether the documentation is effective. This was decided based on the term “effective” being used when successfully producing an intended or desired result. The intended or desired result was obtaining a correct outcome. Therefore, if an expert or novice could read, follow, and correctly identify a match between the documentation notes and an impression, this would suggest the documentation strategy was effective.

### 3.2.5 – Creation of a new documentation form “Mark Analysis Form”

During professional conversations with the participating bureaux, it is thought that contemporaneous notes will be time consuming, specifically the inclusion of marking up friction ridge detail and assigning confidence levels as seen in Langenburg & Champod (2011) and supported by Bunter (2016). Some bureaux are reluctant due to the lack of cases that go to court. Even though all cases may not go to court, there is still a need for contemporaneous notes, more so for cases where there is a difference of opinion between the original expert and the verifier. Section 3.5 investigates inter and intra examiner variance which highlights the importance of contemporaneous notes to understand the thought processes of an expert when there is a difference of opinion either within or between experts. If the examination is not recorded, then this difference of opinion cannot be investigated.

As previously explained in the earlier section of 3.2, the pre-accreditation documentation strategy only included the analysis outcome, anatomical source of which the unknown impression may have originated from, the level one detail i.e., pattern type and the geographical search area which could be police elimination database, local, regional or national database. The search area is usually case dependent. Therefore, to gain insight into the thought process of the expert, a free text box was included, and experts were instructed to produce contemporaneous notes regarding the decision-making process. No further instruction was provided. During the initial observations when the documentation was cross checked to best practice guidance (ACE-V checklist), the consensus of experts’ documentation and chi-square test for associations was evidence of what could be included within documentation strategies and suggested to a collaborating bureau.

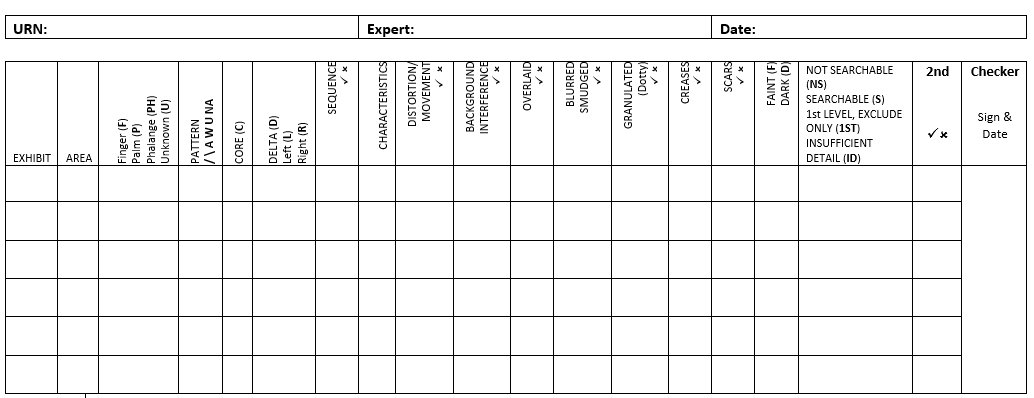
To define the consensus, based on previous research by Diamond *et al* (2014), the most common definition for consensus was percentage agreement with 75% being the median threshold. Therefore, any documentary suggestions that were recorded with a response rate >75%, it was suggested to be included in the potential documentation strategy. This included; unique reference number, date of examination, start and end time of the examination, anatomical source i.e., if and which finger or area of palm, if the impression was a single impression or multiple impressions, assessment of quality of the impression, distortion factors in particular substrate interference and overlapping impressions, level one detail present i.e., pattern type, level two detail present i.e., individual characteristics such as bifurcations and ridge endings, assigning confidence levels to the identified characteristics and inclusion of annotations of the identified characteristics. It must be noted that the consensus for the above documentary suggestions is from both documentation strategies. However, the decision-making notes from pre-accreditation, did not include the annotations and assigning confidence levels to the identified characteristics. The results from the chi-square tests for association demonstrated significant associations between the grade of the mark and generic information and more factors. Within the generic information category, the documentary suggestions include if finger or palm, which finger or palm, single or multiple and relationship between the marks. The unique reference number and date of examination was removed from the test due to the high response and potentially skewing the data. Within the more factors category, the documentary suggestions include presence of movement, substrate interference, pressure distortion, overlapping marks and development technique affects. When there is a significant association, it suggests there is less consistency in experts’ reporting the documentary suggestions over the different grades.

In addition, there is an observation within the notes implying that the grade of the mark is impacting the response rate for recording an element of generic information and more factors. The residual counts are used to investigate the associations. In the generic information, the “relationship between the marks” had the largest residual implying this documentary suggestion was recorded more times than expected. In the more factors category, “overlapping” and “movement” had the largest residuals. It is likely these two distortion factors had the higher response rates, when necessary, i.e., when an impression was affected by such, these factors are easy to identify than pressure distortion. These two distortion factors were most recorded in decision making of fingerprint practitioners in a study conducted by Earwaker (2017).

The categories with a significant association were strongly advised to be within the new documentation strategy when discussed with the senior fingerprint expert. The reasons for conclusions are likely to be within these categories i.e., identifying the anatomical source and the distortion factors impacting the impression. The consensus of experts was also presented as evidence of additional categories to include. It contains all documentary suggestions from the ACE-V checklist omitting location of the mark, orientation of the mark, identifying the matrix and determining its affects, and level one annotations.

It appears that participants recorded more information when prompted rather than free text. Therefore, during professional conversations, it was decided to create a form which included tick boxes and include the generic information and more factors. Generic information such as if finger or palm, sequence of marks i.e., relationship between marks. Level of detail and annotations included presence of level one detail which did not come up as a significant association within the results, but it was a part of the consensus and strongly recommended by the senior fingerprint expert therefore included into the form. Also, within this category is the number of characteristics observed and annotations, the consensus implied to record this information, however the senior fingerprint expert was against the marking up of minutiae due to time constraints but felt it necessary to include the number of characteristics to contribute to the reasons for conclusions. As seen by Ulery *et al* (2012), the operational threshold by an expert is 8 characteristics, and as there is no numeric standard in the UK, the senior fingerprint expert suggested to cap the minutiae at 10 and anything over this would be recorded as 10+. In the consensus, the assessment of mark quality was included supported by the reasons for the quality decision. However, the senior fingerprint expert stated that the assessment of quality would be implied by the distortion factors noted.

Therefore, there is no box relating to quality but there are the distortion factors individually to allow experts to record the presence of any. It also gives the expert opportunity to acknowledge the possibility of distortion factors within the impression.



*Figure 9 - Mark Analysis Form*

To adhere to ISO accreditation requirements, the form is currently used for every mark within a case at the bureau. The form prompts experts to record the following information about the mark;

* **Where the mark has originated from** - For example, finger, palm, phalange or whether it is unknown. Fingerprint experts will be familiar with this aspect due to its requirement within ident1 search, this can help to narrow down the search.
* **Pattern type** - Fingerprint experts will be familiar with this aspect due to its requirement within ident1 search, this can help to narrow down the search.
* **Whether the core and delta are present** - This information can help to determine the pattern type and show how the decision was made.
* **If the mark was a part of a sequence or not** - The spatial relationship between marks can help with identifying the finger or phalange.
* **Characteristics** - This refers to the number of minutiae observed. A total of 10 minutiae is recorded, any more is recorded as 10 +.
* **Factors affecting the mark** - Such as distortion/movement, background interference, overlaid, blurred/smudged, granulated. These factors are like the suggestions within the published literature and policies, yet different terminology is used. For example, overlaid refers to marks which appear to have been placed over each other, overlapped, double tapped. The continuity of ridge detail can be disrupted and affect the ability to identify minutiae. Blurred or smudged refers to marks that appear to have movement within it.   
    
  Granulated refers to the ridge detail broken or dotty as seen in ninhydrin marks. During discussions about the form, it was apparent that participants were more confident with these terms.
* **Creases/scars** - This is to indicate whether there is any third level detail present within the mark. This information can be unique to an individual and help to identify or eliminate suspects.
* **Faint/dark** - There could be a lack of ridge detail visible and difficult to interpret due to the pressure, whether this is faint or dark.
* **Analysis outcome** - Not searchable, searchable, 1st level exclude only and insufficient detail. The terms were decided by management at the bureau, this may vary between bureaux as seen in this study. However, each bureau will have a term for a mark which is insufficient for comparison, a mark which has only the pattern visible and used to exclude or sufficient detail to compare manually or on ident1.
* **Second checker** - If the analysis of the first checker is being reviewed or subjected to verification by a second checker, the second checker will sign and date the form and tick the box if in agreement. If there is a disagreement, the checker will put a cross in the box. A new mark analysis form will be completed by the second checker.
* **Free text box** - The form has a free text box for any additional information that experts feel pertinent to a mark.

The second checker will only sign the contemporaneous notes when an open verification has been completed and therefore the conclusions of the initial checker are made available for the second checker to review.

For a blind verification, the second checker will not see the initial checkers notes as the initial checkers conclusions from the process are not made available to the 2nd checker. They will complete a separate mark analysis form and do not sign the initial checker’s form. Blind verifications are done on the “lead” mark for each suspect identified. The lead mark is the first mark identified to one individual within the case. The other marks identified to the same suspect will be checked via “open” verification.

Blind verification is also used in cases whereby multiple marks are submitted and only one single mark from the case is suitable for comparison when the others are insufficient or 1st level detail only. For example, three marks in a case with a suspect. One of the marks reveals insufficient detail for comparison, another is 1st level detail only and the third is a searchable mark. The searchable mark is submitted with the suspect prints for blind verification. When this has been completed & returned to the initial Expert the QA the other marks are then handed to the 2nd checking Expert for the QA decision to be checked (open).

For all other occasions all checks are carried out as ‘open’ verifications, i.e., where the second checker is aware of the initial checkers conclusions at any stage of the ACE process.

In chapter 1.3, sub-section 1.3.3.2 an overview of ISO 17025 accreditation and the United Kingdom Accreditation Service is provided. UKAS is the organisation appointed by the government to assess and accredit organisations who provide services including certification, testing, calibration, and inspection. As discussed, this quality standard in part refers to the production of contemporaneous notes.

The form was adopted by the collaborating fingerprint bureau and used for three months prior to a UKAS accreditation visit to demonstrate working practices as part of the review.   
An overview of UKAS was previously provided in section 1.3.3.2. As a result of the visit, the form was accepted as a documentation strategy for recording contemporaneous notes. For this work, a letter of appreciation was given to the researcher from the collaborating fingerprint bureau. However, the form is only one documentation strategy and there are more strategies adopted within the UK. These strategies will be explored in section 3.3. Taking this into consideration, the mark analysis form and other UK strategies will be included into the effectiveness study, in section 3.4.

## 3.3 – Objective 5 to 7: The establishment of post-accreditation documentation strategies for Fingerprint Examinations; study two

This study follows a similar structure to study one in which the main findings focus on the consensus of fingerprint experts’ reporting outcomes and how this decision is recorded. However, the difference in this study is that it includes the reporting outcomes at the Analysis and Evaluations stage and additional documentation strategies as each bureau participating has adopted their own approach since the implementation of ISO 17025 accreditation. Again, these strategies can be categorised into semi-structured and structured. These new approaches will be outlined to give an understanding of the approaches which have been adopted.

The online digital software deemed as a structured approach is used again in this study as a means of documenting the whole ACE process. There were no modifications to the software following the first study, but participants were provided with a refresher session and informed of the tools required for the comparison and evaluation stage. The additional tools included producing a quality map for the known print and indicating the individual characteristics which were present on the mark. If there were any additional characteristics seen in the unknown mark after seeing the known print, then these were also marked up but in a different colour to see they were placed during the comparison stage. It may be argued that these have only been seen because of bias, experts must be cautious when using additional characteristics observed only after viewing the control print. Similarly, to study one, all documentation produced was cross-checked against the ACE-V checklist which included documentary suggestions from literature and policies. This list was created using the guidance for documenting fingerprint examinations as discussed in section 1.3 ([See appendix 3 for the full checklist).](#Appendix_3)

The results provided evidence of new approaches since the implementation for bureaux which had gained accreditation. Comparisons were made between semi-structured and structured approaches to determine the consistency of experts. Furthermore, chi-square tests for association were used to determine if there is an association between the ACE-V checklist and the grade of the mark. This is the same statistical test used in study one therefore previously explained in section 3.2.

### 3.3.1 – An overview of a selection of UK post-accreditation documentation strategies

Participants were required to complete a full examination of the same eight images used in study one. However, this study involved the new ISO 17025 accredited documentation strategy. All the bureaux and companies taking part were accredited, and as previously stated, this requires note-taking procedures to be put into place for day-to-day cases, not only court reports as previously seen. Despite all participants being accredited, there are clear differences in how the examinations are being recorded based on documentation strategy approach i.e., tick sheets or written notes, and the extent of the notes within these. UKAS agreed that the note-taking procedures were open to interpretation which explains the difference between forces, however the results from this thesis could suggest a more standardised approach that all UK fingerprint bureaux can adopt and work uniformly. If there was a standardised documenting procedure this would be particularly useful for defence experts as they would not be receiving different approaches and having to interpret the different kinds of notes. Furthermore, if the standardised documenting procedure was a computer software accessible for all bureaux, then this could reduce costs in the production of documentation as all approaches include additional paperwork whether this is forms to be completed or photographs being printed off.

The different approaches witnessed by the research team include producing notes on the back of a photograph, a standardised tick sheet form to complete, a free text form or the use of a software particularly for annotating the impressions. Table 7 shows the number of participant groups which have adopted the different approaches.

The most common approach appears to be a standardised tick sheet form. This approach has been suggested in other forensic disciplines such as handwriting analysis (ENFSI, 2020). The perceived benefits of a form are quick and easy to complete, however the extent of note taking within these tick sheets could still be insufficient for an auditable trail and to understand how the expert reached a decision, therefore the effectiveness of these should be investigated. The Forensic Science Regulator (2020) states that the documentation should be sufficient for an auditable trail, therefore this is the defining measure of effectiveness which is later used and detailed further in Chapter 3.4. Although three out of the seven participant groups are using a standardised form, these differ within each other. The inclusion criteria within these strategies will be explained throughout this chapter.

*Table 7 – Different types of post-accreditation documentation strategies*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Post accreditation documentation strategies | | | |
| Participant group | Back of photograph | Tick sheet | Free text form | Software |
| 1 | ü |  |  |  |
| 2 |  | ü |  |  |
| 3 |  |  | ü |  |
| 4 |  | ü |  |  |
| 5 |  |  | ü | ü (CSIpix) |
| 6 |  | ü |  |  |
| 7 | ü |  |  | ü (FSH) |

Strategy 2 is the mark analysis form created using study one results. As a result of the trends in the data, the chi-square test for association results and professional conversations with the senior fingerprint expert from a participating bureau, the tick sheet includes documentary suggestions as seen in the ACE-V checklist. The documentary suggestions include generic information, factors affecting the mark and the presence of level one, two and three detail. These suggestions are supported by CPS (2010), The Fingerprint Inquiry (2011), SWGFAST (2013), ENFSI (2015) and The Forensic Science Regulator (2020). There is a box for the analysis outcome to report either ‘sufficient for comparison’, ‘suspect comparison only’ or ‘insufficient for comparison’. If sufficient for comparison, there is an additional sheet for the comparison stage. Reporting the outcome is part of the minimum requirements from the Forensic Science Regulator (2020). Within the comparison sheet, the expert will record the name of the suspects who have been compared to the unknown impression, if the mark is identified, the anatomical source is recorded as well as the presence of a coincident sequence.

As suggested by CPS (2010), The Fingerprint Inquiry (2011), SWGFAST (2013) and The Forensic Science Regulator (2020). One important element to note here about this strategy is the inclusion of number of minutiae as a documentation strategy suggested by The Fingerprint Inquiry (2011), SWGFAST (2013) and ENFSI (2015). It does not include any annotations of impressions to show individual characteristics identified in the analysis stage and then in the comparison stage to show the coincident sequence, which has been observed in other strategies and suggested by Langenburg & Champod (2011) and Bunter (2016).

The form includes an additional notes page for any other notes that the expert may feel needs to be recorded, but this is not mandatory to complete. [See appendix 4 for an example.](#_Appendix_4_–)

Strategy 6 is another type of tick sheet which includes the presence of level one detail, digit determination and the orientation of the impression, these are suggested by SWGFAST (2013) and ENFSI (2015) followed by the analysis outcome suggested by The Forensic Science Regulator (2020). Immediately there is a difference on the quantity of notes taken between these two forms. If the impression is identified, then the identifier and anatomical source is recorded which conforms to suggestions by CPS (2010), The Fingerprint Inquiry (2011), SWGFAST (2013) and The Forensic Science Regulator (2020). An additional notes page is available for experts to record any notes, if necessary, but again there is no requirement to provide an explanation on how the comparison was made within this strategy. [See appendix 5 for an example.](#_Appendix_5_–)

Strategy 4 is the third type of tick sheet observed and has the least information recorded. For the analysis stage, experts are only required to record the outcome, not how the outcome was reached. This is very similar to the pre-accreditation documenting procedure whereby experts would record the outcome on the case management system.

As an outcome is reported, it meets the minimum requirements from the Forensic Science Regulator (2020) but the reasons for this conclusion are not always clear. If the impression is identified, then the identifier and anatomical source is recorded as suggested by CPS (2010), The Fingerprint Inquiry (2011), SWGFAST (2013) and The Forensic Science Regulator (2020). An additional notes page is available for experts to record any notes, if necessary, but as seen elsewhere, there is no requirement for this. [See appendix 6 for an example.](#_Appendix_6_–)

These tick sheets have been categorised as a structured approach as the experts are being prompted on what to record in their examinations. Similarly, to the online software, which has been used throughout these studies. There is other software available which have been witnessed by the research team to record the mark up of impressions and as evidence of documenting the characteristics in agreement or disagreement. The software is used in conjunction with the documentation strategy adopted by the force. For example, strategy 7 is a printed photograph of the unknown impression and the contemporaneous notes are written on the back of the photograph. When it is free text, it is the opinion of the expert to write as little or as much as possible. Therefore, there can be variability amongst experts but also variability in the notes depending on quality. Within strategy 7, the following was recorded at least once over the eight impressions by experts. Generic information including unique reference number, date of examination, if finger or palm, which finger or palm and single or multiple marks as suggested by CPS (2010) and the Forensic Science Regulator (2020). Factors affecting the mark including orientation, quality of the mark, presence of movement, overlapping, pressure, substrate interference and development technique effects as suggested by The Fingerprint Inquiry (2011), SWGFAST (2013) and ENFSI (2015). Level of detail and annotations such as presence of level one, level two and any other detail as suggested by SWGFAST (2013) and ENFSI (2015). The “any other detail” appeared within image 8 which was the palm impression. The presence of creases was recorded by only one expert. Outcomes at analysis and evaluation were reported which the subsequent identifier and anatomical source as suggested by the Forensic Science Regulator (2020). As stated, the use of a software can be used within this bureau to annotate impressions and show features in agreement. However, due to the nature of the software, research images could not be uploaded and therefore have been omitted from the documentation examples. [See appendix 7 for an example of this strategy](#_Appendix_7_–).

Strategy 1 is like strategy 7 as it includes text on the back of a photograph. Again, there can be variability amongst experts but also variability in the notes depending on quality.

Notes are seen for impressions deemed as ‘sufficient’ or ‘suspect comparison only’ at analysis. However, when the mark is deemed as ‘insufficient’, this outcome is reported, and no further notes are provided. Having this approach does not explain the reasons for the conclusion and does not comply with documentary suggestion by CPS (2010), The Fingerprint Inquiry (2011), SWGFAST (2013) and The Forensic Science Regulator (2020). Although, the documentation procedure is accepted by UKAS. Within this strategy, the following was recorded at least once over the eight impressions by experts. Generic information including unique reference number, date of examination, if finger or palm, which finger or palm and single or multiple marks as suggested by CPS (2010) and the Forensic Science Regulator (2020). Factors affecting the mark including what the matrix is (although this was only reported once), location of the mark, orientation, quality of the mark, presence of movement, overlapping, pressure, substrate interference and development technique effects as suggested by The Fingerprint Inquiry (2011), SWGFAST (2013) and ENFSI (2015). Interestingly, not recorded for marks that were deemed as insufficient. Within other strategies, these factors affecting the mark were recorded more frequently in the insufficient (grade 4) marks. The orientation of the mark was indicated by an arrow drawn next to the impression. Level of detail and annotations such as presence of level one, number of characteristics and features in agreement as suggested by SWGFAST (2013) and ENFSI (2015). This bureau would “prick out” the characteristics within the impression on the photograph during the analysis. Followed by reporting features in agreement but these are not shown. The analysis outcome if ‘sufficient for comparison’ or ‘suspect comparison only’ was not stated, as the identifier and anatomical source were used to imply progression to comparison and evaluation. As stated, the analysis outcome ‘insufficient for comparison’ would be recorded. The evaluation outcome “identified” is implied by the identifier and anatomical source. All other outcomes would be recorded. This complies with The Forensic Science Regulator (2020) requirements. [See appendix 8 for an example](#_Appendix_8_–).

Another documentation strategy adopted is free text on a word document rather than the back of a photograph. The strategies which involve free text have devised an approach to help with the time producing contemporaneous notes. This involves using abbreviations and all experts within the force are aware of the different abbreviations of what can be used and the meanings for each abbreviation. Strategy 1 and 7 were dependent on the expert, which shows the flexibility of note taking in the bureaux. However, strategy 3 used this more frequently, there was only one participant from this group though. [See appendix 9 for an example](#_Appendix_9_–). One issue with the use of abbreviations is another expert, in particularly the defence expert, may not understand what these abbreviations refer to. The notes must be combined with a glossary of the abbreviations.

Within strategy 3, the following was recorded at least once over the eight impressions by the expert. Generic information such as unique reference number and date of examination marks as suggested by CPS (2010) and the Forensic Science Regulator (2020). The presence of single or multiple marks was recorded but only for image 2 when there was a sequence of marks. This strategy included an illustrations box to show the orientation of the impression. Factors affecting the mark such as movement, pressure and substrate interference were reported as suggested by The Fingerprint Inquiry (2011), SWGFAST (2013) and ENFSI (2015). These were reported more in the poor-quality marks. The presence of level one and level two were recorded however, level one was recorded more so than level two. Similarly, to strategy 1, the analysis outcome was reported only for an insufficient mark. The identifier and anatomical source reported would suggest progressing to the comparison and evaluation stage.

The final strategy to discuss and the only documentation taken by a defence expert is strategy 5. These are handwritten notes with the addition of potentially using software to indicate the characteristics identified during the analysis and comparison. Of the free text strategies, this is the only strategy to clearly differentiate between each stage, analysis, comparison and evaluation.

Within the notes, there is presence of generic information such as unique reference number and date of examination is recorded as suggested by CPS (2010) and the Forensic Science Regulator (2020). The orientation and quality of the mark however, the factors affecting the mark such as pressure and movement were rarely recorded. In relation to the level of detail and annotations, the presence of level one was recorded but this strategy stood out from the rest in particularly due to the number of characteristics being recorded at analysis. The expert provided confidence levels regarding the characteristics present by stating “clear” ridge characteristics and “indicative” ridge characteristics. These were later confirmed during the comparison and highlighting the features in agreement which is supported by The Fingerprint Inquiry (2011), SWGFAST (2013) and ENFSI (2015). The outcome at analysis and evaluation was recorded as suggested by the Forensic Science Regulator (2020) [see appendix 10 for an example.](#_Appendix_10_–)

The forces which have documentation strategies using free text including abbreviations and/or software this has been deemed as semi-structured, as they are to some extent prompted with the abbreviations and marking up the characteristics. It must be noted that the extent of note taking is different depending on the grade of the mark. This is explored further within this chapter by observing trends in the data and using chi-square tests for association between documentary suggestions and the grade of the mark. The effectiveness of all the strategies from study one and study two is assessed in chapter 3.4.

### 3.3.2 – Consensus of reporting outcomes during ACE-V process

The outcomes that have been reported during the ACE process in the second study are taken only from the documentation that was produced by participants in the study, it must be noted that these notes may be in conjunction with the case management system as previously used before accreditation. This has not been made clear to the research team. The possible analysis outcomes were explained in section 3.2.1. These terms have been used again in this study. Table 8 shows the number of participants which recorded the analysis outcome in the post-accreditation documentation strategy.

Out of 33 participants, none of the images had all participants’ analysis outcome reported. This suggests that is not common practice within all bureaux to record this outcome on the form or photograph, however it does appear that most bureaux will record the outcome if it is insufficient. If the impression is deemed sufficient for ident1 search or only suitable for comparison purposes, then the additional notes suggest one of these outcomes rather than stating the analysis outcome. The higher response in ‘suspect comparison only’ seen here in comparison to the previous study is more than likely due to the design of this study. Participants were provided with ten-print forms, like elimination or suspect prints within casework, therefore this would be deemed as comparing against a suspect.

In the new documentation strategies, the evaluation outcome is more likely to be recorded which is consistent across all bureaux. There was little variance of opinion between experts for the evaluation outcomes which highlights consistency between experts. As expected, and seen in study one, there is less consistency between experts for the poor-quality marks (grade 3 and 4). Refer to table 9 for the evaluation outcomes in post-accreditation documentation strategies. There is a grade 1 mark that resulted in insufficient and inconclusive, this was due to the poor-quality of the ten-print form rather than the unknown impression. This is evident from the notes that were produced by the participants and demonstrates the importance of note keeping as it is more likely that justifications can be understood. More importantly, if the case was taken to court, which can be months after the original examination, the notes would be useful to understand the thought process at the time of the examination and a new set of ten prints could be requested to carry out the examination again.

There are significantly fewer participants who completed the second part of the study which involved the use of the structured approach (online digital software) to document to ACE process. There was a drop in participants because of workload within the bureaux and independent companies, especially with the accreditation deadline approaching. The sample size should have been higher for the purpose of the chi-square tests to achieve significant associations between the documentation and the grade of the mark but as previously seen and explained in study one, the “correct” number of participants might not achieve a significant association due to the nature of the study. The research team acknowledges the low number of participants used is a limitation but still believes it is appropriate to include this strategy as a means of documenting the ACE process. The structured approach requires participants to record the analysis outcome before continuing to the comparison stage or the next image (if the impression is deemed insufficient for comparison), this explains the consistency between participants recording this aspect. Refer to table 10 for the analysis outcomes using the structured approach. The immediate similarities can be seen in the poor-quality marks as all participants recorded the impression was insufficient when using the structured approach as seen in the post-accreditation documentation strategy. The inconsistencies between experts are apparent in grades 4 when using the structured approach, refer to table 11. Participants were provided with the correct known mark during the comparison when using the structured approach. The outcomes should either be identified because there are enough characteristics in agreement, insufficient if the known print is poor quality and a decision cannot be made, or inconclusive if there is not enough characteristics in agreement or disagreement. Not excluded.

*Table 8 – Analysis outcomes in the post-accreditation documentation strategy*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Image | Grade | Sufficient | Suspect Comparison Only | Insufficient |
| 1 | 1 | 11 | 4 | - |
| 2 | 2 | 10 | 6 | - |
| 3 | 4 | 9 | - | 12 |
| 4 | 1 | 9 | 5 |  |
| 5 | 4 | - | - | 31 |
| 6 | 3 | 11 | 5 |  |
| 7 | 2 | 10 | 5 |  |
| 8 | 1 | 9 | 5 | - |

*Table 9 – Evaluation outcomes in the post-accreditation documentation strategies*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Image | Grade | Identified | Excluded | Insufficient | Inconclusive |
| 1 | 1 | 33 | - | - | - |
| 2 | 2 | 33 | - | - | - |
| 3 | 4 | - | 2 | 3 | 15 |
| 4 | 1 | 33 | - | - | - |
| 5 | 4 | - | - | - | - |
| 6 | 3 | 31 | - | - | 2 |
| 7 | 2 | 30 | - | 1 | 2 |
| 8 | 1 | 31 | - | 1 | 1 |

There were several instances where experts had reported “excluded” this was also seen in the post-accreditation strategies. This is where it would be useful to have contemporaneous notes to obtain information on the thought process of the original examiner reaching that decision. Section 3.5 addresses the inter and intra examiner variance in reporting outcomes during analysis and comparison/evaluation stage.

Pre-accreditation documentation strategies involved case management systems which would prompt experts to record the analysis outcome. However, the results indicate that not all bureaux are recording the analysis outcome within their post-accreditation strategy, unless it is recorded on the case management system still.

It is suggested in literature and policies to record the outcome as previously explained in section 1.3. There is no evidence within this study to suggest it affects the auditable trail and that it should be recorded, this will be assessed in the effectiveness study in section 3.4.

*Table 10 - Analysis outcomes using the structured approach*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Image | Grade | Sufficient | Suspect Comparison Only | Insufficient |
| 1 | 1 | 13 | - | - |
| 2 | 2 | 13 | - | - |
| 3 | 4 | 3 | 8 | 1 |
| 4 | 1 | - | 12 | - |
| 5 | 4 | - | - | 11 |
| 6 | 3 | - | 11 | - |
| 7 | 2 | - | 8 | - |
| 8 | 1 | - | 9 | - |

*Table 11 – Evaluation outcomes using the structured approach*

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Image | Grade | Identified | Excluded | Insufficient | Inconclusive |
| 1 | 1 | 13 | - | - | - |
| 2 | 2 | 12 | 1 | - | - |
| 3 | 4 | - | 6 | - | 5 |
| 4 | 1 | 12 | - | - | - |
| 5 | 4 | - | - | - | - |
| 6 | 3 | 11 | - | - | - |
| 7 | 2 | 8 | - | - | - |
| 8 | 1 | 6 | 1 | - | 2 |

The evaluation outcome is expected to be recorded as it’s the outcome of the examination, and as previously seen established procedures i.e., streamlined forensic reporting and in infamous cases such as R-v-Smith it is the only information to be recorded. The purpose of this research is to determine what supporting information goes alongside the outcome for a sufficient auditable trail, in other terms, enough information for another expert to follow. The upcoming sections address what documentary suggestions are being recorded now, post-accreditation, and whether there is more consistency between experts and if there is an association between the level of documentation and the grade of the mark.

If there is an association, this may provide evidence to support the use of a different documentation procedure for complex marks as seen in the Codes of Practice and Conduct for Fingerprint Comparisons (2020).

### 3.3.3 – Initial observations and chi-square tests for association for “Generic information” from the ACE-V checklist

One of the categories of the ACE-V checklist is “generic information”. This includes the unique reference number which refers to the case (CPS, 2010, Forensic Science Regulator, 2017). The date, start and end time of the examination to allow for an auditable trial of the case (CPS, 2010). If the case needed to be reviewed, it can be searched using information such as URN and date. The materials used throughout the examination could be recorded to allow for the verifier to have the option to use the same materials as the original expert (CPS, 2010, Forensic Science Regulator, 2017). Following this, participants may record if the mark is originated from the finger or palm and which finger or area of palm is present. This information is required when a mark is sufficient for an ident1 search or can aid with the ‘suspect comparison only’ marks. Additional generic information about the mark is suggested by ENFSI (2015) such as if the mark is single or multiple marks present and if there is more than one mark present, participants may record the relationship between the marks present. The spatial relationship may indicate the marks are part of a sequence. Table 12 shows the response rate of recording these documentary suggestions using the new documentation strategy (left figure) and the structured approach (right figure). The grades for each image are also provided to show the relationship between the quality of the mark and what is recorded.

**Initial observations**

In relation to the generic information, when observing what is recorded for this category there is little difference between pre and post accreditation. The same sub-categories are frequently recorded across all images which are unique reference number, date, if finger or palm, which finger or palm and if the impression is single or multiple. This information was previously recorded in the case management system and as explained in literature and policies are mental observations and should be recorded to show the thought process of examiners, especially with the mandatory move towards ISO accreditation (The Fingerprint Inquiry, 2011, SWGFAST, 2013, ENFSI, 2015). There does appear to be a drop in the numbers recording if finger or palm and which finger or palm at the analysis stage. It is primarily recorded in the evaluation stage once an impression has been identified.

Occasionally, the start time of the examination and the materials used are recorded. Although, the Forensic Science Regulator (2017) stated recording the materials used throughout the examination must be included in the documentation. The purpose of this is for another expert to repeat the examination using the same materials as the original expert as features may be missed during the verification stage, if the full equipment list is not used in the same way as the original examination.

As explained, the online software prompts the participants to record if finger or palm is present and which finger or palm is present which explains the consistency between this category, however the other sub-categories are not asked of experts in the software which explains the lack of response.

The Chi square test for association has been used to determine if there is an association between the documentary suggestions and the grade of the mark. There are individual tests for the data referring to post-accreditation documentation and the structured online software.

If there is no significant association this implies there is consistency between the expert’s response for that category despite the grade of the mark, in other terms, there is no relationship between a documentary suggestion and the quality of a mark. However, if there is a significant association there is a relationship between the documentary suggestion and a particular grade/quality mark.

*Table 12 - Response rate for generic information in both approaches.*

A picture containing text, shoji, crossword puzzle

Description automatically generated

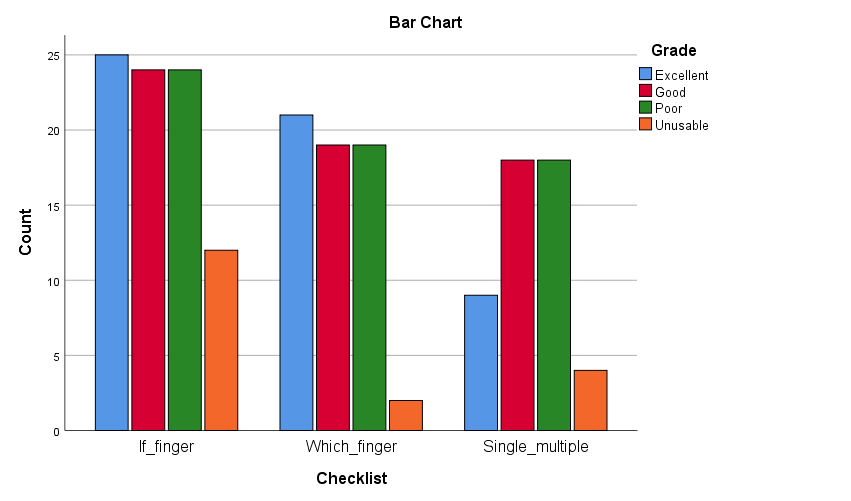
The standardised residual counts are used to determine where this relationship falls. The results can be used to suggest that the categories with a significant association are essential within the documentation and evidence to be included on a day-to-day basis.

The Chi square test for association has been used to determine if there is an association between the documentary suggestions and the grade of the mark. There are individual tests for the data referring to post-accreditation documentation (after) and the online software.

If there is no significant association this implies there is consistency between the expert’s response for that category despite the grade of the mark, in other terms, there is no relationship between a documentary suggestion and the quality of a mark. However, if there is a significant association there is a relationship between the documentary suggestion and a particular grade/quality mark. The standardised residual counts are used to determine where this relationship falls. The results can be used to suggest that the categories with a significant association are essential within the documentation and evidence to be included on a day-to-day basis.

**Chi-square test for association weighted by after data: Generic information.**

The chi square test indicated no significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting the generic information during the examination after accreditation, ꭓ2 (6) =8.600, p=0.107, two-tailed test. This implies there is consistency between the experts’ response for that category despite the grade of the mark. This is explainable as all these categories are a requirement either by guidance from the Forensic Science Regulator or prompted during an Ident1 search, despite the grade of the mark. However, it can be expected that if finger or palm and/or which finger or palm are not recorded for the poor-quality marks such as grade 4 as this detail is not usually present. Refer to figure 10 to see the drop in participant response for this grade and categories. Also seen in before and after data.   
w = 0.54 is a large effect size but is under powered at 0.54. To achieve a power of 0.80, 55 participants are required. It must be noted that even if the correct number of participants were recruited, this would not necessarily produce a significant association between the checklist and grade of the mark if experts are expected to record these aspects of the checklist.

*Figure 10 - Participant response rate for the generic information against the grade of the mark   
(after accreditation)*

**Chi-square test for association weighted by online software data: Generic information.**

The Chi square test indicated no significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting the generic information during the examination using the online software, ꭓ2 (3) =0.000, p=1.000, two-tailed test. This implies there is consistency between the experts’ response for that category despite the grade of the mark. This is explainable as all these categories are a requirement of the online software.

As the effect size was less than small, the power of the study could not be calculated correctly. However, when using the effect size tables (Clark-Carter, 2010) it is suggested that the study is under powered as to achieve 0.80, 55 participants were required for a significant association. This study had 13 participants, not 55. As previously explained throughout, if the optimal sample size was recruited it may not have created a significant association as the software prompts the participant to record the presence or no presence of the categories above.

### 3.3.4 – Initial observations and chi-square tests for association for “Factors affecting the mark” from the ACE-V checklist

This category relates to the recording of information about the position of the mark, the mark quality and quantity and subsequently what is affecting the quality of the mark, if any.

This could be substrate interference, development technique, pressure applied during deposition, movement and overlapping of marks as suggested by The Fingerprint Inquiry (2011), SWGFAST (2013) and ENFSI (2015).

At a crime scene, it is unlikely to have a perfect mark i.e., a complete impression in terms of surface area with no areas of distortion. Fingerprint experts will be familiar with marks that are affected by the above factors affecting ridge detail. These factors could be the reason why a mark is deemed insufficient for comparison or inconclusive. It appears that this information is not routinely recorded (Bunter, 2016) but is suggested in certain literature and policies (The Fingerprint Inquiry, 2011, SWGFAST, 2012, ENFSI, 2015). In the previous study, the consensus of experts demonstrated the assessment of mark quality, and some distortion factors were used in the decision-making process and recorded within the notes. This study assessed the likely change of the content in documentation since adopting new strategies post-accreditation. This is novel research as there is currently no evidence to suggest what documentation strategies have been introduced since gaining accreditation within UK fingerprint bureaux.

Table 13 shows the response rate of recording these documentary suggestions using the new documentary strategy (left figure) and the structured approach (right figure). The category “location of the mark” has only two responses across the eight images, using both approaches and only one response across the eight images for the category “what the matrix is”. The response rate for “location of the mark” is higher than pre-accreditation documentation. Although it is only two occasions, one participant for two different impressions, it is two more than previously seen in the pre-accreditation documentation and it demonstrates experts may record this information if pertinent. Bunter (2017) supports studying the location of the impression can aid with interpretation by explaining how the location and orientation of fingerprints on a security gate showed that it was virtually impossible for the fingerprints to have been placed during the offence. It was suggested that contact was more likely to have occurred before the gate was put in place, ten years earlier. This shows the importance of recording the location of the mark for such cases as this.

In addition, the matrix was not recorded. “Matrix” refers to what the mark is made up of, or left in. This is the substance that is deposited by the finger and eventually developed such as sweat or blood (The Forensic Science Regulator, 2013). All test marks were sweat marks, but this was not mentioned by any of the participants. However, if there was a mark in blood, there is potential for this information to be recorded due to The Forensic Science Regulator (2017) stating a mark in blood is deemed as a “complex” mark. As previously mentioned, a complex mark may have a different documentary process, therefore the presence of a mark in blood would be documented. Despite neither of these documentary suggestions have been used in the test marks, there is evidence to suggest that there will be cases in which they are required, so should not be disregarded.

One immediate difference observed between pre- and post-accreditation documentation strategies is the documenting of the orientation of the mark. Almost all the strategies now include this as an aspect to record, despite the type of strategy whether is semi-structured or structured. The use of arrows indicates how the impression is oriented, this is useful for verifying experts to observe the impression in the same way as the original examiner.

It appears that more factors affecting the mark are taken into consideration when producing the documentation post-accreditation. More specifically in the documentation strategies which prompt the experts to record the presence of the factors affecting the mark as seen in one of the strategies.

Experts are more likely to record the quality of the mark post-accreditation opposed to pre-accreditation strategies. There appears to be a trend between grade 1 impressions only stating a “good quality mark”, with no factors affecting it as expected then grades 3 and 4 impressions the “poor-quality” is not recorded, but the factors which are affecting the impression is recorded implying it is a poor-quality mark. However, it must be noted that not all strategies will require the factors affecting the mark to be recorded, this is evident in the number of responses. It is common practice in at least two of these strategies to only record that an impression, particularly grades 4, are insufficient for detail, then no explanation is given. This is the reverse to what was seen in pre-accreditation documentation strategies. As implied in the R-v-Smith case, the expert reported a one sentence report stating the impression was identified to the suspect and no explanation provided of how this conclusion was reached. In this post-accreditation documentation strategy, it is seen that the impressions of poor quality are deemed insufficient for comparison, are given this conclusion with no explanation. Therefore, is this sufficient for an auditable trail?

*Table 13 - Response rate for factors affecting the mark in both approaches.*

A picture containing calendar

Description automatically generated

The structured approach has similar results as seen in study one due to no changes to the software. Participants are prompted to record the quality and quantity of the mark when using the quality map tool which explains the high consistency between experts here.

Overall, between pre- and post-accreditation documentation strategies the only differences are the inclusion of the orientation of the mark and the quality of the mark is stated. Some forces have adopted approaches which prompt experts to record they have considered the possibility of factors affecting the mark which has increased the numbers for this aspect opposed to pre-accreditation. The chi square tests are used to determine if there is a relationship between the categories and the grade of the mark.

**Chi-square test for association weighted by after data: Factors affecting the mark.**

The Chi square test indicated no significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting the factors affecting the mark during the examination after accreditation, ꭓ2 (3) =1.395, p=0.707, two-tailed test. This implies there is consistency between the experts’ response for that category despite the grade of the mark. This suggests that participants are recording the orientation of the mark and the quality of the mark despite the grade of the mark.

As it may be expected, the orientation of the mark is less likely to be recorded for grade 4 marks (unusable) as the mark isn’t clear to orientate. As for the quality of the mark, there appears to be no trend, as this is a category that can be recorded whether it’s an excellent mark or unusable mark. The results suggest it is most likely to be recorded for a poor-quality mark, this may be due to the more likely factors affecting the mark are present but still enough detail to exclude perhaps.

w = 0.21 is a small to medium effect size but is under powered at 0.11. To achieve a power of 0.80, 150 participants are required. It must be noted that even if the correct number of participants were recruited, this would not necessarily produce a significant association between the checklist and grade of the mark if experts are expected to record these aspects of the checklist.

**Chi-square test for association weighted by online software data:   
Factors affecting the mark.**

The Chi square test indicated no significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting the generic information during the examination using the online software, ꭓ2 (3) =0.000, p=1.000, two-tailed test. This implies there is consistency between the experts’ response for that category despite the grade of the mark. This is explainable as all these categories are a requirement of PiAnoS. Users of the software are asked to produce a quality map using colours like the GYRO system (Langenburg & Champod, 2011). Under GYRO, the quality of the mark is indicated by using green for high quality, yellow as medium quality and red for low quality. Orange is used during the comparison stage if the user identifies other additional features. The user will cover the areas in which they believe can and cannot be used which shows the quantity of the mark.

The effect size calculated to 0, which is below low effect size, this led to 0 for the power analysis. This may be due to the low number of participants. However, as previously seen throughout the results and discussed, the correct number of participants may never result in a significant association due to the nature of the software i.e., the quality and quantity of the mark will be required despite what grade the mark is.

**Chi-square test for association weighted by after data: More factors.**

The Chi square test indicated no significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting more factors during the examination after accreditation, ꭓ2 (12) =6.102, p=0.911, two-tailed test.

This implies there is consistency between the experts’ response for that category despite the grade of the mark. This suggests that there is no factor within the list which is associated with a grade.

w = 0.43 is a medium to large effect size but is under powered at 0.28. To achieve a power of 0.80, 100 participants are required. It must be noted that even if the correct number of participants were recruited, this would not necessarily produce a significant association between the checklist and grade of the mark if experts are expected to record these aspects of the checklist.

**Chi-square test for association weighted by online software data: More factors.**

The Chi square test indicated no significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting more factors during the examination using the online software, ꭓ2 (12) =18.108, p=0.112, two-tailed test. Again, the results imply there is consistency between the experts’ response for that category despite the grade of the mark. This suggests that there is no factor within the list which is associated with a grade.

w = 1.28 is a large effect size but is under powered at >0.80. To achieve a power of 0.80, 20 participants are required. The number of participants would have been achievable however, all the other tests have indicated 100+ participants were required which would not have been achievable. In contrast to the pre-accreditation study results, there was a significant association between the documentary suggestions and the grade of the mark for this category (more factors). This implies there is less consistency between the experts’ response for that category, across all grades of the marks. Even though the online software encourages experts to record the different factors affecting a mark, there appears to be inconsistencies between what experts are identifying as affecting the mark across the different grades. This can be expected as a grade 1 mark will have minimal distortion whereas a grade 4 will have excessive amount of distortion to the extent of non-identifiable pattern and ridge characteristics. The highest residual is in the category pressure for a poor grade which was 3.9, a lot more people recorded it than expected. When observing the expected counts, there is a clear increase, it suggests experts are more likely to record factors affecting the mark (movement, pressure, substrate interference, overlapping and development technique) for the poor and unusable marks. From the results, participants are most likely to record the substrate interference over the other factors, this is potentially due to the ease of identifying this as a factor opposed to pressure applied during deposition or development technique affects. However, in a previous study conducted by Earwaker (2017) the terms background or substrate where not recorded, only 4 occasions whereby ‘interference’ was recorded which could be relating to background/substrate interference.

### 3.3.5 – Initial observations and chi-square tests for association for “Level of detail and annotations” from the ACE-V checklist

The final category for the analysis stage relates to a significant aspect to fingerprint examinations, identifying fingerprint patterns and ridge characteristics. Literature and policies suggest that the presence of level one (SWGFAST, 2012, ENFSI, 2015), level two and other features (The Fingerprint Inquiry, 2011, SWGFAST, 2012, ENFSI, 2015) could be recorded within contemporaneous notes. This information is ultimately what is used to make a comparison between an unknown mark and a known mark which shows its importance within the original experts’ contemporaneous notes. Level one refers to the fingerprint pattern which could be arch, loop or whorl. The experts would be required to input this information into ident1 when searching the database. This explains the high consistency between expert’s response for this element during the semi-structured approach. Table 14 shows the response rate for the level of detail and annotations recorded during the post-accreditation documentation strategies and the structured approach.

**Initial observations**

There appears to be high consistency between experts recording the level one detail present and the analysis outcome for only the insufficient (grade 4) impressions. The other categories have very low consistency and little change since the adoption of the new documentation strategy within all bureaux. One bureau has included annotating the photograph by pricking out the level two detail. The method of “pricking out” characteristics has been discussed by Bunter (2016). It is used to show where the expert believes usable characteristics are visible on the unknown mark and would use during the comparison stage. However, many other bureaux have not adopted such approach as it was stated that the characteristics in an unknown mark may not always be visible on a known print therefore cannot be used. It may be deemed as a “safer” approach to not record these characteristics and annotate the impressions once completing a side-by-side comparison, however this can potentially cause bias. Using a system like the GYRO system can be useful to indicate what characteristics have been observed during the Analysis stage, and any other characteristics during the Comparison stage.

The decision of the expert can be easily explained in court if using this approach. This is used in the online software which explains the consistency between experts recording the level two detail, the number and annotations. Other bureaux have included how many features have been observed during the Analysis within their form but not documented the location of these characteristics. It may be argued that the inclusion of only the number of features is a pointless exercise, as the exact location of these are not documented anywhere. This will be assessed in the effectiveness study.

The analysis outcome was always recorded by participants in the pre-accreditation documentation as it was recorded on the case management system, this may still be the case, but it is not always recorded for grades 1 to 3 impressions on the written notes.

*Table 14 - Response rate for level of detail and annotations in both approaches.*

A picture containing calendar

Description automatically generated

**Chi-square test for association weighted by after data: Level of detail and annotations.**

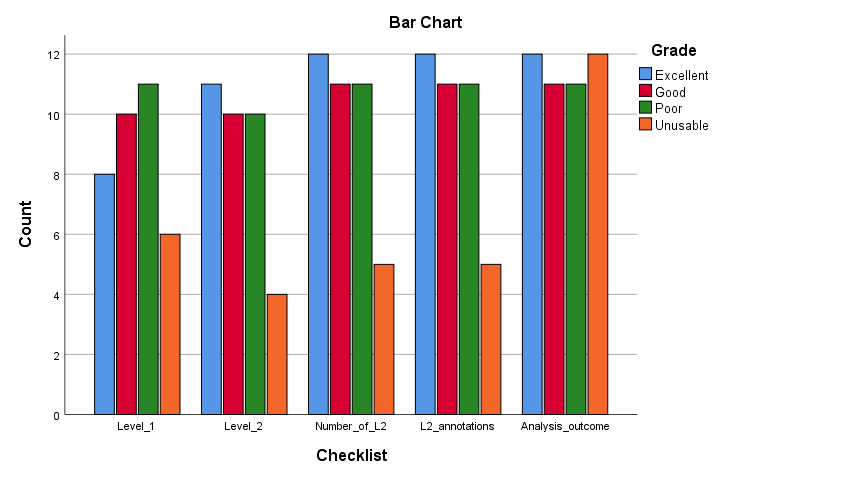
The Chi square test indicated a significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting the level of detail and annotations during the examination after accreditation, ꭓ2 (12) =41.555, p<0.001, two-tailed test. This implies there is less consistency between the experts’ response for that category, across all grades of the marks. The inconsistencies could possibly stem from certain fingerprint bureaux now recording this information and others not recording this information. Also, it may be expected that level one and level two detail are more likely to be recorded if the mark is excellent or good quality. The highest residual is in the category level 2 annotations for an unusable mark and analysis outcome for a good mark which were both -2.6, refer to figure 11. The counts were significantly lower than what is expected. It is surprising to see the lack of reporting for analysis outcome over the different grades as each grade would have an outcome to report. As there is no set standard in the UK, this explains the inconsistencies between participants despite the grade of the mark. w = 1.12 is a large effect size and is overpowered at 0.97. Only 20 participants were needed to achieve a power of 0.80. It is a robust finding that the grade of the mark has an impact on the recording of the level of detail and annotations.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Checklist \* Grade Crosstabulation** | | | | | | | |
|  | | | Grade | | | | Total |
| Excellent | Good | Poor | Unusable |
| Checklist | Level\_1 | Count | 26 | 25 | 28 | 10 | 89 |
| Expected Count | 24.5 | 20.4 | 26.5 | 17.6 | 89.0 |
| Standardized Residual | .3 | 1.0 | .3 | -1.8 |  |
| Level\_2 | Count | 7 | 5 | 4 | 5 | 21 |
| Expected Count | 5.8 | 4.8 | 6.2 | 4.2 | 21.0 |
| Standardized Residual | .5 | .1 | -.9 | .4 |  |
| Number\_of\_L2 | Count | 6 | 5 | 6 | 3 | 20 |
| Expected Count | 5.5 | 4.6 | 5.9 | 4.0 | 20.0 |
| Standardized Residual | .2 | .2 | .0 | -.5 |  |
| L2\_annotations | Count | 9 | 12 | 13 | 0 | 34 |
| Expected Count | 9.3 | 7.8 | 10.1 | 6.7 | 34.0 |
| Standardized Residual | -.1 | 1.5 | .9 | -2.6 |  |
| Analysis\_outcome | Count | 13 | 4 | 15 | 26 | 58 |
| Expected Count | 15.9 | 13.3 | 17.2 | 11.5 | 58.0 |
| Standardized Residual | -.7 | -2.6 | -.5 | 4.3 |  |
| Total | | Count | 61 | 51 | 66 | 44 | 222 |
| Expected Count | 61.0 | 51.0 | 66.0 | 44.0 | 222.0 |

*Figure 11 – Standardised residual counts highlighted to show significant association*

**Chi-square test for association weighted by online software data:   
Level of detail and annotations.**

The Chi square test indicated no significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting the level of detail and annotations using structured approach, ꭓ2 (12) =5.177, p=0.952, two-tailed test. This implies there is consistency between the experts’ response for that category despite the grade of the mark. This is explainable as experts are prompted to record this information on the online software. When observing the graph, see figure 12, it is less likely to record the level 1 and 2 detail along with annotations for the unusable marks. The quality is so poor that these features cannot be observed. The analysis outcome is always asked of participants despite the grade, the fall in numbers comes from the number of participants completing that image to analyse rather than not recording this information.   
w = 0.69 is a large effect size but is under powered at 0.39. To achieve a power of 0.80, the study would need 45 experts. It must be noted that even if this number of participants were recruited, this would not necessarily produce a significant association between the checklist and grade of the mark if people do not record these categories ever, or very often.



*Figure 12 - Participant response rate for the level of detail and annotations against the grade of the mark (after accreditation)*

### 3.3.6 – Initial observations and chi-square tests for association for “Comparison and evaluation notes” from the ACE-V checklist

The previous categories of the checklist which have been described relate to the Analysis stage and therefore, seen in both studies. A comparison has been made of what was recorded pre-accreditation and the post-accreditation strategies in relation to these categories. This is the first occurrence of the comparison and evaluation notes within the thesis. Refer to table 15 for the response rate of comparison and evaluation notes in both approaches.

As explained, the comparison stage involves a side-by-side comparison of the unknown impression to a known print either from a database or ten-print form from eliminations or suspects. If there are enough features in agreement, then the expert will report this as an identification. If it is identified, the person of whom it belongs to, and the anatomical source should be recorded as the evaluation notes. As previously explained, there can be additional features seen during the comparison stage which aid with the conclusion, and it is debated whether this should be recorded or not (Champod & Langenburg, 2010). All the documentation produced post-accreditation was analysed to determine which of these categories are recorded.

As expected, there is high consistency between experts recording the outcome and subsequently the identifier (who the impression belongs to), the anatomical source (which finger or area of palm) and a copy of the reference prints which were compared. Grades 1 to 3 were deemed sufficient for comparison during the Analysis stage and due to the quality of the unknown impression and the known prints, an identification was made. However, there is approximately half of the participant group which explains how this conclusion was made using features in agreement. These features in agreement were either stated on the back of the photograph or within the form, dependent on strategy. Only two of the participant groups visually showed the features in agreement using software such as CSIpix or FSH and taking a photo of the screen. It was very uncommon for experts to provide additional notes explaining the conclusion. The Forensic Science Regulator (2020) stated that reasons for conclusions must be documented as part of the minimum requirements for documentation. It is evident in the results that not all participants recorded their reasons for the conclusion, and only stated the outcome. However, the notes made prior in the analysis stage may be enough for another expert to follow and see how the overall decision was made. This will be assessed in the effectiveness study.

*Table 15 - Response rate for comparison and evaluation notes in both approaches.*

Calendar

Description automatically generated

That said, in section 12 of Codes of Practice and Conduct it states when reporting an inconclusive outcome, there should be an explanation as to why the outcome is inconclusive. Also, where relevant, the notes should include the “mark status” for example, searched with a negative result and remains unidentified. “Should” indicates generally accepted practice, whereas “shall” is a corresponding requirement in ISO 17025 standards or FSR Codes of Practice and Conduct.

The structured approach prompted experts to record the features during the comparison stage and pair these to the features seen in the analysis stage which explains the consistency between experts. There is a free text box in the software for experts to record any additional notes, however this was rarely used apart from image 3. This impression was graded 4 but participants believed there was enough detail present to exclude so therefore continued to the comparison stage. Due to the lack of clarity of the impression and known print, the experts reported this as exclusion or inconclusive. The reasons for this were explained in the free text box. This implies that there are impressions that experts do have additional notes to make and feel the need to use the text box, taking this into consideration it would be useful to have a free text box for this reason. Most strategies observed do include a free text box as a safety net. The correct known print was provided to the experts when using the online software to reduce time in completing the study. It wasn’t important to assess if the experts were identifying the mark correctly, as this is not novel. The study was investigating the software as a means of documenting the process, therefore the correct known print was always provided to the expert. This explains the zero response for identifier and anatomical source as the experts were not asked to record this.

The results show what is most recorded during the comparison and evaluation stage and the consistency between experts. Following this, a series of chi square tests was conducted to see if there was an association between what is recorded within the comparison and evaluation stages and the grade of the mark. These are discussed below.

**Chi-square test for association weighted by after data: Comparison notes.**

The Chi square test indicated a significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting the comparison notes during the examination after accreditation, ꭓ2 (12) =39.874, p<0.001, two-tailed test. This implies there is less consistency between the experts’ response for that category, across all grades of the marks. The highest residual is in the category explanations for an unusable mark which was 4.9. The counts were higher than expected. The test calculated the expected count is 1.3 but the actual count was 7. The 7 participants that recorded explanations would have explained why they believed the mark was unidentifiable due to the mark being grade 4 / unusable. Explanations could be given for any grade mark as participants could explain how they reached their conclusion whether it’s an excellent mark or unusable mark. The effectiveness of including explanations will be assessed in the effectiveness study.

w = 1.1 is a large effect size and is overpowered at 0.97. Only 20 participants were needed to achieve a power of 0.80. It is a robust finding that the quality of the mark has an effect on the comparison notes. Features in agreement are less likely to be recorded for an unusable mark as the features are not clear enough in this type of mark. As expected, participants are less likely to state the anatomical source and identifier for the unusable marks due to the quality and subsequently not making an identification.

**Chi-square test for association weighted by online software data: Comparison notes.**

The Chi square test indicated no significant association between the documentary suggestions (ACE-V checklist) and the grade of the mark when documenting the comparison notes when using PiAnoS, ꭓ2 (9) =11.955, p=0.216, two-tailed test. This implies there is consistency between the experts’ response for that category, across all grades of the marks. This is likely due to the software prompting participants to record the categories, no matter what the grade is.

w = 1 is a large effect size but is under powered at >0.80. 20 participants were needed to achieve a power of 0.80. As these are a requirement of PiAnoS, obtaining the correct number of participants may not result in a significant association between the checklist and the grade of the mark. The participants did not complete this part during study 1 therefore I cannot cross compare the results from this study to study 1.

## 3.4 – Objective 8 to 11: The determination of effective documentation strategies for Fingerprint Examinations; study three

The term “effective” is used when successfully producing an intended or desired result. For example, when a property is broken into, the homeowner’s fingerprints are taken for elimination prints. Ink strips and ten print cards are used to recover the victim (homeowners) fingerprints for exclusion purposes. The intended result is to obtain a good set of quality prints using ink strips and ten print cards. The professional taking the elimination prints will determine the quality and decide whether the ten prints are sufficient for elimination purposes. If successful, this has demonstrated that the process to recover the victim’s elimination prints and therefore, effective. In this study, the effectiveness of documentation strategies was investigated. The intended or desired result was obtaining a correct outcome of identifying a match between a documentation strategy and impression. If an expert or novice could read, follow, and correctly identify a match between the documentation notes and an impression, this would suggest the documentation strategy was effective. If a novice can understand the notes of the original expert, then this could potentially be used towards court reports. Before accreditation, court reports were created retrospectively (Bunter, 2016).

In section 3.3, new documentation strategies post- ISO 17025 accreditation was outlined. These are only a selection of UK fingerprint bureaux strategies. There are differences in strategy on how to record notes such as on the back of a photograph or in a text box form, whether this is free text or a structured tick sheet, or the use of a software to annotate impressions. Despite the differences in strategy, the results have highlighted the most common documentary suggestions throughout them all. These are a unique reference number, the date, if the impression is a finger or palm, anatomical source, the orientation of the impression, quality of the impression, level one detail present, level two annotations during the analysis stage, the analysis outcome, features in agreement at comparison stage and the evaluation outcome. These relate to generic information and level of detail and annotations from the ACE-V checklist. However, there are strategies which include prompts to record factors affecting the mark, therefore it is important that these feature within the example documentation strategies in the effectiveness study. As the participant numbers from each participating bureau differed. To achieve a fair and representative sample of each documentation strategy for the effectiveness study, if the element from the ACE-V checklist was recorded by more than 50% of the participant group, this was included into the example documentation strategy. The justification for this is if you were to dip sample half of the participant groups, it is likely the elements of the checklist would be within the notes. Dip sampling is an adopted approach within bureaux as means of non-conforming tests to comply with quality standards such as ILAC G-19 and ISO 17025.

Within the effectiveness study, an interactive PowerPoint which was disseminated to experts and novices included an example of the nine strategies from study one and study two. This includes pre-accreditation, the online software and the seven participating groups. For each documentation strategy there is an example for each grade (1 to 4). This totals to 36 sets of notes and impressions which all should have a “match” as the correct answer.

As stated in section 3.1, the impressions used in study one and study two cover all the documentary suggestions in literature and policies. To achieve all these suggestions, there were multiples of grades within the test selection. There were three grade 1s, two grade 2s, one grades three and two grade 2s. Within the effectiveness study, it was important to have at least one of each grade for each documentation strategy to determine the similarities and differences between UK fingerprint bureaux practice and the extent of notes depending on the grade of the mark. Therefore, the images used were image 1 (grade 1), image 2 (grade 2), image 6 (grade 3) and image 5 (grade 4). To allow for “non-match” answers a further 12 questions were included, these were the same documentation examples but presented alongside another impression of the same grade therefore, deemed as close non-matches. The rationale behind this decision was whether the notes for the correct impression were detailed enough to differentiate between another impression of the same grade.

In section 3.3, the results from the post-accreditation strategies demonstrated the response rate for each documentary suggestion in the ACE-V checklist, for all the strategies collectively. This was to understand the difference in strategies post accreditation and to determine if there were any associations between the notes and the grade of the mark. Within this chapter, the results are broken down further to determine which strategies recorded what documentary suggestions from the ACE-V checklist. Followed by investigating the association between obtaining a correct outcome to group, outcome to strategy, outcome to grade of the mark. As the variables are categorical, chi-square test for association is deemed most appropriate but also maintain consistency throughout the studies. As previously mentioned in 2.2.1, it was the original intentions to predict obtaining a correct outcome dependent on the documentation used, expert group and grade of the mark. A correct outcome is successfully identifying a fingermark to its corresponding notes. The statistical test used for this prediction was binominal logistic regression as it predicts the probability that an observation falls into one of two categories of a dichotomous dependent variable based on one or more independent variables that can be either continuous or categorical. After completing the tests, it was determined using the Hosmer-Lemeshow Test (as part of the testing procedure) that the model does not fit the data (Minitab, 2013). As the variables are all categorical in this study, there appears to be no statistical test to predict a correct outcome based on documentation strategy, expert group and grade of the mark. As prediction study was unsuccessful, it was decided to determine if there was an association between outcome (correct or incorrect) and documentation strategy, outcome and expert group, outcome and grade of the mark. Therefore, the test used was chi square tests for association, as previously used in study one and study two.

To recap from the previous results on how to interpret chi-square tests for association results. The *count* value is the actual number of observations in a sample that belong to a category, and the *expected count* value is the projected frequency that would be expected in a cell if the variables are independent. It is calculated by the row and column totals divided by the sample size.

The *standardized residuals* are used to see which variables have the largest difference between the expected counts and actual counts relative to the sample size. Therefore, it is the standardized residuals which are of interest when there is a significant association to pinpoint exactly where this lies in the data (Minitab, 2019). The bar charts produced as part of the output can be used to illustrate the counts more easily.

To investigate the impact of the significant associations the effect size is calculated. Cramer’s V was used to show the strength of the relationship between the two nominal variables. To interpret, Cohen (1988) is used; this refers to a weak association as 0.1, moderate association as 0.3 and strong association as 0.5.

Retrospective power analysis was carried out to determine how much power the study had after completion. This was only required if the results were non-significant.

When there is a non-significant association, power analysis is reported to discern the optimal number of participants, based on the results achieved, to obtain a power of 0.8 and increase the likelihood of then getting a significant result in future studies. 0.8 is a common power value used to find significant results. In other terms, 80% chance of detecting a difference or association. To calculate the required number of participants for future research, effect size tables based on degree of freedom were used (Clark-Carter, 2010). However, it must be noted that the optimal sample size may be unrealistic. In relation to this research, recruiting 400 fingerprint experts to get a significant association may be impossible due to workload and other commitments. In addition, even the optimal sample size may result in a non-significant association.

The p-value is reported for the Chi-square test for association, the significant association is explored and discussed using the crosstabulation and bar charts from the output, followed by the effect size and power.

### 3.4.1 – Exemplar post-accreditation documentation strategies for the effectiveness study

During the review of the documentation produced, for each strategy and each image, the notes were cross checked against the ACE-V checklist. Section 3.3 discusses the response rate for inclusion of the documentary suggestions from the checklist, in the post-accreditation documentation strategies collectively. Whereas this chapter focuses on what each individual post-accreditation documentation strategy includes from the ACE-V checklist and selecting a documentation from the participant groups to be the exemplar for that strategy and grade.

The raw data were split into eight different spreadsheets, one for each image, and it shows the response rate for the documentary suggestions from the ACE-V checklist for each individual strategy. The green indicates when more than 50% of the participant group recorded that criterion from the ACE-V checklist. [See appendix 11 which shows the full raw data for images 1 to 8.](#_Appendix_11_–)

As the study involved four images, the following tables (16-19) show all the documentary suggestions that was recorded by 50% or more of the participant group, for each strategy and for each grade, and omits all the suggestions not included (which can be seen in the raw data). Subsequently, these were the exemplar documentation strategies included within the study. It must be noted that due to the methodology in which bureaux have adopted to annotate images (pricking out or use of a comparator/software), the raw documentation could not be disseminated to participants of the study and therefore, have not been included into the exemplar. This is one limitation of the research, but the research team acknowledge this and is addressed in the further work section.

Table 16 shows the different strategies and the most common documentary suggestions included within a grade 1 impression. As shown, the extent of the notes varies between strategies.

There is consistency regarding generic information such as unique reference number and date of the examination and the reporting outcomes at Evaluation stage. Followed by the identifier and anatomical source identified to, with a copy of the reference prints attached. However, this does not imply that the same reporting outcome has been selected. The variance of opinion by experts is investigated in section 3.5.

The least information is recorded by strategy 3, which is a free text strategy. Other than generic information and reporting outcome, the only other information recorded is the level one detail present and orientation of the impression. This is less than what was required in the pre-accreditation strategy. The pre-accreditation also prompted experts to record which finger or area of palm the impression could originate from. The Forensic Science Regulator (2020) states as part of the minimum requirements, reasons for conclusions must be within the notes. The only justification within the notes for this image is the level one detail and the orientation of the impression. These notes could be used for another grade 1 impression which is an arch pattern and no factors affecting the impression. This strategy (and strategy 1) has adopted a similar approach regarding the analysis outcome.

If identified, then it is assumed the reporting outcome at Analysis was sufficient. The only case where an analysis outcome is reported, is an insufficient impression.

*Table 16 – ACE-V checklist inclusion criteria for Image 1 / Grade 1*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Participant group | | | | | | |
| Checklist | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| URN | ü | ü | ü | ü | ü | ü | ü |
| Date | ü | ü | ü | ü | ü | ü | ü |
| Start time |  |  |  |  | ü |  |  |
| If finger/palm | ü | ü |  | ü |  | ü |  |
| Which finger/palm | ü |  |  | ü |  | ü |  |
| Single/multiple | ü | ü |  |  |  |  |  |
| Orientation | ü |  |  |  | ü | ü | ü |
| Quality | ü |  |  | ü | ü |  | ü |
| Movement |  | ü |  |  |  |  |  |
| Pressure |  | ü |  |  |  |  | ü |
| Substrate |  | ü |  |  |  |  |  |
| Overlapping |  | ü |  |  |  |  |  |
| Development |  | ü |  |  |  |  |  |
| Level 1 | ü | ü | ü |  | ü | ü | ü |
| Level 2 |  | ü |  |  |  |  |  |
| Other features |  | ü |  |  |  |  |  |
| Number of features |  | ü |  |  | ü |  |  |
| Confidence level |  |  |  |  | ü |  |  |
| Level 2 annotations | ü |  |  |  | ü |  |  |
| Analysis outcome |  | ü |  | ü | ü | ü | ü |
| Features in agreement |  | ü |  |  | ü |  | ü |
| Explanations |  |  |  |  | ü |  |  |
| Outcome | ü | ü | ü | ü | ü | ü | ü |
| Identifier | ü | ü | ü | ü | ü | ü | ü |
| Anatomical source | ü | ü | ü | ü | ü | ü | ü |
| Copy of reference prints | ü | ü | ü | ü | ü | ü | ü |

Strategy 4 is a tick sheet strategy with optional notes. This strategy contains the same information as strategy 3 but also includes if and which finger or area of palm within the notes. This information is not expected as part of the tick sheet form. The tick sheet only includes analysis outcome and evaluation outcome with the subsequent identifier and anatomical source. This strategy is the only strategy to not record the level one detail.

From the notes, the justifications for the outcome are explained by stating “plenty” of ridge characteristics and pattern type visible. However, it does not give the number of characteristics identified or which pattern type. Similarly, to strategy 3 this could relate to different impressions received. The only strategies to include a number of characteristics is strategy 2 and strategy 5. Strategy 2 is the mark analysis form created as part of study one. During reviewing the results and professional conversations with the senior fingerprint expert from the participating bureau, it was decided to include the number of characteristics within the tick sheet to give more weight towards justifying the identification. Although UK fingerprint experts do not work to a numeric standard, in previous literature, a threshold has been observed. When the number of minutiae reported versus the decision of suitability for a given fingermark was plotted, a threshold for suitability appeared to be between 7 and 8. In addition, a threshold for impressions identified was observed to be between 8 and 9 (Langenburg, 2012). Taking this research into consideration, the bureau policy states to record the number of characteristics up to 10, if the expert exceeds this then 10+ is recorded rather than the exact number.

Within the comparison notes of strategy 2, it prompts experts to record the presence of a coincident sequence. This reaffirms that the characteristics are the same and in the same position with the same ridge count between the unknown impression and known print.

It is important to note, strategy 2 has all factors affecting the mark ticked off as the form prompts the expert to record if present, therefore it can be asserted that the factors are acknowledged during every case. For all other strategies, the factors affecting the mark are only recorded occasionally. It is likely the experts’ have considered factors affecting the mark, but due to the absence within the notes, this cannot be confirmed. As previously seen in the white box study (study one) and other research such as Stevenage & Pitfield (2011) and Earwaker (2017) when experts were asked to provide their thought processes, the factors were only recorded when clearly visible. Such studies have demonstrated that experts will consider these factors but is there sufficient evidence to support expert experience is enough to suggest the absence within the notes means there are no factors impacting the mark or should there still be an acknowledgment of these such factors have been considered but not present. As mentioned previously, within healthcare, it is commonly known “if its not recorded, it didn’t happen”. This statement could transpire into the forensic field.

Strategy 5 goes a step further with the notetaking than strategy 2 in relation to the level of detail and annotations category. This strategy provides a number of characteristics which are clear, and which are indicative at analysis stage, followed by the characteristics in the comparison. The strategy also includes the confidence level of these characteristics similarly to the GYRO system (Champod & Langenburg, 2011).

Table 17 shows the different strategies and the most common documentary suggestions included within a grade 2 impression. This impression is the only sequence of marks within the study. The response rate for single or multiple marks is higher when there is a sequence as opposed to a single impression. The immediate differences between grades 1 and grades 2 notes is there are more factors recorded, as it may be expected due to the quality degrading.

*Table 17 – ACE-V checklist inclusion criteria for Image 2 / Grade 2*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Participant group | | | | | | |
| Checklist | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| URN | ü | ü | ü | ü | ü | ü | ü |
| Date | ü | ü | ü | ü | ü | ü | ü |
| Start time |  |  |  |  | ü |  |  |
| If finger/palm | ü | ü |  | ü |  | ü |  |
| Which finger/palm | ü |  |  | ü | ü | ü |  |
| Single/multiple | ü | ü | ü | ü |  |  | ü |
| Orientation | ü |  | ü |  | ü | ü | ü |
| Quality |  |  | ü | ü |  |  | ü |
| Movement |  | ü | ü |  |  |  | ü |
| Pressure |  | ü |  |  |  |  | ü |
| Substrate |  | ü |  |  |  |  |  |
| Overlapping |  | ü |  |  |  |  |  |
| Development |  | ü |  |  |  |  |  |
| Level 1 | ü | ü | ü |  | ü | ü | ü |
| Level 2 |  | ü | ü |  |  |  |  |
| Other features |  | ü |  |  |  |  |  |
| Number of features |  | ü |  |  | ü |  |  |
| Confidence level |  |  |  |  | ü |  |  |
| Level 2 annotations | ü |  |  |  | ü |  |  |
| Analysis outcome |  | ü |  | ü | ü | ü |  |
| Features in agreement |  | ü |  |  | ü | ü | ü |
| Explanations |  |  |  |  | ü |  |  |
| Additional features |  |  |  |  | ü |  |  |
| Outcome | ü | ü | ü | ü | ü | ü | ü |
| Identifier | ü | ü | ü | ü | ü | ü | ü |
| Anatomical source | ü | ü | ü | ü | ü | ü | ü |
| Copy of reference prints | ü | ü | ü | ü | ü | ü | ü |

As the quality of the impressions diminish, the immediate differences between grade 3 impressions and grade 1 / 2 are the lower likelihood of experts stating if finger or palm and which finger or palm and more factors are being recorded. This can be seen in table 18 which shows the different strategies and the most common documentary suggestions included within a grade 3 impression. It is strategy 3 and 7 that starts to record the factors affecting the mark. Interestingly, strategy 3 started with the least information recorded and now has more information as the quality of the impressions deteriorate. Both strategies are free text/back of the photograph strategies and appear to record the factors when observed. Most strategies have remained consistent between what is recorded for grades 1 to 3. These are all impressions that are sufficient for identification.

*Table 18 – ACE-V checklist inclusion criteria for Image 6 / Grade 3*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Participant group | | | | | | |
| Checklist | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| URN | ü | ü | ü | ü | ü | ü | ü |
| Date | ü | ü | ü | ü | ü | ü | ü |
| If finger/palm | ü | ü |  |  |  | ü |  |
| Which finger/palm | ü |  |  |  |  | ü |  |
| Single/multiple | ü | ü |  |  |  |  |  |
| Orientation | ü | ü | ü |  |  | ü | ü |
| Quality | ü |  |  | ü |  |  | ü |
| Movement |  | ü | ü |  |  |  | ü |
| Pressure |  | ü |  |  |  |  | ü |
| Substrate |  | ü | ü |  |  |  |  |
| Overlapping |  | ü |  |  |  |  |  |
| Development |  | ü |  |  |  |  |  |
| Level 1 | ü | ü | ü |  | ü | ü | ü |
| Level 2 |  | ü |  |  |  |  |  |
| Other features |  | ü |  |  |  |  |  |
| Number of features |  | ü |  |  | ü |  |  |
| Confidence level |  |  |  |  | ü |  |  |
| Level 2 annotations | ü |  |  |  | ü |  |  |
| Analysis outcome |  | ü |  | ü | ü | ü | ü |
| Features in agreement |  | ü |  |  |  |  |  |
| Outcome | ü | ü | ü | ü | ü | ü | ü |
| Identifier | ü | ü | ü | ü | ü | ü | ü |
| Anatomical source | ü | ü | ü | ü | ü | ü | ü |
| Copy of reference prints | ü | ü | ü | ü | ü | ü | ü |

Table 19 shows the different strategies and the most common documentary suggestions included within a grade 4 impression. Strategy 1 and strategy 6 are two different types of strategies, one being tick sheet and one free text on the back of the photograph.   
Both only record generic information such as unique reference number and the date of the examination, followed by the Analysis outcome. There are no explanations, which contradicts the guidance from the Forensic Science Regulator.

If there is a difference of opinion or an error i.e., a misidentification, this cannot be resolved as the thought process cannot be determined. Arguably, it is as important to explain an insufficient impression as it is to explain an identification. It also reverts to the Court of Appeal judgement, in the R-v-Smith case, where it was stated that, “no competent forensic scientist in other areas of forensic science these days would conduct a forensic examination without keeping detailed notes of his examination and the reasons for his conclusions”. The remaining strategies include reasons for their conclusions, stating the inability to determine if finger / palm or which finger / palm and the contributing factors affecting the mark. Strategy 2 acknowledges the absence of level one, level two and other detail and reports the 0 characteristics identified.

*Table 19 – ACE-V checklist inclusion criteria for Image 5 / Grade 4*

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Participant group | | | | | | |
| Checklist | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
| URN | ü | ü | ü | ü | ü | ü | ü |
| Date | ü | ü | ü | ü | ü | ü | ü |
| Start time |  |  |  |  | ü |  |  |
| Materials used |  |  |  |  | ü |  |  |
| Single/multiple |  | ü |  |  |  |  |  |
| Orientation |  | ü |  |  | ü |  |  |
| Movement |  | ü |  | ü |  |  | ü |
| Pressure |  | ü |  |  |  |  |  |
| Substrate |  | ü | ü |  |  |  | ü |
| Overlapping |  | ü |  | ü |  |  | ü |
| Development |  | ü |  |  |  |  |  |
| Level 1 |  | ü |  |  |  |  |  |
| Level 2 |  | ü |  |  |  |  |  |
| Other features |  | ü |  |  |  |  |  |
| Number of features |  | ü |  |  |  |  |  |
| Analysis outcome | ü | ü | ü | ü | ü | ü | ü |
| Copy of reference prints | ü | ü | ü | ü | ü | ü | ü |

The next step is to identify the differences in strategies over the different grades. Tables 20 to 26 shows individually the most common documentary suggestions for each strategy and grade of the marks from 1 to 4.

When referring to table 20, strategy 1, the notes are similar across grades 1 to 3, the most difference occurs within the grade 4. This strategy only records the analysis outcome for a grade 4 impression. include unique reference number and date of examination for all grades. Further generic information such as if finger or palm, which finger or palm, if the impressions are single or multiple are recorded for grades 1 to 3. This information would be expected as the impressions are sufficient for comparison.

The orientation of the impression is indicated using an arrow and is present for grades 1 to 3 again. Factors affecting the mark category of the ACE-V checklist is not well documented in this strategy. The quality of the impression is not consistently recorded, more than half of the participant group recorded the quality for grade 1 and 3 but not for grade 2 and 4. Grade 4 are poor quality impressions and therefore it may be expected to see quality noted. Although some participants would record the presence of factors affecting the mark, overall, these did not appear in any of the exemplars due to the low response rate.

In relation to the level of detail and annotations, this strategy recorded level one detail for grades 1 to 3 and pricked the photograph to show level two detail during the analysis stage. The reporting outcome was recorded with the subsequent identifier and anatomical source for grades 1 to 3. The grade 4 impression only had the outcome reported.

*Table 20 - ACE-V checklist inclusion criteria for strategy 1 for all grades*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Strategy 1 | | | | |
|  | **Grade** | | | |
| Checklist | 1 | 2 | 3 | 4 |
| URN | ü | ü | ü | ü |
| Date | ü | ü | ü | ü |
| Start time |  |  |  |  |
| If finger/palm | ü | ü | ü |  |
| Which finger/palm | ü | ü | ü |  |
| Single/multiple | ü | ü | ü |  |
| Orientation | ü | ü | ü |  |
| Quality | ü |  | ü |  |
| Movement |  |  |  |  |
| Pressure |  |  |  |  |
| Substrate |  |  |  |  |
| Overlapping |  |  |  |  |
| Development |  |  |  |  |
| Level 1 | ü | ü | ü |  |
| Level 2 |  |  |  |  |
| Other features |  |  |  |  |
| Number of features |  |  |  |  |
| Confidence level |  |  |  |  |
| Level 2 annotations | ü | ü | ü |  |
| Analysis outcome |  |  |  | ü |
| Features in agreement |  |  |  |  |
| Explanations |  |  |  |  |
| Outcome | ü | ü | ü |  |
| Identifier | ü | ü | ü |  |
| Anatomical source | ü | ü | ü |  |
| Copy of reference prints | ü | ü | ü | ü |

Due to the design of this strategy, there is consistency between what is recorded within the notes. Although the experts may differ on their decisions for example, all experts may not agree movement is present but from the use of a tick sheet, it can be asserted that all experts have acknowledged the presence of this. The only difference between the notes and the grades is a grade 4 impression which lacks clarity and inability to identify, the comparison sheet is not used and therefore the features in agreement, outcome, identifier and anatomical source is not completed.

*Table 21 - ACE-V checklist inclusion criteria for strategy 2 for all grades*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Strategy 2 | | | | |
|  | **Grade** | | | |
| Checklist | 1 | 2 | 3 | 4 |
| URN | ü | ü | ü | ü |
| Date | ü | ü | ü | ü |
| If finger/palm | ü | ü | ü | ü |
| Which finger/palm |  |  |  |  |
| Single/multiple | ü | ü | ü | ü |
| Orientation | ü | ü | ü | ü |
| Quality |  |  |  |  |
| Movement | ü | ü | ü | ü |
| Pressure | ü | ü | ü | ü |
| Substrate | ü | ü | ü | ü |
| Overlapping | ü | ü | ü | ü |
| Development | ü | ü | ü | ü |
| Level 1 | ü | ü | ü | ü |
| Level 2 | ü | ü | ü | ü |
| Other features | ü | ü | ü | ü |
| Number of features | ü | ü | ü | ü |
| Confidence level |  |  |  |  |
| Level 2 annotations |  |  |  |  |
| Analysis outcome | ü | ü | ü | ü |
| Features in agreement | ü | ü | ü |  |
| Explanations |  |  |  |  |
| Outcome | ü | ü | ü |  |
| Identifier | ü | ü | ü |  |
| Anatomical source | ü | ü | ü |  |
| Copy of reference prints | ü | ü | ü | ü |

Strategy 3 relates to a free text box for notes and illustrations, there are some similarities throughout the different grades. Illustrations are provided to show the orientation of the impression despite the grade. It is apparent that key information is recorded when clearly visible.

*Table 22 - ACE-V checklist inclusion criteria for strategy 3 for all grades*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Strategy 3 | | | | |
|  | **Images** | | | |
| Checklist | 1 | 2 | 3 | 4 |
| URN | ü | ü | ü | ü |
| Date | ü | ü | ü | ü |
| If finger/palm |  |  |  |  |
| Which finger/palm |  |  |  |  |
| Single/multiple |  | ü |  |  |
| Orientation | ü | ü | ü | ü |
| Quality |  | ü |  |  |
| Movement |  | ü | ü |  |
| Pressure |  |  |  |  |
| Substrate |  |  | ü | ü |
| Overlapping |  |  |  |  |
| Development |  |  |  |  |
| Level 1 | ü | ü | ü |  |
| Level 2 |  | ü |  |  |
| Other features |  |  |  |  |
| Number of features |  |  |  |  |
| Confidence level |  |  |  |  |
| Level 2 annotations |  |  |  |  |
| Analysis outcome |  |  |  | ü |
| Features in agreement |  |  |  |  |
| Explanations |  |  |  |  |
| Outcome | ü | ü | ü |  |
| Identifier | ü | ü | ü |  |
| Anatomical source | ü | ü | ü |  |
| Copy of reference prints | ü | ü | ü | ü |

For example, grade 2 impression was a sequence of marks therefore multiple marks was recorded and cores and deltas were noted to justify the pattern type. If the impression is affected by factors, these were recorded. These were more prevalent in grade 3 impressions; these are the borderline impressions which explains the more information being recorded. The only impressions to have the analysis outcome reported is a grade 4, like witnessed in strategy 1. As it was insufficient at analysis, the evaluation outcome is not required. However, this strategy does explain the conclusions. Whether the lack of notes for a grade 4 will be assessed using the chi-square tests for associations and using the counts further in this chapter.

Strategy 4 is a tick sheet with an optional notes section. The tick sheet prompts experts to record the analysis outcome and the evaluation outcome. Any reasons for conclusions are a part of the optional notes section. If finger or palm impression and which finger or area of palm is recorded for the better-quality impressions i.e., grades 1 and 2. Whereas this information is not recorded for grades 3 and 4 as it may be expected. Although the quality is discussed in grades 1 to 3, if there are any factors affecting the mark this is recorded only for grades 4.

*Table 23 - ACE-V checklist inclusion criteria for strategy 4 for all grades*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Strategy 4 | | | | |
|  | **Images** | | | |
| Checklist | 1 | 2 | 3 | 4 |
| URN | ü | ü | ü | ü |
| Date | ü | ü | ü | ü |
| If finger/palm | ü | ü |  |  |
| Which finger/palm | ü | ü |  |  |
| Single/multiple |  | ü |  |  |
| Orientation |  |  |  |  |
| Quality | ü | ü | ü |  |
| Movement |  |  |  | ü |
| Pressure |  |  |  |  |
| Substrate |  |  |  |  |
| Overlapping |  |  |  | ü |
| Development |  |  |  |  |
| Level 1 |  |  |  |  |
| Level 2 |  |  |  |  |
| Other features |  |  |  |  |
| Number of features |  |  |  |  |
| Confidence level |  |  |  |  |
| Level 2 annotations |  |  |  |  |
| Analysis outcome | ü | ü | ü | ü |
| Features in agreement |  |  |  |  |
| Explanations |  |  |  |  |
| Outcome | ü | ü | ü |  |
| Identifier | ü | ü | ü |  |
| Anatomical source | ü | ü | ü |  |
| Copy of reference prints | ü | ü | ü | ü |

For example, if it is a good quality impression this is stated but not necessarily why it is a good quality impression, no inclusion of level of detail to support this conclusion. If it is a poor-quality impression the reasons for this are recorded. Again, similarly to previous strategies, if there are more than one impression present, multiple marks is recorded, but if it is a single impression, single is not recorded. The justification behind this could be that the statement of multiple impressions and labelling the impressions aids the reader to understand which impression is being referred to.

It may be important to note that strategy 5 is the only defence expert within the participants. Arguably, defence experts do fewer analyses and are always involved in disputes and therefore, may provide more detailed explanations because of this. The strategy involves a free text and the use of a software to annotate impressions. The characteristics identified are provided a confidence level to indicate the usability. Other strategies that have included software’s to annotate the impressions have not been included in the exemplars due to the software only accessible for operational needs. Therefore, to remain a fair test, these annotations have also not been included.

*Table 24 - ACE-V checklist inclusion criteria for strategy 5 for all grades*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Strategy 5 | | | | |
|  | **Images** | | | |
| Checklist | 1 | 2 | 3 | 4 |
| URN | ü | ü | ü | ü |
| Date | ü | ü | ü | ü |
| Start time | ü | ü |  | ü |
| Materials used |  |  |  | ü |
| If finger/palm |  |  |  |  |
| Which finger/palm |  | ü |  |  |
| Single/multiple |  |  |  |  |
| Orientation | ü | ü |  | ü |
| Quality | ü |  |  |  |
| Movement |  |  |  |  |
| Pressure |  |  |  |  |
| Substrate |  |  |  |  |
| Overlapping |  |  |  |  |
| Development |  |  |  |  |
| Level 1 | ü | ü | ü |  |
| Level 2 |  |  |  |  |
| Other features |  |  |  |  |
| Number of features | ü | ü | ü |  |
| Confidence level | ü | ü | ü |  |
| Level 2 annotations | ü | ü | ü |  |
| Analysis outcome | ü | ü | ü | ü |
| Features in agreement | ü | ü |  |  |
| Explanations | ü | ü | ü |  |
| Outcome | ü | ü | ü |  |
| Identifier | ü | ü | ü |  |
| Anatomical source | ü | ü | ü |  |
| Copy of reference prints | ü | ü | ü | ü |

The research team have taken this into consideration. In conjunction with the annotations, the hand-written notes also imply the number of features which are clear and indicative and then the corresponding characteristics at comparison stage. The level of detail and annotations is present for grades 1 to 3, all of which are sufficient for comparison. Within the four sets of notes, the only occasion whereby the expert reported which finger or palm was a part of the multiple marks to indicate the one in which the notes referred to.

The materials used within the examination are a part of the Forensic Science Regulator (2020) minimum requirements. This is the only occasion whereby an expert has reported the materials used. The expert noted a magnifying glass was used on the grade 4 impression; this could suggest the expert attempted to improve their chances of identifying by using a magnifying glass to zoom in to the ridges. However, as the analysis outcome suggests the impression was insufficient for comparison. The lack of quality was recorded, but the reasons for this were not.

*Table 25 - ACE-V checklist inclusion criteria for strategy 6 for all grades*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Strategy 6 | | | | |
|  | **Images** | | | |
| Checklist | 1 | 2 | 3 | 4 |
| URN | ü | ü | ü | ü |
| Date | ü | ü | ü | ü |
| If finger/palm | ü | ü | ü |  |
| Which finger/palm | ü | ü | ü |  |
| Single/multiple |  |  |  |  |
| Orientation | ü | ü | ü |  |
| Quality |  |  |  |  |
| Movement |  |  |  |  |
| Pressure |  |  |  |  |
| Substrate |  |  |  |  |
| Overlapping |  |  |  |  |
| Development |  |  |  |  |
| Level 1 | ü | ü | ü |  |
| Level 2 |  |  |  |  |
| Other features |  |  |  |  |
| Number of features |  |  |  |  |
| Confidence level |  |  |  |  |
| Level 2 annotations |  |  |  |  |
| Analysis outcome | ü | ü | ü | ü |
| Features in agreement |  |  |  |  |
| Explanations |  |  |  |  |
| Outcome | ü | ü | ü |  |
| Identifier | ü | ü | ü |  |
| Anatomical source | ü | ü | ü |  |
| Copy of reference prints | ü | ü | ü | ü |

Strategy 6 is the final tick sheet form from the selection of documentation strategies within the study. This strategy prompts experts to record if finger or palm impression, which finger or palm impression and the level one detail present at analysis stage, followed by the analysis outcome, similarly to before accreditation. The additional documentary suggestion included is the orientation of the impression. After the comparison is complete, the evaluation outcome with the subsequent identifier and anatomical source is reported.

This explains the consistency between experts for grades 1 to 3 when the impressions are deemed as sufficient for comparison. On the other hand, a grade 4 impression will not have this information recorded due to the lack of quality. However, the analysis outcome is reported. The tick sheet has an optional notes form. However, from the participants within this group, these additional notes were rarely included and therefore have not been included in the exemplar. Therefore, the results demonstrate outcomes but not necessarily how these decisions were made despite in literature and policies the suggestions to record this and the expectation from the Forensic Science Regulator. That said, this information could be enough for a sufficient auditable trail and will be assessed further in this chapter.

*Table 26 - ACE-V checklist inclusion criteria for strategy 7 for all grades*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Strategy 7 | | | | |
|  | **Grade** | | | |
| Checklist | 1 | 2 | 3 | 4 |
| URN | ü | ü | ü | ü |
| Date | ü | ü | ü | ü |
| If finger/palm |  |  |  |  |
| Which finger/palm |  |  |  |  |
| Single/multiple |  | ü |  |  |
| Orientation | ü | ü | ü |  |
| Quality | ü | ü | ü |  |
| Movement |  | ü | ü | ü |
| Pressure | ü | ü | ü |  |
| Substrate |  |  |  | ü |
| Overlapping |  |  |  | ü |
| Development |  |  |  |  |
| Level 1 | ü | ü | ü |  |
| Level 2 |  |  |  |  |
| Other features |  |  |  |  |
| Number of features |  |  |  |  |
| Confidence level |  |  |  |  |
| Level 2 annotations |  |  |  |  |
| Analysis outcome | ü |  | ü | ü |
| Features in agreement | ü | ü | ü |  |
| Explanations |  |  |  |  |
| Outcome | ü | ü | ü |  |
| Identifier | ü | ü | ü |  |
| Anatomical source | ü | ü | ü |  |
| Copy of reference prints | ü | ü | ü | ü |

Strategy 7 is the final documentation strategy witnessed that used free text on the back of a photograph, with the additional use of a software to annotate the impressions. However, due to the software only accessible for operational needs, these annotations have not been included. The research team have taken this into consideration.

Generic information is recorded consistently throughout. For grades 1 to 3 impressions, the quality of the impression and the even pressure within the impression is acknowledged. Again, like previously seen, if there are multiple marks, this is reported. Any factors that may be affecting the mark are more likely to be reported for grades 2 to 4 as the quality diminishes. Grade 4 has the most factors affecting the mark reported as it may be expected to explain the conclusions. The presence of level one detail is seen in grades 1 to 3 and can contribute towards the reasons for the outcome but it is only assumed that the pattern matches the pattern in the known print by the identification outcome. Interestingly, there is no relation between reporting outcome and grade of the impression. It may be reported, it may not. It is the expert’s decision. When an identification is made, the notes state “characteristics in agreement”, there is no number assigned to this. However, as stated, the corresponding characteristics could be indicated using the software.

To conclude, the exemplar strategies are as outlined above. The tables indicate the similarities and dissimilarities between strategies post-accreditation in relation to the documentary suggestions. There are parts of the ACE-V checklist that have not been utilised by any strategy, for any grade in the exemplars. These include end time of the examination, relationship between marks, location of the mark, quantity of the mark, volume of matrix, contaminants, selecting a target area, level one annotations, any other annotations and additional features observed in the comparison stage. However, it must be noted that some of these were reported by some individuals just not a part of the consensus. These include relationship between marks and location of the mark. The latter tables show the strategies adopted individually and how the extent of the notes may differ over the different grades. Collectively, the strategies differ between each other however there appears to be consistency between the individual strategies demonstrating experts following bureau policy on what to record as part of the examinations. The effectiveness of each strategy is investigated in the next section of this chapter and the results will provide evidence of suggested minimum requirements for effective documentation.

### 3.4.2 – Initial observations and chi-square tests for associations between a correct or incorrect outcome and expertise group

As a reminder, the final study of this research was to determine the effectiveness of the documentation strategies witnessed. This contributes to the overall aim of providing evidence-based recommendations for the content of contemporaneous notes that are recorded by identification experts as part of fingerprint examinations in casework.

A new participant group was recruited to avoid the original experts recognising their previous notes and impressions and identifying matches, which would skew the data. The new participant group included 30 worldwide fingerprint experts and 30 non-experts.

Each participant was provided with an interactive PowerPoint and it contained 48 questions. Instructions and consent forms were provided at the start of the study. For each question, the participant was required to read the notes and look at the impression alongside and record if the two matched (were related), did not match (unrelated) and unsure. The unsure option was included as a precaution if the participant was hesitant of confirming a match or non-match. Within the data analysis, the unsure decisions were partnered with the non-matches as an ‘incorrect outcome’. To recap, the definitions of a correct and incorrect outcome are correct outcome refers to a participant identifying the documentation and impression match based on the ground truth, incorrect outcome refers to a participant not identifying the documentation and impression match based on the ground truth. Obtaining a correct outcome is the measure of effectiveness in this study. Therefore, choosing the unsure option participants did not correctly identify a match between the notes and impression. The outcome is one variable, along with three other variables which are participant group, documentation strategy and grade of the mark. As the outcome is the measure of effectiveness, the other three variables have been cross tabulated against the outcome variable and later assessed using chi-square tests to determine whether there is an association between outcome and group, outcome and strategy and outcome and grade of the mark. Therefore, the results are separated into three sections in this thesis.

Within this section, the results focus on outcome and group, firstly discussing crosstabulations of the variables, followed by chi-square tests for association. The results are used to suggest if experts and non-experts can follow the notes of another and imply this is effective documentation. However, this can be investigated further, and the effectiveness of each individual strategy can be assessed. This is covered in section 3.4.3 as it focuses on the outcomes and each documentation strategy. The results from this are used to recommend effective documentation strategies and content within the notes to contribute to the aim of the PhD research. The Fingerprint Inquiry (2011) recommends producing notes for complex marks and when a fresh comparison is being made for court purposes. However, to suffice under ISO 17025 accreditation quality standards and as stated by the Forensic Science Regulator (2020) all fingerprint examinations should be recorded contemporaneously.

Langenburg (2012) recommends documentation must be done in each case to the extent that is appropriate for the complexity of the case and to the extent that is sufficiently transparent how the expert arrived at the conclusion. The recommendation mirrors the suggestion by SWGFAST (2010). In both the authors state more detail within the notes should be provided for complex cases. Complexity is a subjective term however the Forensic Science Regulator (2020) has provided a definition of complexity, although the list is not exhaustive. Within the samples in this study, impressions graded 3 and 4 would be deemed as complex due to the lack of quality impacting level one and two detail. However, as seen in the exemplar strategies, the extent of note taking for these poor-quality impressions vary with some strategies only stating an outcome and no reasons for the conclusion. This does not correlate to the recommendations seen in previous literature. Therefore, the final part of the research was to observe the counts of correct and incorrect outcomes for each grade (1-4) and to determine any associations between the outcome and grade of the mark. This can be found in section 3.4.4.

*Table 27 – Counts for correct and incorrect outcomes from experts and novices*

|  |  |  |
| --- | --- | --- |
|  | Incorrect | Correct |
| Expert | 560 | 520 |
| Novice | 596 | 484 |
| Total | 1156 | 1004 |

Table 27 shows the number of correct and incorrect outcomes from experts and novices across all varied quality impressions and documentation strategies. As it can be seen, fingerprint experts performed better than novices, there were 520 correct outcomes out of a possible 1080 by experts and 484 correct outcomes out of 1080 by novices. It may be expected that experts perform better than novices due to the expertise and knowledge within the area, however the counts are similar. The terminology used by experts will be understood by experts more so than novices and this could explain the higher count of incorrect outcomes by novices. Despite experts performing (marginally) better than novices, the experts and novices had more incorrect outcomes than correct outcomes. In total, there were 560 incorrect outcomes from experts and 596 incorrect outcomes from novices. As previously stated, the incorrect outcomes include participant responses of non-match and unsure of a match. When the incorrect outcomes are broken down further, of the 560 incorrect outcomes by experts, 322 were confirmed non-matches and 238 were unsure, this demonstrates that experts were more confident in stating a non-match than reverting to an unsure decision. Whereas, of the 596 incorrect outcomes from novices, 190 were confirmed non-matches and 406 were unsure. This demonstrates the uncertainty of novices and more likely to say unsure than confirm a non-match between the documentation and impression.

As the measure of effectiveness in this study is correctly identifying a match between notes and impressions, the higher number of incorrect outcomes would imply ineffective documentation strategies regardless of expertise group. All participants were asked to read the notes and state whether it matched or non-matched or unsure of their decision, but there is no guarantee that own judgements influenced the decisions which may explain the incorrect outcome. The high count of incorrect outcomes by novices could imply collectively post-accreditation documentation are not sustainable for court reports as a novice would not understand these or be able to use the notes correctly. There is potential for further work in this area to determine the extent of notes required within a court report for novices i.e., jurors to understand.

A chi-square test for association was conducted between group (expert or non-expert) and obtaining a correct or incorrect outcome. Any correct outcome indicates the notes are sufficient to follow and therefore, effective documentation. Of the 60 participants, there was no significant association between group and outcome, χ2(1) = 2.412, p = .131.

These findings suggest obtaining a correct or incorrect outcome is not associated with being an expert or non-expert and correlates to the cross tabulations previously discussed. The findings of the correct or incorrect outcomes for expertise group is illustrated in figure 13.

As this is a two by two, Yates correction is used. To show the strength of the relationship between the two nominal variables, Cramer’s V calculation was used. There was a weak association between group and outcome, w = 0.033, p = .120. Using prospective power analysis, with this research design to potentially find a significant association, with a medium effect size, 90 experts and 90 novices is suggested optimum in (1) degree of freedom table (Clarke-Carter, 2010). However, power analysis was not originally conducted, due to the original data analysis involved binary logistic regression. This was unsuccessful due to the research design and variables, as a result chi-square tests for association was carried out, and therefore no power analysis was completed.

Using retrospective power analysis, the results were under powered at <0.20. To achieve a power of 0.80 and to maximise the possibility of achieving a significant association, based on the 60 participants in the current sample, the study would require 800 participants in each condition i.e., 800 experts and 800 non-experts. This sample size is unrealistic to achieve and even if the optimal sample size was obtained then a significant association still may not be found due to the method design. To understand why this non-significance is occurring, the author refers to the counts, illustrated in figure 13. There is little difference in counts by experts and novices, getting an incorrect or correct outcome, which explains why there isn’t a significant association observed. If the study was carried out in the same way again i.e., making no changes to the notes or the impressions, there is the chance of a similar result as seen here and therefore would remain non-significant. However, if the test was adapted and included more documentation samples rather than one exemplar from each strategy for each grade, there could be information within the notes that aid the participants with obtaining a correct outcome. In addition, the only exemplar within the effectiveness study which involved annotating impressions i.e., marking up characteristics to demonstrate points in agreement or disagreement was the online software. Some exemplars include statements inferring characteristics in agreement but no annotations. The absence of these notes could be a reason why experts and novices reported more incorrect outcomes. If the annotations were included, it could aid with following the notes from the original expert and match to the corresponding impression. This would provide more weight in supporting Langenburg (2012) and SWGFAST (2010) documentary suggestion to include annotations. This is an area of interest to the author to investigate further.

However, from this current sample, which does include an exemplar of each grade, from each strategy, there is no association between getting a correct or incorrect outcome dependent on expertise group.

Chart, bar chart

Description automatically generated

*Figure 13 – Bar chart showing the rate of outcomes to expertise group.*

As previously mentioned, experts were more confident than novices when stating a non-match between documentation and an impression. Therefore, chi-square test for association was carried out again to find a potential significant association between expert group and outcome. The variables within this test were group (expert, non-expert) and outcome (incorrect, correct and inconclusive). To define the terms; incorrect is a result whereby experts/novice reported a “non-match” for a mated pair, correct is a result whereby experts/novice reported a “match” for a mated pair, and inconclusive is a result whereby experts/novice reported an “unsure” decision in instances where a match or non-match could not be committed.

Of the 60 participants, there was a significant association between group and outcome, χ2(2) = 79.14, p = <0.001. Using Cramer’s V, there was a weak association between expert and outcome, w = 0.191, p <.001. To understand where the association lies, the standardised residuals were assessed. The largest differences could be seen in non-experts for inconclusive outcomes (residual, 4.7) indicating there were more inconclusive outcomes than expected. In addition, there was a large difference in experts for incorrect outcomes (residual, 4.1) indicating there were more incorrect outcomes than expected. These findings support that experts perform statistically significantly better than novices and more confident in their decision between a match or non-match. It must be noted that despite experts making a definitive decision between match or non-match, overall, there were 322 (30%) of incorrect decisions as opposed to 520 (48%) correct decisions.

The numbers are not vastly different and therefore could imply ineffective documentation. There must be missing information or misinterpretation, which led to the incorrect outcome. The level of documentation and the outcome is reported and discussed further in section 3.4.3. For a further work suggestion, it would be interesting to repeat the study and removing the “unsure” option and seeing how the results are impacted.

### 3.4.3 – Initial observations and chi-square tests for associations between a correct or incorrect outcome and documentation strategy

In the previous section, the results indicated experts performed better than novices in correctly identifying related documentation and impressions. Although, there were more incorrect outcomes from both participant groups suggesting ineffective documentation collectively. However, this section discusses the counts for incorrect and correct outcomes for individual documentation strategies to highlight which strategies are most effective. Followed by chi-square tests for association to assess relationships between obtaining a correct or incorrect outcome and strategies. In turn this provided evidence of minimum requirements for fingerprint examination documentation. Table 28 shows the counts for correct and incorrect outcomes for all documentation strategies. As a reminder, strategies 1 to 7 are the post-accreditation strategies from participating bureaux, strategy 8 is the online software and strategy 9 is the pre-accreditation documentation strategy. In total, there are more incorrect outcomes than correct outcomes. However, there are strategies which achieved more correct outcomes than incorrect outcomes individually. These are strategies 2, 3, 7 and 8. The strategy with the most correct outcomes was strategy 7 (n = 146), followed by strategy 3 (n = 137) and strategy 2 (n = 129).

*Table 28 – Counts for correct and incorrect outcomes for all the documentation strategies.*

|  |  |  |
| --- | --- | --- |
|  | Incorrect | Correct |
| Strategy 1 | 141 | 99 |
| Strategy 2 | 111 | 129 |
| Strategy 3 | 103 | 137 |
| Strategy 4 | 146 | 94 |
| Strategy 5 | 123 | 117 |
| Strategy 6 | 177 | 63 |
| Strategy 7 | 94 | 146 |
| Strategy 8 | 114 | 126 |
| Strategy 9 | 147 | 93 |
| Total | 1156 | 1004 |

A chi-square test for association was conducted between documentation strategy and obtaining a correct or incorrect outcome. Any correct outcome indicates the notes are sufficient to follow and therefore, effective documentation. There was a statistically significant association between documentation strategy and outcome, χ2(8) = 92.864, p <.001. Using Cramer’s V, there was a moderately strong association between strategy and outcome, w = 0.207, p <.001.

To identify where the significant association lies within the data. The residuals can be used to determine the largest difference between the expected counts and actual counts relative to the sample size. The largest differences are within strategy 6 with 6.7 for incorrect outcomes, indicating this strategy achieved more incorrect outcomes than expected. The second largest difference is within strategy 7 with 4.7 for correct outcomes indicating this strategy achieved more correct outcomes than expected. This is further supported when exploring the actual and expected counts of correct outcomes. Using the residual counts, this indicates the most effective and least effective strategies. However, within the discussions the top three strategies and the bottom three strategies are discussed to determine what is recorded within these strategies to deem them most and least effective.

Based on the findings from the initial counts and the chi-square tests for association, the top three strategies with the most correct outcomes and therefore effective are strategy 7, 3 and 2. All of these strategies are different types of documentation strategy. Strategy 7 is free text on the back of a photograph, strategy 3 is free text with illustrations and strategy 2 is a tick sheet form. It is important to determine what documentary suggestions were within these strategies to evidence what should be included as part of the minimum requirements. Refer to table 29 to see the documentary suggestions recorded throughout the different grades.

From the ACE-V checklist, over the four grades, strategy 7 recorded generic information such as the unique reference number and date of examination. This information would help to identify a case if required in the future i.e., for court proceedings as outlined in the streamlined forensic reporting section, not all cases go to court and court reports are made retrospectively when asked to attend court. The final criterion from the generic information category was presence of single or multiple impressions. It was only stated if there were multiple impressions present on the exhibit which can be useful to gain an understanding of which notes referred to what impression. As part of the factors affecting the mark category, the orientation of the impression and quality of the impression was recorded.

Demonstrating the orientation of an impression can be useful to aid with the ridge flow and identifying level one and two detail. This strategy stated if an impression was of good quality and with the absence of factors affecting the mark, together this implied no factors affecting the mark. Alternatively, if the mark was affected by factors such as substrate interference, overlapping, movement, pressure this was recorded opposed to stating a poor-quality mark.

Taking this into consideration and that most correct outcomes come from this strategy, it is strongly advised to record the quality and the factors affecting the mark for poor quality impressions. The analysis outcome was not routinely recorded in this strategy and with strategy 3 being a close second effective strategy (whom also did not record the analysis stage), it can be suggested that only recording an insufficient impression would be acceptable as the presence of the evaluation outcome affirms a sufficient for comparison or suspect comparison only analysis outcome. However, it could be argued that if a ‘suspect comparison only’ outcome was reported, there could be a misidentification if the impression was in fact suitable for an AFIS search. In terms of the level of detail and annotations, this strategy records the level one (pattern type) alongside recorded features in agreement. Despite the exemplars not including the matching characteristics, it can only be assumed that this would provide more weight to achieving a correct outcome. The research team sees no disadvantages to the inclusion of annotations.

Finally, as expected in any forensic examination, the outcome is reported. This includes the official outcome either identification, exclusion, inconclusive and insufficient. As stated by the Forensic Science Regulator (2020), experts must record their conclusions for an inconclusive impression. However, these findings have also supported the inclusion of explanations for all other outcomes is effective for another individual to follow the notes of an original expert.

*Table 29 – ACE-V checklist inclusion for most effective strategies*

|  |  |  |  |
| --- | --- | --- | --- |
|  | Strategies | | |
| Checklist | 7 | 3 | 2 |
| URN | ü | ü | ü |
| Date | ü | ü | ü |
| If finger/palm |  |  | ü |
| Single/multiple | ü | ü | ü |
| Orientation | ü | ü | ü |
| Quality | ü | ü |  |
| Movement | ü | ü | ü |
| Pressure | ü |  | ü |
| Substrate | ü | ü | ü |
| Overlapping | ü |  | ü |
| Development |  |  | ü |
| Level 1 | ü | ü | ü |
| Level 2 |  | ü | ü |
| Other features |  |  | ü |
| Number of features |  |  | ü |
| Analysis outcome | ü | ü | ü |
| Features in agreement | ü |  | ü |
| Outcome | ü | ü | ü |
| Identifier | ü | ü | ü |
| Anatomical source | ü | ü | ü |
| Copy of reference prints | ü | ü | ü |

When comparing the most effective strategy to the other strategies. There are similarities in what is recorded and therefore provides more support to include these within the minimum requirements for fingerprint examination documentation. The criterion which was not routinely recorded in the most effective strategy was if finger or palm impression, development technique affects, presence of level two detail and any other details such as scars and creases, and the number of features observed at analysis. Arguably, these have influenced obtaining a correct outcome and therefore these criterions would be suggestions to record, rather than a part of the minimum requirements. Inclusion of these criterions would provide more information of the decision-making process. For example, stating there are level two and other details in the analysis stage, this could help inform decisions at comparison if observed in the comparison print. Furthermore, despite UK fingerprint bureaux working to a non-numeric standard, the number of features seen in the analysis and then agreed or disagreed in the comparison stage can aid with decision making.

The research team have observed the number and placement of characteristics differ between experts and are interested in establishing whether this has an impact on the comparator list in an AFIS, this is explained further in section 4.5.

The strategy with the most incorrect outcomes was strategy 6 (n = 177), strategy 9 / pre-accreditation (n = 147) and strategy 4 (n = 146). It is interesting to see these within the three top strategies obtaining incorrect outcomes as they are all very similar and strategy 6 and 4 show little change in documentation strategies post accreditation. Strategy 6 and strategy 9 (pre-accreditation) record generic information such as date, unique reference number and the anatomical source and level one detail to aid with the comparison. Although strategy 4 is relatively like strategy 6, there are additional notes provided, when necessary. The documentation example shows the notes of an examiner acknowledging the quality of the impression and the level one detail. Using these results, it indicates that this level of notes is ineffective and less likely to be followed by another individual. It provides more support to include documentary suggestions such as the orientation of the impression, the quality and subsequent factors affecting the mark and features in agreement.

### 3.4.4 – Initial observations and chi-square tests for associations between a correct or incorrect outcome and grade of the mark

As stated previously, Langenburg (2012) and SWGFAST (2010) recommends documentation must be done in each case to the extent that is appropriate for the complexity of the case and to the extent that is sufficiently transparent how the expert arrived at the conclusion. In both the authors state more detail within the notes should be provided for complex cases. The results in this next section were used to inform the extent of notes for each grade of impression.

*Table 30 – Counts for incorrect and correct outcomes for each grade.*

|  |  |  |
| --- | --- | --- |
|  | Incorrect | Correct |
| Grade 1 impression | 248 | 292 |
| Grade 2 impression | 233 | 307 |
| Grade 3 impression | 288 | 252 |
| Grade 4 impression | 387 | 153 |
| Total | 1156 | 1004 |

In table 30, more correct outcomes occurred for grades 1 and 2 impressions. Interestingly, participants achieved more correct outcomes for grade 2 impressions than grade 1 impressions. However, as previously mentioned in study one and study two, as one of the grade 2 marks was a sequence of marks, this could be the reason for the slighter increase in counts for correct outcome.

If the notes indicated the presence of sequence of marks, this would lead to potentially getting the correct outcome. Five out of the seven strategies reported the grade 2 impression was a multiple of marks. If this outlier was removed, there could be a trend of most correct outcomes to least correct outcomes for grades 1 to 4 impressions.

More incorrect outcomes occurred for grades 3 and 4. This could be expected due to the quality of the impressions; however, it also indicates that there are little notes provided for someone to follow and understand the thought process of the original expert. Grades 3 are any marks with 1st and 2nd unclear, limited quantity, distortion present but sufficient for identification purposes. Therefore, this information could be recorded for example, the quality and the subsequent factors affecting the impression but the presence of level one allows for sufficient to identify.

Furthermore, grades 4 are any marks with poor quality 1st and 2nd level detail, quantity insufficient for identification purposes. Again, the lack of quality and affects could be noted. As previously discussed, a mark can be affected due to a number of factors such as pressure, movement, development technique. In the previous section, the effective documentation strategies include these factors for grade 4 impressions. Four of the seven (strategy 1, 3, 5 and 6) new documentation strategies provided little notes or no notes for insufficient marks which does not comply with the Forensic Science Regulator guidance.

When investigating the experts’ number of incorrect outcomes for these strategies with no notes, the range was between 14 to 30 dependents on strategy. Strategy 1, 5 and 6 only reported the analysis outcome and these all had between 21 and 30 incorrect outcomes for grade 4. Strategy 3 had almost 14 incorrect outcomes, this strategy included a factor affecting the mark which supports the suggestion of recording more information as opposed to only stating the analysis outcome “insufficient”. SWGFAST (2011) and Langenburg (2012) state the extent of documentation will be dependent on the complexity of the mark and is sufficiently transparent to how the expert reached the conclusion. Both suggesting a complex case, will have more detail within the documentation. A complex case will be a borderline impression which is very close to the threshold. As stated, the UK work towards a non-numeric standard, so this will be a personal threshold. Langenburg (2012) goes further than SWGFAST (2012) to suggest including the confidence level of observed characteristics. However, strategy 8 included marking up characteristics and confidence level but this did not aid with the interpretation of notes and resulting in a correct outcome. In further work, it would be beneficial to include the annotated images to determine the impact this has on the outcome.

The information recorded in the grades 1 and 2 and therefore contribute to obtaining a correct outcome is generic information such as unique reference number and date of examination, with the inclusion of if the impression is single or multiple.

The orientation of the impression and the quality with any obvious factors affecting the impression. The presence of level one detail and level two with the number of features in agreement. Finally, the reporting outcomes at analysis and evaluation with the subsequent identifier and anatomical source.

A chi-square test for association was conducted between grade of finger/palm marks and obtaining a correct or incorrect outcome. Any correct outcome indicates the notes are sufficient to follow and therefore, effective documentation. There was a statistically significant association between grade and outcome, χ2(3) = 107.361, p <.001.

Using Cramer’s V, there was a moderately strong association between grade of the mark and outcome, w = 0.233, p <.001. The significant association lies amongst the grade 2 and grade 4 impressions. The adjusted residual for grade 2 impressions obtaining a correct outcome was 5.6 demonstrating there were more actual counts than expected counts. As previously mentioned, this could be due to the sequence of marks included in the test samples. Within the notes, if a single or multiple of marks are noted, this allows the reader to identify a clear difference between impressions if there are more than one. The adjusted residual for grade 4 impressions obtaining a correct outcome was -9.8 demonstrating the actual counts were less than the expected count. An explainable reason for this low count could be based on some bureaux only reporting the outcome and not explaining the conclusion therefore it is difficult to distinguish between grade 4 impressions. The results support Langenburg (2012) and SWGFAST (2010) statement regarding more detail is required for complex cases. Using the strategies which did produce notes for grades 3 and 4 impressions, it is suggested to include; unique reference number, date of examination, if finger or palm, orientation of the impression, quality of the impression, presence of distortion, level one detail if present, level two detail if present, outcome at analysis and comparison (if progressed to this stage).

## 3.5 – Objective 12: The assessment of inter and intra-examiner variation of reporting outcomes at Analysis and Comparison/Evaluation stage

The aim of the PhD research is to provide evidence-based recommendations for the content of contemporaneous notes that are recorded by identification experts as part of fingerprint examinations in casework. Objectives 1 to 11 are related to this aim and have been discussed in sections 3.2, 3.3 and 3.4.

The research has facilitated data which has been used to assess the inter and intra-examiner variance of reporting outcomes. Although it does not relate to this aim, the research contributes to existing research and the novelty of this research is the use of contemporaneous notes to understand expert decision making.

For this research and to remain consistent with previous work (Needham *et al,* 2022), the same definitions of grades 1 to 4 are as follows; Grade 1 is an impression with clear 1st and 2nd level detail, sufficient to see a pattern or area of palm and clearly identifiable. Grade 2 is an impression with clear 1st and 2nd level detail, sufficient to identify, may not be able to establish a pattern or area of palm. Some areas not clear or distorted. Grade 3 is an impression with 1st and 2nd detail unclear, limited quantity available, presence of distortion but sufficient for identification purposes. Grade 4 is an impression with poor quality and quantity of 1st and 2nd level detail and is insufficient for identification purposes.

Although it is recognised that the grades 1 - 4 are the opposite to the CAST grading system, which is used frequently in research. The grading system used in this doctoral research is used by a collaborating UK fingerprint bureau. The bureau used this system to grade impressions created for ground truth data to ensure the samples varied in quality to replicate casework. To determine the sample of impressions varied in quality and therefore, replicated case work, the grading system was used to provide each impression with a grade. Eldridge *et al,* (2021) stated the impressions used in their study replicated casework by including a range of substrates, matrices, and development techniques, but no further explanation of how the quality was assessed. Whereas the method adopted in this study provided a more objective approach to determine at least one impression for each category. Out of the eight impressions, 3 of them are a grade 1, 2 of them are grade 2, 1 is a grade 3 and 2 of them are a grade 4 (see table 31).

The results will focus on each grade independently, highlighting the reliability and reproducibility of experts and identifying trends between the data. The contemporaneous notes produced by the experts will be analysed to determine if the notes can indicate the thought processes of the potential outcome changes. As previously mentioned in chapter 3.2, the pre-accreditation documentation strategy was on the case management systems on force computers and experts were expected to note anatomical source, pattern type and geographical search area. Also, the outcome at Analysis stage. However, this lacked detail on sufficiency decisions therefore the experts participating in this research were asked to produce notes which captured their thought process.

In addition, another documentation strategy used in both study one and two was an online software called PiAnoS. This is a picture annotation software, online, available for research purposes. As a reminder, it prompts users to document their examination according to suggestions from literature and policies, therefore is relevant to include as a documentation strategy in this research. The software has the capability of annotating marks and assigning confidence levels on the overall quality of the mark and the minutiae identified. The confidence level also known as the GYRO system (Champod & Langenburg, 2011) highlights the confidence in usability of an area within the impression and minutiae identified by using a colour coding scheme. Green indicating high confidence, yellow indicating medium confidence and red indicating low confidence. Orange is used for characteristics seen in the comparison stage to address bias. The impact of contextual and confirmation bias in forensic science has been researched, Dror & Charlton, (2006); Dror *et al*, (2006);Langenburg *et al*, (2009) specifically investigated bias within fingerprint evidence.

Contextual bias is used to describe the tendency for forensic analysis to be influenced by background information for example, case information. Confirmation bias is used to describe the tendency to search for or interpret information that confirms or supports beliefs for example, seeing minutiae in the known print and trying to find it in the unknown impression, sometimes forcing the fact (Langenburg *et al,* 2009). It is believed that confirmation bias and reverse reasoning was the cause of the misidentification in the Madrid bombing case (PCAST, 2016). Therefore, orange in the GYRO system acknowledges this, and uses orange to demonstrate any additional minutiae seen in the comparison.

PiAnoS also has a series of multiple-choice questions relating to anatomical source, pattern type, factors affecting the impression, quality of level 1, 2 and 3 details and suitability of the impression.The suggestion of recording the presence of distortion and the quality of the mark comes from The Fingerprint Inquiry (2011), ENFSI (2015), SWGFAST (2012) as it aids with the justification of how the outcome was made. Following the questions, further notes can be entered into the text box available, however this is not compulsory.

As previously discussed in chapter 3.3, the results of study two highlighted new documentation strategies. The different strategies adopted since accreditation included written notes on the back of a photograph, tick sheet form to prompt experts on what to record, free text form with illustrations to show orientation of the impression and use of digital software’s such as CSIpix and FSH. These are software that can document the examinations including annotating the impressions to show minutiae in the unknown impression and corresponding minutiae or new minutiae seen at comparison stage.

*Table 31 – The eight impressions used within this study*

A picture containing text

Description automatically generated

### 3.5.1 - Analysis and Evaluation outcomes for Grade 1 impressions

Grade 1 impressions in this study were single impressions (x2) and a palm impression, all with little or no interference causing distortion to the ridge detail. Experts have recorded their outcomes at Analysis for the same impressions four times, unknowingly. Once using the pre-accreditation documentation strategies, followed by using an online digital software to document fingerprint examinations known as PiAnoS.

After this, the new post-accreditation documentation strategy was used to document the fingerprint examination, followed by again the use of PiAnoS.

As mentioned previously, the Analysis outcomes are not defined by the Forensic Science Regulator. During the studies, the author has witnessed different terminology in UK fingerprint bureaux but ultimately have the same meanings. Table 32 provides each reporting outcome at Analysis and Evaluation with the numerical value which will be used in the figures to visually present the reliability and reproducibility of the experts. In some cases, experts may not record their Analysis outcome. This will be identified as 5 in the graphs to completely isolate the outcome from the others. That said, it is assumed in most cases, no analysis outcomes mean sufficient as the expert did continue to comparison and evaluation.

*Table 32 – Key to interpret the graphs*

|  |  |  |
| --- | --- | --- |
| Value | Analysis outcome | Comparison/Evaluation outcome |
| 1 | Insufficient | Identification |
| 2 | Suspect comparison only | Insufficient |
| 3 | Sufficient for comparison | Inconclusive |
| 4 |  | Exclusion |
| 5 | Not stated |  |

Referring to figure 14, this graph shows the repeatability and reproducibility of experts for a grade 1 impression. Grade 1 is an impression with clear 1st and 2nd level detail and suitable for comparison. One participant (P28) changed their decision from ‘sufficient for comparison’ to ‘suspect comparison only’ during the post-accreditation notes. The negative implications of this outcome change could potentially lead to a missed opportunity to identify As a grade 1 impression would be suitable for an AFIS search, if an expert compares the impression only to a set of suspect or elimination ten prints and does not confirm an identification. There is potential that an identification could have been made using the national database. When comparing the examination notes of P28 from post-accreditation and PiAnoS (2) there are no justifications for the Analysis outcome. From the notes, it can be determined that the expert agreed on the quality of the impression, the anatomical source, and the presence of level one detail, in both stages. There is no indication of why the impression was only suitable for ‘suspect comparison only’. As part of the methodology, in the post-accreditation study, experts were provided with ten print cards which could explain why the one participant changed their outcome to ‘suspect comparison only’ as in casework fingerprint experts can manually compare unknown impressions to ten print cards when there is a suspect in custody or elimination prints from a crime scene.

Eldridge *et al* (2021) states that value for exclusion only (VEO) which is an American term for ‘suspect comparison only’ was never used as a majority vote. From that study, experts are more confident in using the impression for a comparison or disregarding the impression, there are little uses of the middle route of using the impression for exclusion purposes only. Immediately, this is a difference between Eldridge *et al* (2021) participant group and this current UK participant group. Another change observed in the data, P1, P3, P4, P5, P6 and P13 all from one force which will be known as group one, did not record the Analysis outcome indicated by value 5 on the graph. The new documentation strategy procedure demonstrated an absence of analysis outcome reported if continuing to comparison stage. Rather than stating sufficient for comparison, if there is a conclusion outcome reported then it is assumed that the outcome at analysis was sufficient for Ident1 search or manual comparison.

The second grade 1 impression is another single mark, which has some movement to the left-hand side of the impression (figure 15). It does not affect the quality of the ridge detail within the impression. Similarly, to the previous grade 1 impression, all experts agreed on the Analysis outcomes during pre-accreditation and PiAnoS (1). All stated sufficient for comparison. A difference of opinion occurs again in post-accreditation, on this occasion one more expert changed their outcome. The same expert (P28) changed their decision from sufficient for comparison to ‘suspect comparison only’ again, which could be due to the method design as mentioned above. Group one did not report the analysis outcomes as per bureau guidelines, but also P21 did not report an Analysis outcome. This expert did continue to comparison stage therefore suggests the impression was sufficient for comparison.

The third grade 1 impression is a palm impression (figure 16). The same difference of opinion can be seen in this impression as to the previous single grade 1 impression. Group one and P21 recorded no Analysis outcome which is indicated with value 5 on the graph, but it assumed this mark is sufficient for comparison and P28 changed the outcome from ‘sufficient for comparison’ to ‘suspect comparison only’. If it is assumed that group one and P21 remained with ‘sufficient for comparison’ and P28 only reported ‘suspect comparison only’ due to the method design. It can be reported that all experts repeated and reproduced the Analysis outcomes for the grade 1 impressions. However, there are differences of opinion with the evaluation outcomes and as the ground truth is known, the accuracy of experts can also be investigated. Refer to figures 17 to 19 for the reliability and reproducibility of experts’ evaluation outcomes.

In figures 17 and 18, all experts repeated and reproduced their outcome as all experts reported identification. All experts accurately identified the unknown fingermark in image 1 to the known print (donor 1 – right forefinger). However, despite all experts stating identification for image 4 (figure 18). There are some issues with the contemporaneous notes from one participant. Within this force, the approved documentation procedures include when there are multiple marks from the same donor, in one case, then only one impression must contain notes. P5 identified the impression correctly (donor 10 – right forefinger), however the notes state “in bureau circumstances, no notes required as notes written on image 2”. However, image 2 is the same donor but not the same finger. The sufficiency of these notes could be argued as image 2 is a sequence of impressions and from the other hand. The notes of image 2 specifically discuss the anatomical source, patterns, and coincidence sequence for the sequence of impressions.

For PiAnoS (2), the software was pre-loaded with the correct print for comparison i.e., mated pairs and therefore, all outcomes from PiAnoS could have been identification, unless the marks were of poor quality. In figure 19, all experts reported an identification correctly using the post-accreditation strategy. However, there was difference of opinion when PiAnoS (2) was used. One expert reported inconclusive. When referring to the notes, there is clear justifications for this outcome. When using PiAnoS (2), the software prompts participants to record the outcome and if inconclusive, it asks to clarify if the conclusion is leaning more towards an identification or exclusion or neither. Furthermore, it asks if the conclusion would change if a better-known print was provided. There is also an option to record if the presence of level 3 detail was used to aid with the conclusion and if there are any additional notes necessary to record. The expert who changed their conclusion to inconclusive, stated the outcome is leaning more towards an identification which corresponds to the previous outcome of identification. It is clear from the additional notes section the reason for this inconclusive outcome. The expert was unable to alter the scale of the impression and both images were pixelated under magnification which made it difficult to compare and differentiate minutiae to image noise. That said, it is likely that the expert would have remained consistent in their decision. One expert reported exclusion; however, this is incorrect. Using the notes from this expert, it is clear the reasons for this outcome. The expert believes there is issues with the scaling, which has made it difficult to compare. Also, there is excessive creasing in the known print which affects the comparison. It could be argued the outcome should be inconclusive, rather than excluded, as there is still potential of an identification if the scale was appropriate and fewer creases within the known print.

*Figure 14 – Analysis outcomes for Grade 1 impression (image 1)*

Figure 15 – Analysis outcomes for Grade 1 impression (image 4)

Figure 16 – Analysis outcomes for Grade 1 impression (image 8)

Figure 17 – Evaluation outcomes for Grade 1 impression (image 1)

Figure 18 – Evaluation outcomes for Grade 1 impression (image 4)

Figure 19 – Evaluation outcomes for Grade 1 impression (image 8)

### 3.5.2 - Analysis and Evaluation outcomes for Grade 2 impressions

Grade 2 impressions in this study were a sequence of impressions and a single impression on newspaper to cause some background interference to impact the quality of ridge detail. However, there was still sufficient quality and quantity ridge detail to identify, as judged by the experts. As previously seen in grade 1 impressions, the experts agreed with the Analysis outcomes during pre-accreditation and PiAnoS (1) studies however, the difference of opinion occurred in the post-accreditation study. Refer to figures 20 and 21.

Group one did not report an Analysis outcome but did report an identification therefore suggests the outcome at analysis was sufficient for comparison. Likewise, P21 and P22 (both from the same bureau) provided no analysis outcome but continued to comparison stage.

P28 changed from ‘sufficient for comparison’ to ‘suspect comparison only’ which suggests the method design is impacting the Analysis outcome. Over the two studies whereby experts provided the Evaluation outcomes for the sequence of impressions, all reported identification as seen in figure 22. Experts correctly identified the impression to donor 10, left ring left middle and left forefinger.

However, the other grade 2 impression had a difference of opinion with one expert, see figure 23. Following on from a sufficient for comparison outcome at the Analysis stage, P3 reported the outcome as insufficient at comparison during the post-accreditation strategy which is a potential misidentification. When referring to the notes of P3, the information recorded relates to generic information such as unique reference number and date, followed by the orientation of the impression and the evaluation outcome. There is no mention of factors affecting the impression which would correlate to the sufficient for comparison decision, however there are no mention of pattern types or characteristics. It is difficult to determine the thought process to reach this conclusion with these notes. Potentially, if the impression is sufficient for comparison at analysis stage and insufficient at comparison stage, there could be an issue with the ten print cards. Although, the remaining experts within the group managed to successfully identify using the ten print cards. The same expert reported the outcome as identified to the correct impression (donor 5 – right forefinger) whilst using PiAnoS (2). The notes demonstrate the minutiae in agreement from the analysis to comparison and within the additional notes section, the quality of the mark is mentioned stating both good quality marks identified. It is clear how the decision was made based on matching characteristics between the unknown and known impression. The notes on the quality of the impression relate to the confidence levels assigned. Almost all the unknown impression and known print were assigned green, suggesting high confidence in usability. As green was used for the ten prints, it would contradict the previous conclusion regarding the ten-print quality in post-accreditation and suggest the ten print was usable therefore, the justification for the insufficient outcome in post-accreditation for P3 is still unknown. There is potential of a misidentification (a false negative or failure to identify when this was possible), and if the notes do not explicitly state the decisions, then the issue cannot be rectified. During discussions in UK bureaux, it is apparent that only identifications are verified. No expert is double checking insufficient impressions due to workload however, this study and previous studies (Ulery *et al,* 2011, Pacheco *et al,* 2014*,* Eldridge *et al*, 2021) have shown the inconsistencies between experts and therefore, the potential of misidentifications. A possible reason for this decision change could be the use of a digital software. It is likely the tools accessible on the software enhanced the visibility of the ridge detail either within the unknown impression or known print. However, there is not enough evidence from this study to recommend the use of a digital software to aid with identifications.

*Figure 20 - Analysis outcomes for Grade 2 impression (image 2)*

Figure 21 - Analysis outcomes for Grade 2 impression (image 7)

Figure 22 – Evaluation outcomes for Grade 2 impression (image 2)

Figure 23 – Evaluation outcomes for Grade 2 impression (image 7)

### 3.5.3 - Analysis and Evaluation outcomes for Grade 3 impressions

Grade 3 impression has 1st and 2nd detail which is unclear, limited quantity available, presence of distortion but sufficient for identification purposes. It would be expected to see experts reporting sufficient for comparison at Analysis stage for this type of impression. As previously seen in grade 1 and grade 2 impressions, the experts agreed with the Analysis outcomes during pre-accreditation and PiAnoS (1) studies however, the difference of opinion occurred in the post-accreditation study.

P28 changed from ‘sufficient for comparison’ to ‘suspect comparison only’ which suggests the method design is impacting the Analysis outcome. Also, group one and P21 provided no Analysis outcome which is indicated as value 5 in figure 24 but continued to comparison stage suggesting sufficiency. This has been a common occurrence through grade 1 to 3 impressions.

When looking at the Evaluation outcomes in figure 25, all experts agreed on the reporting outcome as identification during post-accreditation and PiAnoS (2). All experts correctly identified the impression to donor 5 right forefinger. Referring to the notes, the most common documentary suggestions from the ACE-V check list used in study one and study two are unique reference number, date of examination, the anatomical source, orientation of the impression and presence of level one detail.

The linearity of the graphs (figure 14 to figure 25) indicates the reproducibility and repeatability of experts. It is the grades 3 into grade 4 impressions which significantly change. The variability within each other and between each other in borderline impressions has been observed in a previous study (Ulery *et al,* 2012).

Figure 24 – Analysis outcomes for Grade 3 impression (image 6)

Figure 25 – Evaluation outcomes for Grade 3 impression (image 6)

### 3.5.4 - Analysis and Evaluation outcomes for Grade 4 impressions

Grade 4 impressions are poor quality impressions with limited quantity of 1st and 2nd level detail and is insufficient for identification purposes. However, when referring to figure 26, the results show only one expert (P20) reported insufficient within the pre-accreditation and PiAnoS (1) study. P20 did not provide any sufficient notes to outline the thought process for this outcome. The lack of notes is one reason for the change to fingerprint examinations and implementation of ISO 17025 accreditation. However, when looking at the notes on PiAnoS (1), P20 reported insufficient due to the factors affecting the impression such as low contrast and colour reversed. Also, due to the multiple-choice questions in the system, it is noted that the characteristics are indistinct which relates to the definition of insufficient and grade 4 impressions. The remaining experts reported ‘sufficient for comparison’ or ‘suspect comparison only’ which contradicts the definition of a grade 4 impressions. These ten experts all stated the presence of level one detail which is one of the discerning factors when looking at the definitions of grades. That said, this impression could be borderline grade 3 / 4. Out of the eight impressions, image 3 has the most variance in repeatability and reproducibility of experts, this can be seen in figure 26. The reproducibility of outcomes for this grade 4 impression can be seen clearly due to the inconsistencies within the plots. However, it is difficult to determine the repeatability of the 11 experts in figure 26. Therefore, the repeatability of the experts is shown in table 33. There were 17 out of 44 occasions (39%) where an expert changed their analysis decision. Only one expert out of eleven remained consistent over the four analyses. Furthermore, there were 2 out of 15 (13%) occasions where an expert changed their evaluation decision. See figure 28.

During study one pre-accreditation, the consensus for analysis outcome was ‘suspect comparison only’. 8 out of 11 experts reported this. When investigating the notes of each expert, all experts would record the generic information and presence of level one detail. As mentioned, the pattern type must be visible for a ‘suspect comparison only’ impression. This also reflects the documentation strategy before accreditation, as mentioned in chapter 3.2, the case management system would store case information, level one detail, and geographical search area. In the additional notes, which were used to capture the thought process, the amount of information varied between experts. The fact the impression was of poor quality was mentioned by 6 of the 8 experts. However, further discussions of what was affecting the impression was only mentioned by 2 of these 6 experts.

After completing PiAnoS (1), 8 out of 11 reported ‘suspect comparison only’. However, these were not the same 8 experts. 2 of the experts changed their decisions during study one pre-accreditation and PiAnoS (1); one changing from ‘suspect comparison only’ to ‘sufficient for comparison’ and the other changing from ‘sufficient for comparison’ to ‘suspect comparison only’. In both situations, there was the potential of a misidentification. If the individual was not a suspect in custody but was on the national database. Therefore, it is crucial to have notes which can explain how the decision was made and understand where the error within the analysis occurred and to prevent potential misidentifications occurring again by providing training where necessary. The notes from pre-accreditation were limited as previously witnessed, acknowledged, and criticised in court (R-v-Smith, 2011). Therefore, when the study was designed the research team included a free text box and asked the experts to record the thought process to determine how the decision was made. Referring to P1 notes from pre-accreditation, the generic information was recorded such as unique reference number and the date of the examination. Followed by, which finger the impression potentially originated from to input into an AFIS to narrow the search. Also, the quality of the impression with any obvious factors affecting the impression, in this case pressure and development technique was recorded. Finally, the presence of level one detail was recorded. As the expert can see level one detail, this may be enough information to exclude an individual which would explain the outcome ‘suspect comparison only’. This expert changed their decision when presented with the same impression on PiAnoS, unknowingly. On this occasion, the analysis outcome was reported as sufficient for comparison. As a reminder, this outcome is used when the impression is suitable for an AFIS search, not only manual comparison. When referring to the notes on PiAnoS, they are contradicting the reporting outcome. The expert recorded the mark is of poor quality and needs enhancing before it can be compared if possible. There is little contrast between the impression and the substrate, and the characteristics are indistinct. However, reported the outcome as sufficient for comparison. There are some similarities between the two documentation notes from pre-accreditation and PiAnoS. The expert stated the presence of pressure interfering with the impression and the presence of level one detail on both. Referring to P3 notes from pre-accreditation, the expert justified the sufficient outcome by noting generic information, which finger the impression potentially originated from, level one detail present and the quality of the impression. However, these notes could relate to a ‘suspect comparison only’ outcome too. This expert changed their decision to ‘suspect comparison only’ when using PiAnoS (1). The system prompts experts to record factors affecting the mark, anatomical source, level one detail, quality of detail and outcome. This is like what was seen in the pre-accreditation notes; however, the expert goes into much further detail in the notes section of the system. It states, “this is a single impression showing a whorl pattern but very little second level detail characteristics can be seen. There appears to be areas of high pressure where the ridges cannot be distinguished. I have done an approximate ridge count of 10 but this may not be accurate due to distortion on the mark. The ridges do go quite high and are potentially fault ridges flowing to the left at the top of the mark therefore my digit determination for this mark is Left Thumb. I think this mark could be used to exclude as the pattern type can be seen but it would be difficult to identify this mark due to the lack of clear second level detail.”

Interestingly, after post-accreditation the consensus of experts changed from ‘suspect comparison only’ to ‘insufficient for comparison’. Only 2 experts reported ‘suspect comparison only’, whereas 6 experts reported ‘insufficient for comparison’.

These 6 experts previously reported ‘sufficient for comparison’ or ‘suspect comparison only’ changed their decision to ‘insufficient for comparison’. As part of bureau procedures since accreditation, some bureaux only report insufficient and do not provide any notes to explain this decision. Others may provide additional notes when necessary. 5 of the 6 experts provided no notes. The expert which did provide notes stated “not enough clear characteristics to identify” this also relates to the definition of insufficient previously discussed. If there are no notes, it makes it difficult to determine the thought process. Especially for a borderline impression, it would be useful to understand this outcome when 3 other experts have stated the impression is sufficient for comparison during this study.

These 3 experts although did not explicitly report sufficient for comparison, it can be assumed that no analysis outcome reported, which can be seen as value 5 on the graph is sufficient for comparison due to continuing to the next stage. These experts along with 2 other experts who stated ‘suspect comparison only’ all reported inconclusive at Evaluation showing repeatability. When observing the notes of these experts, two did not provide any justifications. However, the others vary in explanation but can understand the inconclusive decision. The notes indicate lack of clarity in the impression and ten prints which resulted in inconclusive. P14 goes one step further and states “cannot exclude left thumb of donor 1 due to similarities in 1st level detail”.

Referring to P20 who was the only expert to previously report the impression as insufficient during pre-accreditation and PiAnoS (1), changed their decision to ‘suspect comparison only’ and remained with this outcome in PiAnoS (2). The notes during post-accreditation indicate the reason for this change. In the Analysis, the pattern was identified as a whorl due to two visible deltas. As the pattern is visible, the expert is unable to exclude therefore progresses to comparison. There is mention of heavy background interference affecting the impression. In the Comparison and Evaluation stage in post-accreditation, P20 reports the outcome as insufficient. The factors affecting the impression noted would explain this outcome. When cross comparing the analysis notes from both parts, they are very similar and justify the same decision as for a ‘suspect comparison only’ outcome the pattern should be visible. P20 stated the pattern was a whorl based on two deltas visible. Again, these were not indicated which could have been useful. No clear characteristics are annotated on the unknown impression, and the low quality of the impression is indicated with the use of red when using the quality tool. However, as a pattern type is seen, the expert continues to comparison. The outcome at comparison changed from inconclusive to exclusion during PiAnoS (2). Exclusion is incorrect. Using the comparison notes the reason for the decision change and where the error occurred can be identified. P20 states “can exclude due to scene mark being an inner whorl and the donor finger is an outer whorl”. However, this is a mated pair. In previous studies, false negatives have been seen similarly to this. Yet, this study has the potential to understand inaccuracies. The reason for this incorrect conclusion can be determined due to the incorrect pattern type observed. This demonstrates the importance of documentation as any errors can be seen and training can be provided where necessary.

In addition, another expert changed the Evaluation outcomes for this same impression, when using PiAnoS (2) to exclusion, which is incorrect. In the analysis notes, P1 reports the potential anatomical source based on ridge direction and size of the impression and the pattern type as a whorl due to core and both deltas visible. Although it would be useful to indicate where the delta and core are positioned in case there was a difference of opinion between experts. As this impression was graded 4, but according to the consensus of the group, it is borderline grade 3. It is expected to see some acknowledgement of poor quality and reasons for this. In the notes, P1 reports presence of extreme deposition pressure which has made the ridges spread and obliterated many visible characteristics. This has been supported with the use of the quality tool to indicate low confidence for the entire impression. That said, the expert stated suitable for comparison. The notes from the comparison stage indicate pattern types are similar but the ridge tracing and orientation of pattern are complete opposites. However, if the image was flipped, it may have made the comparison less difficult. Due to these issues, it was unlikely to find a coincident sequence and there was no match of ridge flow at all which resulted in a non-identification i.e., exclusion. It could be argued that the outcome should be inconclusive due to the issues arose but can understand why the exclusion outcome was given due to no coincident sequence found, which is required for an identification.

Ultimately, the results indicate the large variance of opinion between experts and within experts for this impression. This can be anticipated for a poor-quality mark. As the Analysis stage has no objective “ground truth”, there is no correct or incorrect answer. However, as mentioned it can have an impact on a case and as stated by Eldridge *et al* (2021) it’s only by chance who receives the impression on the day as to whether it is submitted for a comparison or not.

When observing the full data, experts agree with each other and between each other more for impressions graded 1-3 than grade 4. Unless the grade 4 is of extreme poor quality, such as image 5. Figure 27 demonstrates the consistency between experts with only one expert stating ‘suspect comparison only’, the other experts all reported insufficient for all four occurrences. P5 changed their decision from insufficient in study one pre-accreditation to ‘suspect comparison only’ in PiAnoS (1). In the pre-accreditation notes, P5 mentioned the impression was affected by substrate interference and overlapping and no pattern visible therefore reports the impression as insufficient. This coincides with the definitions outlined previously. However, when looking at the notes from PiAnoS (1), the same expert reports “patchy ridge flow that looks to be overlapping in areas. No clear ridge characteristics, possible core visible, but no delta area visible. The ridge flow seems to go anticlockwise and its size would suggest a left thumb first, then all other left digits and RF. If this is the edge of a piece of paper, it more strongly supports that this could be a left thumb as it would support 'gripping it'. This mark is not good enough to search, but because there is a pattern type visible, it may be of use to exclude someone.” There are similarities in both notes i.e., acknowledging the impression affected by overlapping. However, on this occasion, a pattern is seen and justifications for this have been provided. As there is a pattern, the expert states it could be used to exclude an individual, therefore suitable for ‘suspect comparison only’.

As most experts reported insufficient at analysis and the only occurrence of a ‘suspect comparison only’ outcome was during PiAnoS (1) study whereby there was no comparison and evaluation stage included, there are no reporting outcomes concluded at Evaluation stage.

Figure 26 – Analysis outcomes for Grade 4 impression (image 3)

Table 33 – Repeatability and reproducibility of experts reporting outcomes on a Grade 4 impression (image 3)

Table

Description automatically generated

Figure 27 – Analysis outcomes for Grade 4 impression (image 5)

Figure 28 – Evaluation outcomes for Grade 4 impression (image 3)

Overall, the results have shown the reliability and reproducibility of experts regarding the Analysis and Evaluation outcomes for marks graded 1 to 3. Despite fingerprint examinations being subjective and criticised for this (PCAST, 2016), the results indicating that fingerprint experts agree with each other, and with themselves on multiple occasions is promising. In most cases, experts agreed that the impressions were sufficient for comparison using an AFIS such as Ident1 or insufficient. However, these findings differ to Eldridge *et al* (2021) however, it must be noted that this paper used palm impressions which may be the reason for variances. The study discovered a high level of variability in the analysis decisions. 599 instances which individual examiners determined a palm impression to be suitable for identification i.e., sufficient for comparison, when the majority stated it be of no value i.e., insufficient. 788 instances where examiners deemed a palm impression to be of no value when most examiners stated suitable for identification.

Furthermore, Dror & Charlton (2006) studied the possible influences of contextual information on different decision types. They also examined the consistency of decisions by representing experts with the same fingerprints previously compared but without providing any contextual information, therefore examining the reliability of fingerprint experts. As a result, from Dror & Charlton’s (2006) research, it is evident that experts are inconsistent with their decisions. However, as this study and other studies have investigated the repeatability and reproducibility of experts without taking into consideration context, it is evidence to suggest experts may differ, whether they are provided with context or not. Both research areas, investigating difference of opinion based on contextual information or not, should continue to be an area of interest within research to contribute to the validity of fingerprint examinations.

In this study, there is a significant difference of opinion within the poor-quality impressions i.e., grade 4, in both Analysis and Evaluation outcomes. The borderline impression, image 3. This impression was graded 4 by the research team, and therefore should be insufficient for comparison. However, the consensus of this group on two occasions reported the impression was suitable for ‘suspect comparison only’ and therefore, the impression is borderline grade 3 / 4. These results and similarly seen within Eldridge *et al,* (2021) it is by chance who receives the impression and can complete a full comparison. Consequently, there is potential for misidentifications as an insufficient impression would not be routinely checked as part of the verification stage.

From this study, using the contemporaneous notes, the presence of level 1 detail is routinely noted when the outcome is ‘suspect comparison only’. With further research into this area, if this is continuously seen within documentation, the cut off between a ‘suspect comparison only’ impression and insufficient for comparison is the presence or lack of level 1 detail then this could be used to demonstrate when a second checker or verification procedure should take place for an insufficient impression to potentially avoid misidentifications.

Furthermore, additional further work could investigate the accuracy of fingerprint experts for borderline grade 3 / 4 impressions to determine the likelihood of correctly identifying a match on a poor-quality impression. As the verification stage of a borderline impression, could be worthless if the accuracy rate is low.

# Chapter 4: Conclusions

## 4.1 – The establishment of pre-accreditation documentation strategies; study one

The aim of this study was to investigate what is recorded within a semi-structured and structured approach for the documentation of fingerprint evidence and whether the quality of the friction ridge impression affects the level of detail of what is recorded within the documentation.

The results have indicated that experts are more consistent using the structured approach, whereby the experts are prompted on what to record. Within the documentary suggestions, almost all of these were recorded by at least one expert, within one test mark. The location of the mark, what the matrix is, and level one annotations were the documentary suggestions not recorded. However, this can be explained. The location of the mark can be provided by the crime scene investigators within their case notes. Bunter (2017) has argued the need for this information to be recorded. It may not be frequently used, but it can be an important factor. What the matrix is may be difficult to identify and experts may not commit to one specific matrix type, unless obvious such as blood. The test marks did not include a bloody mark. Finally, to annotate the level one detail, the expert could indicate where the core or deltas are within the mark. However, this was not a tool available within the digital platform but could be a suggestion to include.

The extent of documentation produced will be dependent on the complexity of the mark or case (SWGFAST, 2012). This statement has been supported by the findings from this study, more information about the mark was recorded for the grade 3 and 4 marks (more complex).

The most common documentary suggestions have been included into a “mark analysis form”. UKAS have accepted this form as a means of recording the analysis stage however the effectiveness of this form is yet to be completed. Similarly, for other documentation strategies within other accredited bureaux. It is of interest to the author to investigate the effectiveness of each documentation strategy adopted since accreditation.

## 4.2 – The establishment of post-accreditation documentation strategies; study two

The aim of this study was to investigate what is recorded in the post-accreditation documentation strategies which have been categories as semi-structured and structured approaches for the documentation of fingerprint evidence and whether the quality of the friction ridge impression affects the level of detail of what is recorded within the documentation.

The results have shown the new documentation approaches within each of the bureaux taking part in this research. There are differences in how they have been recorded i.e., on the back of a photograph or in a form, whether this is free text or a structured tick sheet. However, the results have highlighted the most common documentary suggestions throughout them all.

These are a unique reference number, the date, if the impression is a finger or palm, which finger or palm it is believed to be, the orientation of the impression, quality of the impression, level one detail present, level two annotations during the analysis stage, the analysis outcome, features in agreement at comparison stage and the evaluation outcome. In comparison to the analysis stage from pre-accreditation, the case would be provided with a URN, the date would be available in the audit trail, the expert would be required to input if the impression was a finger or palm, which finger or palm it is being searched against in the database and the level one detail which is present. Therefore, the differences appear to be rather than recording on the case management system, it is now also or instead of written documentation. It is evident that the additions post-accreditation is the orientation of the mark, the quality of the mark and any level two annotations. However, the annotations are not required by all bureaux.

Within the comparison stage, the features in agreement are not always recorded but if they are, it is either stated in text format, or shown by taking a photograph of marked up images. There appears to be little to no explanations of how the expert reaches the conclusion, despite being a minimum requirement by the FSR.

All these documentation strategies have been accepted by UKAS, to try and adopt a uniformed approach, the next study will investigate the effectiveness of each strategy.

## 4.3 – The determination of effective documentation strategies for Fingerprint Examinations

The aim of this study was to determine the effectiveness of the strategies witnessed in pre-accreditation and post-accreditation to suggest minimum requirements within fingerprint examination documentation. This aim was successfully achieved. There is no significant association between outcome and group which implies obtaining a correct or incorrect outcome is not dependent on the group.

On the other hand, there is significant association between the outcome and strategy and the outcome and grade. Therefore, achieving a correct or incorrect outcome is dependent on the strategy and the grade of the impression. The results indicate a strong association but are under powered. Therefore, the results must be taken with caution, and it is of interest to the research team to run the effectiveness study with the required sample size to determine if it has any impact on the findings.

The minimum requirements were outlined by the Forensic Science Regulator (2020) which included name of the expert, date of the examination, unique reference number, materials used in the examination, any communications, reasons for conclusions, sufficient detailed notes for an auditable trail, records of examination, sequence of recording contemporaneous notes and reporting outcomes. The literature review in section 1.3 discusses other guidance from literature and policies as to what could be included within the documentation.

The broadness of the FSR minimum requirements allows for flexibility within UK fingerprint bureaux. However, with the documentary suggestions from literature and policies, as explained in the literature review in section 1.3, and the findings from this research, what to include as part of the reasons for conclusions are.

* Presence of multiple impressions
* Orientation of the impression
* Report a good quality impression, alternatively.
* If the impression is distorted record the factors causing this i.e., movement, pressure, substrate interference. This list is not exhaustive.
* Presence of level one detail
* Report the analysis outcome, if insufficient
* State the features in agreement, with the inclusion of illustrating this if compatible with bureau equipment i.e., comparator or photograph an annotated impression. This list is not exhaustive.

The findings are in support of the previous requirements from the Forensic Science Regulator (2020) including date of examination, unique reference number and reporting outcomes. In addition, there are other suggestions based on consistency in strategies. As these did not contribute to the most effective strategy, they are only suggestions. These are; if finger or palm impression, development technique affects, presence of level two detail and any other details such as scars and creases, and the number of features observed at analysis.

In addition, the results showed a significant association between the outcome and grade of the impressions. More correct outcomes are achieved for grades 1 and 2 impressions as opposed to grades 3 and 4. Grade 1 and 2 impressions have more information recorded and therefore, the following should be considered within documentation for all grades; generic information such as unique reference number and date of examination, with the inclusion of if the impression is single or multiple. The orientation of the impression and the quality with any obvious factors affecting the impression. The presence of level one detail and level two with the number of features in agreement. Finally, the reporting outcomes at analysis and evaluation with the subsequent identifier and anatomical source. These documentary suggestions are relevant for bureaux if a grading approach similar to this study is adopted. Some of these documentary suggestions will not be suitable for grade 4 impressions i.e., the level of detail however, the quality and factors affecting the mark could be included.

## 4.4 – The assessment of inter and intra-examiner variation of reporting outcomes at Analysis and Comparison/Evaluation stage

The data from study one and study two was used successfully to assess the reliability and reproducibility of fingerprint experts reporting outcomes at Analysis and Evaluation stage and to use contemporaneous notes to understand expert decision making.

The findings contribute to existing literature surrounding the reliability and reproducibility of fingerprint experts (Ulery *et al,* 2011, Ulery *et al,* 2012, Tangen, *et al,* 2011, Neumann *et al*, 2013, Pacheco *et al,* 2014*,* Eldridge *et al,* 2021). However, there is more consistency between this current expert group in comparison to previous literature. That said, there is a large variance between and within experts for the borderline impressions (grade 3 - 4) which has also been reported by Ulery *et al* (2012).

When there has been an incorrect outcome for example, exclusion instead of identification or inconclusive, the contemporaneous notes have been used to understand the thought process. Where the error has occurred is clearly visible and training opportunities could be provided to prevent it from happening again, which shows the importance of documentation. However, there are cases such as insufficient at Analysis or inconclusive at Evaluation, whereby experts do not provide explanations and where there was a difference of opinion, the source of the dispute could not be inferred. Additional research should be carried out to understand the border line impressions which is causing this variance of opinion between and within experts to prevent misidentifications.

## 4.5 – Further work suggestions based on the research findings

Suggestions for further work include expanding on this research with a larger sample size, but also additional novel studies investigating the importance of number and placement of minutiae, the impact ISO 17025 accreditation has had on UK fingerprint bureaux and understanding juror’s perspective on fingerprint documentation.

As a result of this study, there is evidence-based research to suggest what is deemed as effective documentation for fingerprint examinations. However, the research team are aware this is a selection of documentation strategies across the UK and therefore, it would be beneficial to review other strategies. A similar approach to study three could be conducted with the inclusion of additional strategies. Also, as some results were under powered within the chi-square tests, to achieve more robust findings a larger sample within the study should be used. The inclusion of other documentation strategies does not have to solely be UK accredited procedures but also documentation procedures outside of the UK. There is potential to broaden the scope of this research and expand into other countries and other methods to document examinations, such as digital platforms. The inclusion of digital platforms in further studies can investigate the benefits of moving from paper-based notes to digital notes and investigate the inclusion of marking up minutiae as part of day-to-day case notes for a sufficient auditable trail which as defined in this research is deemed as effective documentation. Ulery *et al* (2013), Ulery *et al,* (2014) and Langenburg (2012) have provided support that minutiae count is an important factor in decision making, however this proposed continuation from the novel study would highlight the significance of including mark ups in case notes for effective documentation. It would be interesting to compare the outcomes from this study to a study including marking up as part of the documentation strategy to determine if more correct outcomes occur when mark ups are used.

The use of the online software served its purpose as a means of documenting fingerprint examinations with the inclusion of documentary suggestions from literature and policies. It contributed to the relevance of recording certain documentary suggestions such as generic information and factors affecting the mark. However, it unexpectedly indicated the variance of mark up between experts. The variance within experts’ marking up characteristics has been observed in previous literature (Ulery *et al,* 2016). However, investigating the variance in markup was not within the objectives of this research. There are differences within the number and placement of characteristics. It is of interest to the research team to see if these differences alter the comparator list and correlation scores presented within an automated fingerprint identification system. The outcome could be used to suggest whether it is important to record the mark up or not. The proposed methodology is allowing an automated fingerprint identification system to automatically mark up the impression to where it identifies minutiae using the algorithms within the database. This will be altered by using a removal tool to replicate the experts’ mark up from studies one and studies two for each of the impressions. The correlation score and the comparison mark presented at the comparison stage after searching the database for each category will be recorded to determine if the number and placement of characteristics has an impact.

In addition, it is of interest to the research team to understand the impact of ISO 17025 accreditation has had on UK fingerprint bureaux. In part this refers to the production of documentation, which is one justification for this research however, the impact this has had on bureaux is unknown. It could potentially cause a backlog of cases. Alternatively, it may have had a positive impact on the production of court reports. As pre-accreditation, the court reports were made retrospectively as outlined in section 1.3 but with the new approach, producing contemporaneous notes, this could have aided with more reliable court reports. As the deadline to achieve accreditation was 2019, it would be interesting to determine any lessons learnt over the recent years and share experiences amongst UK fingerprint bureaux. A survey could be conducted to determine the impact of ISO 17025, any issues highlighted could lead to future research ideas.

The results from the effectiveness study indicated experts outperformed novices statistically significantly when pairing documentation to the corresponding impressions. However, it must be noted that this significant association between expert group and outcome occurred when all outcomes (correct, incorrect and inconclusive) were used as variables. The significant association was with novices reporting more “unsure” decisions/inconclusive than expected. Experts would commit to a definitive conclusion (match or non-match), rather than novices. However, 30% of the overall decision from experts were incorrect which could imply ineffective documentation and there is still room for improvement on this topic. Further work in this area could be removing the “unsure” option to encourage a definitive conclusion to demonstrate how effective the documentation is. The original reason to include novices into this study was to understand whether the documentation produced during the original examination was sufficient enough for novices i.e., jurors to understand and therefore, take these notes to court as opposed to creating new court reports. Producing sufficient notes at the time of the examination would save time in the future when going to court. Further work into jurors’ perspective on notetaking would be beneficial to understand the extent of note taking required in court reports. To establish this, an extension to the effectiveness study could be conducted. For example, a survey alongside the documentation strategies to understand what the uncertainty of any information within the documentation and therefore requires explanations but also the useful information within the documentation.

# Chapter 5: Key findings from the research and Recommendations for Practitioners

This chapter summarises the key findings from study one (pre-accreditation documentation strategies), study two (post-accreditation strategies) and study three (the effectiveness of these documentation strategies) with the corresponding recommendations listed. The recommendations are in numbered points however, this is not suggesting order of significance or priority. The dissemination of key findings and recommendations has begun. A 1-hour webinar was delivered on study one and study two to the EUIAI members, who are experts inside and outside of fingerprint examination. The purpose of this was to demonstrate the work that has occurred with several UK police fingerprint bureaux and share ideas. There was high interest from the EUIAI members to learn from study three which is motivation to continue the dissemination to professionals inside and outside of fingerprint examinations. The key findings and recommendations will be shared to UK fingerprint bureaux who participated and the Forensic Science Regulators office due to their responsibility of providing best practice in Forensic Science. This could be potentially shared with other UK fingerprint bureaux and provide evidence-based research to contribute to the digital platform, Xchange. Following an international conference such as EUIAI, there is the potential to further this research on an international perspective and learn from each other, in particularly the use of digital platforms to document the examinations as this appears to be more popular in some countries than others.

## 5.1 – The establishment of pre accreditation documentation strategies; study one

Study one involved 31 UK fingerprint experts who analysed eight impressions of varied quality. In the first part, using a form replicating pre-accreditation documentation strategies and cross compared to the ACE-V checklist which comprised of documentary suggestions. [The full list can be seen in appendix 3](#Appendix_3). Two months later, in the second part of the study, the experts performed the analysis on the same eight impressions, unknowingly, with an online software which included documentary suggestions. This documentation was also cross compared to the ACE-V checklist. Chi-square tests for association was used to determine associations between documentary suggestions and the grade of the mark.

**Key findings:**

* This is novel research which established pre-accreditation documentation strategies.
* How fingerprint examinations were documented was established through a gap analysis between professional discussions regarding what was documented in casework to what was suggested to document in literature and policies.

Any notes were reported using a case management system. This included unique reference number, date of examination, which finger or area of palm the unknown mark is believed to originate from and the geographical search area in the national database. The search area was dependent on crime type.

* The reasons for conclusions were not documented as suggested despite being suggested by CPS (2010), The Fingerprint Inquiry (2011) and SWGFAST (2013) all before the accreditation announcement.
* The response rate for documentary suggestions from the ACE-V checklist recorded for the semi-structured and structured approach were investigated initially.
  + Generic information[[4]](#footnote-4) recorded more consistently in the semi-structured approach than the structured approach.
  + Digit determination (if finger or area of palm and which finger or area of palm) recorded more frequently for grades 1 to 3.
  + Relationship between marks only recorded if present.
  + Most common documentary suggestion for factors affecting the mark was quality.
  + Any obvious factors affecting the mark were recorded for example, grade 4 impression affected due to overlapping (n = 16 out of 33).
  + Factors affecting the mark[[5]](#footnote-5) were recorded more frequently with the structured approach as prompted.
  + Less likely to be documented was the level of detail and annotations[[6]](#footnote-6) in semi-structured approach.
  + Overall, more consistency between experts when prompted.
* When there is a significant association, this implies there is a lack of consistency between experts’ recording documentary suggestions and is dependent on the grade of the mark.
  + There was a significant association between generic information reported during the pre-accreditation semi structured strategy and the grade of the mark.
  + There was a significant association between more factors reported during the pre-accreditation semi structured strategy and the grade of the mark.
  + There was a significant association between more factors reporting during the use of an online software, a more structured approach, and the grade of the mark.
* When there is no significant association, this implies there is consistency between experts’ documentary suggestions despite the grade of the mark.
  + All other associations were non-significant therefore consistency between documentary suggestions (factors affecting the mark and level of detail and annotations) despite the grade of the mark during semi-structured and structured approach.

**Recommendations:**

1. Use a structured approach for documenting fingerprint examinations.
2. As a minimum, include generic information and distortion factors impacting marks as these suggest the quality of the mark.
3. Include an additional notes section for all other documentary suggestions within the ACE-V checklist.

Recommendation 1:

Within a semi-structured approach, it allows for experts’ to be open to interpretation on what should be recorded within the documentation but with this, there is the potential of information thought about but not necessarily recorded. Arguably, this information has been used to aid with the overall decision and should be included in the documentation. As stated by the Forensic Science Regulator (2020) to include all reasons for conclusions. However, a structured approach allows for more consistency between experts’ documentation. Not necessarily consistency between exact thought processes, as an opinion-based evidence the interpretations can differ within experts, but a structured documentation approach shows all experts’ have acknowledged the same criterion throughout the thought process of conducting a fingerprint examination. It refers back to the quote “if its not recorded, it didn’t happen”. As a novel study establishing a new documentation strategy, it was important to choose the latter approach to allow for as much information to be recorded as possible. The effectiveness of this approach would be investigated and if necessary, the approach could be altered to meet the requirements of the effective strategy.

Recommendation 2:

The author supports previous literature in their documentary suggestions (CPS, 2010, The Fingerprint Inquiry, 2011, SWGFAST, 2013 & ENFSI, 2015) based on the chi-square tests for association. As a minimum, documentation strategies should include generic information such as unique reference number, date of examination, if finger or area of palm, which finger or area of palm and if there is a relationship between marks. In addition, factors causing any distortion within the mark i.e., substrate interference, development technique affects, overlapping, movement or pressure. These details are recorded dependent on the grade of the mark and therefore contributing factors to conclusions. As a result, these all feature within the documentation strategy created in this research, the “mark analysis form”.

Recommendation 3:

Although, minimum documentary suggestions have been outlined in recommendation 2. Other criteria from the ACE-V checklist were recorded and therefore should be considered in the documentation. During a professional discussion with a senior fingerprint expert, it was decided to include the presence of level one detail, the number of characteristics within the analysis stage, and the reporting outcome. These were strongly advised by the expert to feature within documentation due to the importance in comparing an unknown mark and known print. To allow for the remaining criteria from the ACE-V checklist to be recorded, when necessary, an additional notes section should be included. Guidance on the checklist should be given to all experts to ensure uniformity in note taking.

## 5.2 – Post accreditation documentation strategies; study two

Study two involved 33 UK fingerprint experts who analysed and compared eight impressions of varied quality. These were the same eight impressions from study one. In the first part, the examination was documented using the new post-accreditation documentation strategies and these were cross compared to the ACE-V checklist which comprised of documentary suggestions. Two months later, in the second part of the study, the experts performed the analysis on the same eight impressions, unknowingly, with an online software which included documentary suggestions. This documentation was also cross compared to the ACE-V checklist. The software was pre-loaded with the correct comparison mark and therefore should record identification, or inconclusive/insufficient if poor quality. Chi-square tests for association was used to determine associations between documentary suggestions and the grade of the mark.

**Key findings:**

* Novel research to establish post-accreditation documentation strategies.
* UK fingerprint bureaux documentation strategies come in different approaches i.e.,
  + Free text box form
  + Free text on the back of a photograph
  + Tick box sheet
  + The use of software to annotate impressions.
* All documentary suggestions from the ACE-V checklist were recorded at least once over the eight impressions.
* The extent of the notes differs between UK fingerprint bureaux and between grade of the mark.
* When there is a significant association, this implies there is a lack of consistency between experts’ recording documentary suggestions and is dependent on the grade of the mark.
  + There is a significant association between level of detail and annotations recorded during post-accreditation strategies and grade of the mark.
  + There is a significant association between comparison and evaluation notes recorded during post-accreditation strategies and grade of the mark.
* When there is no significant association, this implies there is consistency between experts’ documentary suggestions despite the grade of the mark.
  + All other associations were non-significant therefore consistency between documentary suggestions (generic information, factors affecting the mark and more factors) despite the grade of the mark during both approaches.
  + Reporting outcomes

**Recommendations:**

1. In addition to recommendation 2, include level of detail and annotations as well as comparison and evaluation notes within the minimum requirements as these suggest the quality of the mark.

Recommendation 4:

This recommendation coincides with recommendation 2 as it provides additional support on the previous recommendation and new statistical findings to recommend more documentary suggestions to be included in the minimum requirements.

When observing the post-accreditation strategies, the most common documentary suggestions included within the notes were from all categories; generic information, factors affecting the mark, level of detail and annotations and comparison and evaluation notes. In the generic information category, the same documentary suggestions were recorded as captured in study one providing more support to include the following: unique reference number, date of the examination, if the impression is a finger or palm, the anatomical source. The relevance of recording the URN and date is to retrieve the case if required in the future. If the finger or palm and the anatomical source can be identified, this can help to narrow the search within a database or sifting through ten print cards.

From the factors affecting the mark category, the consensus showed the orientation of the impression is recorded which can aid with determining ridge flow. In addition, the quality of the impression, whether this is good or poor, but very few acknowledgements to the reasons for this assigned quality for example, clear ridge detail and pattern visible or any distortion factors. As the previous study indicated a significant association between the more factors being recorded and the quality of the mark, the recommendation remains in place to include this within the minimum requirements. Within the level of detail and annotations category, the consensus showed the following recorded; presence of level one detail i.e., pattern type, level two annotations during the analysis stage and analysis outcome. The relevance of recording the pattern type can help to narrow the search in a database or ten print cards. If a pattern is wrongly identified, this can potentially cause a misidentification as the database will not present close matches. If the pattern type is noted and later identified as an error within the verification stage or through dip sampling, training can be provided where necessary. Level two annotations are either indicated by pricking a photograph or using a comparator, however not all bureaux have adopted this approach. The importance of including the level two annotations can aid with the decision of quality assessment and contribute to the reasons for conclusions. This is supported by SWGFAST (2012) and Langenburg (2012). In conjunction with the consensus and the previous documentary suggestions. The chi-square tests for association provided a strong support to include level of detail and annotations within the minimum requirements as there was a significant association between this documentary suggestion and the grade of the mark. Therefore, implying presence of level one, level two annotations and analysis outcome is dependent on the grade of the mark and can aid with the decision making. Lastly, the comparison and evaluation notes frequently recorded were features in agreement at comparison stage and evaluation outcome. The chi-square tests for association have provided strong support to include comparison and evaluation notes within the minimum requirements. Features in agreement and evaluation outcome. The significant association implies these documentary suggestions are dependent on grade of the mark and can aid with decision making.

To conclude, the new minimum requirements based on findings from both studies.

* Unique reference number
* Date of examination
* If finger or area of palm
* Which finger or area of palm
* If there is a relationship between marks.
* Orientation of the mark
* Quality of the mark
* Factors causing any distortion within the mark i.e., substrate interference, development technique affects, overlapping, movement or pressure.
* Presence of level one detail
* Level two annotations
* Reporting outcome at Analysis
* Features in agreement during Comparison
* Evaluation outcome

Although the minimum requirements recommended here are specific to fingerprint examinations, it does align to the broad expectations of technical records outlined in section 7.5 of the ISO 17025 accreditation standard.

## 5.3 – The effectiveness of documentation strategies for Fingerprint Examinations

Study three involved 60 participants, 30 experts (who did not take part in study one or two) and 30 novices. Each participant conducted a matching exercise whereby a documentation strategy and impression were presented side by side and the participant had to decide whether the two matched, did not match or could opt for unsure. If participants correctly identified a match, this would imply effective documentation as the participant successfully followed the notes of the original expert. Chi-square tests for association was used to determine associations between outcome and group, outcome and strategy and outcome and grade of the mark. The most effective strategies have been used to suggest a new minimum requirement in conjunction with the Forensic Science Regulator minimum requirements.

**Key findings:**

* Novel research to establishing the effectiveness of fingerprint documentation strategies.
* There is no significant association between obtaining a correct or incorrect outcome and expertise group.
* There is a significant association between obtaining a correct or incorrect outcome and documentation strategy.
* There is a significant association between obtaining a correct or incorrect outcome and grade of the mark.
* The documentation strategy approach is irrelevant to the effectiveness of documentation. With the correct guidance any approach is effective i.e., free text form, free text back of a photograph or tick sheet.
* The most effective documentation strategies include all criteria from the ACE-V checklist such as.
  + Generic information
  + Factors affecting the mark.
  + Level of detail and annotations
  + Reporting outcome at Analysis
  + Comparison and evaluation notes

**Recommendations:**

1. To adopt a semi-structured or structured approach for documenting fingerprint examinations
2. Continue to produce court reports in addition to the day-to-day case notes.
3. To achieve effective documentation within fingerprint examination, in conjunction with the Forensic Science Regulator minimum requirements, include factors affecting the mark, level of detail and annotations and comparison/evaluation notes.

Recommendation 5:

When establishing the most effective documentation strategy approach, the top three strategies had a different approach. The most effective strategy was free text on the back of a photograph, the second was a free text box with an illustrations box on a word document and the third was a tick sheet. This would imply the approach on how to document whether it is semi-structured or structured is irrelevant if guidance has been given. It is what is included within the documentation that is important.  
  
Recommendation 6:

The novice group incorrectly identified matches between documentation and impressions more than correctly identifying matches between documentation and impressions. This was also supported with the chi-square tests for association as there was no significant association between obtaining a correct or incorrect outcome and expertise group. This would imply the notes are ineffective for a court report. However, having the original documentation is beneficial when producing a court report as the exact thought process at the time of the examination is known and assist with more reliable court reports as opposed to retrospective notes as previously witnessed (Bunter, 2016).

Recommendation 7:

Within published guidance, there a broad requirement to cover all forensic disciplines such as CPS (2010), ISO 17025 accreditation (2005), ILAC-G19 (2022). All of which require forensic examinations to be documented. The Forensic Science Regulator published Codes of Practice and Conduct for Fingerprint Comparisons (2020), this outlines a list of minimum requirements to include a unique reference number, records of materials used, records of examination, the sequence of recording contemporaneous notes, the reporting outcomes, and records of communication. The broadness of the FSR minimum requirements allows for flexibility within UK fingerprint bureaux. When establishing a documentation strategy to adhere to quality standards, previous literature could have supported with what to include as part of the records of examination, or the reasons for conclusions (specifically for inconclusive marks).

There are some similarities and dissimilarities between documentary suggestions by Champod & Langenburg (2011), The Fingerprint Inquiry (2011), SWGFAST (2013), ENFSI (2015) and Bunter (2016). Using these suggestions as a guide and the findings from this research, it is recommended to include.

* Presence of multiple impressions
* Orientation of the impression
* Report a good quality impression, alternatively.
* If the impression is distorted record the factors causing this i.e., movement, pressure, substrate interference. This list is not exhaustive.
* Presence of level one detail
* Report the analysis outcome, if insufficient
* State the features in agreement, with the inclusion of illustrating this if compatible with bureau equipment i.e., comparator or photograph an annotated impression. This list is not exhaustive.
* As well as the generic information expected and stated by the FSR Codes (2020).

## 5.4 – Inter and Intra-examiner variance for reporting outcomes

Over the two studies, 11 experts completed all stages which included four analysis outcomes and two evaluation outcomes. With this data, the inter- and intra-examiner variance could be investigated, with the use of contemporaneous notes to understand the thought processes and potential changes in decisions.

**Key findings:**

* The results from this study contribute to existing research in inter and intra examiner variance.
* There is reliability and reproducibility of experts regarding analysis and evaluation outcomes for grades 1 to 3 impressions.
* Inter and intra variance of reporting outcomes for grade 4 impressions. In particular impressions borderline grade 3 into grade 4
* The variability is too high to accurately predict expert performance
* In the sample population, insufficient impressions are not verified as per bureau policy.
* Subjected to misidentifications if borderline impressions are not checked.
* Through using the documentation to understand the decision making, the presence of level 1 detail is routinely recorded when the outcome is ‘suspect comparison only’.
* The use of contemporaneous notes is supported by understanding the decision making when observing the inter- and intra-examiner variance.

**Recommendations:**

1. Establish a quality scoring assessment for impressions to identify borderline impressions and introduce second checking for insufficient impressions.

Recommendation 8:

Within the UK, fingerprint experts only verify impressions which have been identified. Therefore, any other impressions including insufficient are not verified. The results have showed the inter- and intra-examiner variance of reporting outcomes, in particular grade 3 into 4 impressions. Some experts believe there is enough information within the unknown mark to compare either using an AFIS or against a suspect. On the other hand, others believe there is insufficient information to compare. There is the potential of misidentifications here. If a quality scoring assessment was implemented, and an expert had an impression which fell within the boundaries of a borderline impression, it could be passed to the next available expert to allow for second checking and avoid misidentifications.

The quality scoring assessment could be similar to the grading system used in this research. The quality scoring assessment is split into six criteria. Five of which relate to impression affected due to distorting factors and the final relating to the presence of level one. The first criterion is substrate interference, pressure distortion, deposition pressure, development technique and overall ridge detail affected. These are given a score between 1 and 3 based on how much of the impression is impacted. Score 1 is given to 0-1/3 of the impression affected, score 2 is given to 1/3-2/3 of the impression affected and score 3 is given to 2/3-3/3 of the impression affected. The presence of level one is given 1 if yes and 2 if not. Therefore, the poor-quality marks obtain a higher score, and the good quality marks obtain a lower quality score. If this similar approach was adopted, a borderline impression would fall between 12 and higher.

# Chapter 6: References

Ashbaugh, D., (1999). *Quantitative-Qualitative Friction Ridge Analysis.* CRC Press.

Bandey, H. & Gibson, A., (2006). *Home Office.* [Online]   
Available at: sfdd0919e69639b3c.jimcontent.com/.../name/FingerprintNewsFeb08  
[Accessed 15 04 2020].  
  
Bhavana, D., Ruchi, J., Prakash, T. and JL, K., (2013). Study of fingerprint patterns in relationship with blood group and gender-a statistical review. *Arches*, *1*(1), pp.15-17.

Bunter, S., (2016). ACE-V: Meaningful note-taking during its linear application. *Fingerprint Whorld,* Volume 161, pp. 10-28.

Bunter, S., (2017). Location, Location, Location: Misinterpretation of Fingerprints on a security gate - a case study. *Fingerprint Whorld,* 42(163), pp. 8-25.

Campbell, A., (2011). *The Fingerprint Inquiry.* [Online]   
Available at: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.361.380&rep=rep1&type=pdf   
[Accessed 15 04 2020].

Clark-Carter, D. (2010). *Quantitative Psychological Research: A Student's Handbook (3rd Edition)*. Hove: Psychology Press.

Champod, C. & Langenburg, G., (2011). *The GYRO System - A recommended approach to more transparent documentation..* [Online]   
Available at: https://projects.nfstc.org/ipes/presentations/Langenburg\_GYRO-System.pdf  
[Accessed 15 04 2020].

Champod, C & Lennard, C & Margot, P & Stoilovic, M. (2004) Fingerprints and other Ridge Skin Impressions, CRC Press.

Civil Aviation Authority. (2023) Aviation Safety. [Online] Available at: [Aviation safety | Civil Aviation Authority (caa.co.uk)](https://www.caa.co.uk/Consumers/Guide-to-aviation/Aviation-safety/)

Cowger, J. (1993) *Friction Ridge Skin: Comparison and Identification of Fingerprints*. 1st edition.

Crown Prosecution Service. (1996) Criminal Procedure and Investigations Act. Ministry of Justice, section 23(1)

Crown Prosecution Service. (2010) Disclosure: Experts' Evidence, Case Management and Unused Material, May 2010: Guidance Booklet for Experts. Crown Copyright. <https://www.cps.gov.uk/legal-guidance/cps-guidance-experts-disclosure-unused-material-and-case-management> Accessed: 25th November 2019

Crown Prosecution Service. (2015) Streamlined Forensic Reporting Guidance and Toolkit. Crown Copyright. <https://www.cps.gov.uk/legal-guidance/streamlined-forensic-reporting-guidance-and-toolkit> Accessed: 25th November 2019

Courts and Tribunals Judiciary (2019) Criminal Practice Directions: Amendment No. 8 effective from 1st April 2019. <https://www.judiciary.uk/publications/criminal-practice-directions-amendment-no-8-effective-from-1st-april-2019/> Accessed: 25th November 2019.

Croxton, R.S., Baron, M.G., Butler, D., Kent, T., and Sears, V.G. (2010) Variation in amino and and lipid composition of latent fingerprints. Forensic Science International. 199. (April). p. 93 -102

Department for Environment Food and Rural Affairs. (2022) Statistical Digest of Rural England: Crime. [Online] Available at: [Crime\_Jun\_22\_final.pdf (publishing.service.gov.uk)](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1087018/Crime_Jun_22_final.pdf#:~:text=Average%20crime%20rates%20%28police%20recorded%20crime%29%20are%20lower,3%2C727%20per%20100%2C000%20population%20in%20Predominantly%20Urban%20areas.) [Accessed: 24th August 2023]

Dror, I. (2012) Cognitive Neuroscience in Forensic Science: Understanding and Utilising the Human Element. *Philosophical Transitions of the Royal Society B: Biological Sciences.* 370.

Dror, I & Charlton, D. (2006) Why experts make errors. *Journal of Forensic Identification.* Vol. 56. pp.600-616

Dror, I., Charlton, D & Peron, A. (2006) Contextual information renders experts vulnerable to making erroneous identifications. *Forensic Science International.* Vol. 156. pp.74-78.

Earwaker, H. (2017) An investigation of fingermark submission making. University College London. Available at: chrome-extension://efaidnbmnnnibpcajpcglclefindmkaj/https://discovery.ucl.ac.uk/id/eprint/1542367/1/Earwaker,%20H.%202017.%20An%20Investigation%20of%20Fingermark%20Submission%20Decision%20Making.pdf

Eldridge, H & DeDonno, M & Furrer, J & Champod, C. (2020) Examining and expanding the friction ridge value decision. Forensic Science International. Vol 314. Available at: <https://doi.org/10.1016/j.forsciint.2020.110408>

Eldridge, H & De Donno, M & Champod, C. (2021) Testing the accuracy and reliability of palmar friction ridge comparisons - A black box study. Forensic Science International. Vol 318. Available at: https://doi.org/10.1016/j.forsciint.2020.110457

European Network of Forensic Science Institutes. (2015) Best Practice Manual for Fingerprint Examination. <http://enfsi.eu/wp-content/uploads/2016/09/6._fingerprint_examination_0.pdf>

Evett, I & Williams, R. (1996) A review of the sixteen points fingerprint standard in England and Wales. *Journal of Forensic Identification.* Vol 46. pp.49-73

Executive Office of the President President’s Council of Advisors on Science and Technology (PCAST). (2016) REPORT TO THE PRESIDENT: Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods. <https://www.innocenceproject.org/wp-content/uploads/2017/03/PCAST-2017-update.pdf> Accessed: 25th November 2019.

General Medical Council. (2013) Good medical practice. Available at: [Good medical practice-english (gmc-uk.org)](https://www.gmc-uk.org/-/media/documents/good-medical-practice---english-20200128_pdf-51527435.pdf)

Hockey, D & Wilkinson, D & Kavanagh, O & Milchak, M. (2020) Building a ground-truth fingerprint dataset for proficiency testing and research. Forensic Science International. 312. DOI https://doi.org/10.1016/j.forsciint.2020.110314.

ILAC G-19 (2022). Modules in a Forensic Science Process. [» ILAC Guidance Documents (G Series) International Laboratory Accreditation Cooperation](https://ilac.org/publications-and-resources/ilac-guidance-series/)

Fieldhouse, S. (2011) *Consistency and reproducibility in fingermark deposition.*  Forensic Science International. 207. 96-100.

Forensic Comparison Software. (2018) The future of fingerprints. [Online] Available at [FCS - The future of fingerprints (forensic-software.com)](http://forensic-software.com/)

The Forensic Science Regulator. (2013) Fingerprint Examination - Terminology, Definitions and Acronyms. [Online] Available from - <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/267523/FingerprintTerminology.pdf> Accessed: 24th June 2019.

The Forensic Science Regulator. (2017) Codes of Practice and Conduct: Fingerprint Comparison. Crown Copyright. <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/638254/128_FSR_fingerprint_appendix__Issue2.pdf> Accessed: 25th November 2019

The Forensic Science Regulator. (2020) Codes of Practice and Conduct: Bloodstain Pattern Analysis. Crown Copyright. [Bloodstain Pattern Analysis (publishing.service.gov.uk)](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/917724/FSR-C-102_BPA_Issue_2.pdf) Accessed: 23rd August 2023

The Forensic Science Regulator. (2020) Codes of Practice and Conduct: Fingerprint Comparison. Crown Copyright. [Friction Ridge Detail (Fingerprint) Comparison (publishing.service.gov.uk)](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/914695/FSR-C-128__Issue3.pdf) Accessed: 29th July 2022

The Forensic Science Regulator. (2018) Fingerprint Quality Standards Specialist Group meeting minutes. Crown Copyright. <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/760147/20180918_-_FQSSG_Minutes_September_18.pdf> Accessed: 25th November 2019.

The Forensic Science Regulator Act. (2021) [FORENSIC SCIENCE REGULATOR (publishing.service.gov.uk)](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1065632/Version_2.0_Code_of_Practice_Version_30.03.2022.pdf)

Klein, P (1999) Reopening inquiry into cognitive processes in writing-to-learn,” *Educ. Psychol. Rev.* vol. 11. 203–270.

Langenburg, G. (2004) Pilot study: A statistical analysis of the ACE-V methodology- analysis stage. *Journal of Forensic Identification.* 54. 64-79.

Langenburg, G. (2012) A critical review of the ACE-V process. Ecole des Sciences Criminelles (ESC)-Institut de Police Scientfique (IPS), University of Lausanne. Switzerland. PhD thesis. <http://www.unil.ch/esc/page18345.html>

Mackay, N., 2001. *Fingerprints unreliable as evidence says police chief.* [Online]   
Available at: http://www.shirleymckie.com/documents/Herald8.7.01.pdf

Macrae, C. (2016) *Errors and Near-Misses: What Health Care Could Learn from Aviation.* [Online] Available at: [Errors and Near Misses: What Health Care Could Learn From Aviation | PSNet (ahrq.gov)](https://psnet.ahrq.gov/perspective/errors-and-near-misses-what-health-care-could-learn-aviation#:~:text=The%20ultimate%20purpose%20of%20analyzing%20errors,and%20near-miss%20events%20is%20to%20improve%20safety.)

Medical Protection. (2013) *What makes good clinical records?* [Online] Available at: [What makes good clinical records (medicalprotection.org)](https://www.medicalprotection.org/ireland/booklets/medical-records-in-ireland-an-mps-guide/what-makes-good-clinical-records)

Monson, K., Roberts, M., Knorr, K., Ali, S., Meagher, S., Biggs, K., Blume, P., Brandelli, D., Albert, M., Reneau, R & Tarasi, F. (2019) The permance of friction ridge skin and persistence of friction ridge skin impressions: A comprehensive review and new results. *Forensic Science International.* Vol. 297, pp.111-131. DOI: [10.1016/j.forsciint.2019.01.046](https://doi.org/10.1016/j.forsciint.2019.01.046)

McCartney, C. (2019) Streamlined Forensic Reporting: Rhetoric and reality. *Forensic Science International: Synergy.* 1. 83-85.

McCartney, C. & Amoako, E., 2019. Accreditation of forensic science service providers. *Journal of Forensic and Legal Medicine,* Volume 65, pp. 143-145.

Ministry of Justice. (2015) The Criminal Procedure Rules. Crown Copyright.

Minitab. (2013) *Interpret the key results for Fit Binary Logistic Model.* [Online] Available at: Available at: [Interpret the key results for Fit Binary Logistic Model - Minitab](https://support.minitab.com/en-us/minitab/21/help-and-how-to/statistical-modeling/regression/how-to/fit-binary-logistic-model/interpret-the-results/key-results/#step-4-determine-whether-the-model-does-not-fit-the-data)

National Institute of Standards and Technology (NIST). (2020) *Forensic Handwriting Examination and Human Factors: Improving the Practice Through a Systems Approach.* Available at: [Forensic Handwriting Examination and Human Factors: Improving the Practice Through a Systems Approach (nist.gov)](https://nvlpubs.nist.gov/nistpubs/ir/2020/NIST.IR.8282.pdf)

National Research Council. (2009) *Strengthing Forensic Science in the United States. A Path Forward*. [Online] Available at: [Strengthening Forensic Science in the United States: A Path Forward (ojp.gov)](https://www.ojp.gov/pdffiles1/nij/grants/228091.pdf)

Nursing Times. (2013) *Would your records stand up to scrutiny?* [Online] Available at: [Would your records stand up to scrutiny? | Nursing Times](https://www.nursingtimes.net/opinion/would-your-records-stand-up-to-scrutiny-13-09-2013/)

Pacheco, I., Cerchiai, B & Stoiloff, S. (2014) Miami-Dade Research Study for the Reliability of the ACE-V process: Accuracy and precision in latent fingerprint examinations. [Online] Available at: [Miami-Dade Research Study for the Reliability of the ACE-V Process: Accuracy & Precision in Latent Fingerprint Examinations | Office of Justice Programs (ojp.gov)](https://www.ojp.gov/ncjrs/virtual-library/abstracts/miami-dade-research-study-reliability-ace-v-process-accuracy)

Police Digital Service. (2022) A new digital fingerprint capability for policing. [Online] Available at: [A new digital fingerprint capability for policing – Police Digital Service (pds.police.uk)](https://pds.police.uk/a-new-digital-fingerprint-capability-for-policing/)

Sears, V., Bleay, S., Bandey, H. & Bowman, V., (2012). A methodology for fingermark research. *Science and Justice,* 52(3), pp. 145-160.

Regina v Smith. (2011) EWCA Crim 1296.

Richmond, K. (2018) Streamlined Forensic Reporting “Swift and sure justice?” *The Journal of Criminal Law.* Doi: 10.1177/0022018318772701

Scientific Working Group on Friction Ridge Analysis, Study and Technology (SWGFAST). (2013) Standard Terminology for Friction Ridge Examinations (Latent/Tenprint). [Microsoft Word - STANDARD Terminology-1-2.docx (clpex.com)](http://www.clpex.com/swgfast/documents/terminology/121124_Standard-Terminology_4.0.pdf)

Scientific Working Group on Friction Ridge Analysis, Study and Technology (SWGFAST). (2013) Standard for Reporting Friction Ridge Examinations (Latent/Tenprint). https://www.nist.gov/system/files/documents/2016/10/26/swgfast\_examinations-conclusions\_2.0\_130427.pdf

Stevenage, S. (2016) Fact or friction: Examination of the transparency, reliability and sufficiency of the ACE-V method of fingerprint analysis. *Science and Justice.* 267. 145-156.

Stubbs-Hayes, F. & Fieldhouse, S. & Platt, A. & Bleay, S. (2015) *Latent fingermark interaction with Iron Oxide Wet Powder Suspension; an investigation into inters and intra participant variation*. Available from:<https://www.researchgate.net/publication/314082358_Latent_fingermark_interaction_with_Iron_Oxide_Wet_Powder_Suspension_an_investigation_into_inter_and_intra_participant_variation>

Tangen, J., Thompson, M & McCarthy, D. (2011) Identifying Fingerprint Expertise. *Association of Psychological Science.* Vol. 22. DOI: <https://doi.org/10.1177/0956797611414729>   
Tully, G., 2018. *Forensic Science Regulator: Overseeing Quality (FQSSG).* [Online]   
Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\_data/file/727479/20180510\_-\_\_FQSSG\_Minutes\_May\_18.pdf

Tully, G., (2018) *Forensic Science Regulator: Overseeing Quality (FQSSG).* [Online]   
Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\_data/file/706701/20180208\_-\_FQSSG\_Minutes\_Feb\_18.pdf  
[Accessed 17 04 2020].

Tully, G., (2019). *Forensic Science Regulator: Overseeing Quality (FQSSG).* [Online]   
Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\_data/file/864261/FQSSG-20191017-01\_June\_minutes\_to\_publish.pdf  
[Accessed 17 04 2020].

Tully, G., (2019). *Forensic Science Regulator: Overseeing Quality. Newsletter.* [Online]   
Available at: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\_data/file/782189/FSRNewsletter\_No31\_feb2019\_final.pdf  
[Accessed 17th April 2020].

Ulery, B & Hicklin, A & Buscaglia, J & Roberts, M. (2011) Accuracy and reliability of forensic latent fingerprint decisions. Proceedings of National Academy of Sciences. 108 (19) 7733-7738.

Ulery, B & Hicklin, A & Buscaglia, J & Roberts, M. (2012) Repeatability and Reproducibility of Decisions by Latent Fingerprint Examiners. *PLOS one.*

Ulery, B & Hicklin, A & Buscaglia, J & Roberts, M. (2014) Measuring what latent fingerprint examiners consider sufficient information for individualisation determinations. <https://doi.org/10.1371/journal.pone.0110179>

United Kingdom Accreditation Service. (2005) ISO 17025 accreditation: testing and calibration labortories.

UKAS (2019) Guidance on the Application of ISO/IEC 17025:2017 Dealing with Expressions of Opinions and Interpretations. Crown copyright.

Yager, N., Amin, A. (2004) Fingerprint classification: a review. *Pattern Analysis and Applications.* Vol 7, pp.77–93 DOI: <https://doi.org/10.1007/s10044-004-0204-7>

# Appendices

## Appendix 1 – Scores 6 to 15 fingermarks from Ground Truth Database



Score 6 score 6.



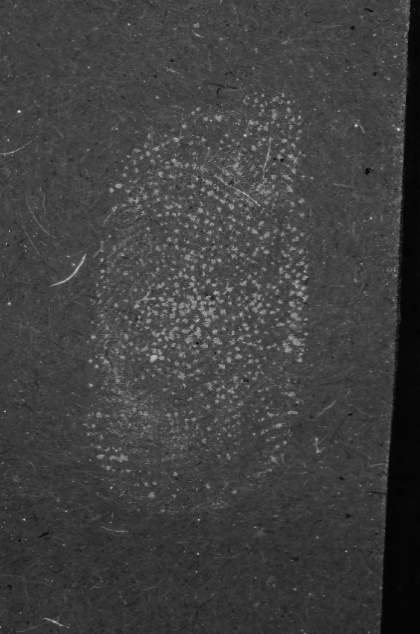


Score 8 score 8.





Score 10 score 10.



Score 12 score 12.





Score 15 score 15.

## Appendix 2 – Example of the pre-accreditation documentation strategy (analysis form)

**Image 1**

**Start date: End date:**

**Mark information:** Burglary

Taken from kitchen windowpane, possible POE   
Mark developed using Aluminium powder - Photographed in situ.

**Case Management System -**

For this part, you should state your outcome of the impression/s for example, insufficient.

**Ident1 Search –**For this part, you should record which finger or area of palm, pattern and geographical search area – local, regional and national.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Item** | **Finger/palm** | **Pattern** | **PE** | **L** | **R** | **N** |
|  |  |  |  |  |  |  |

**Analysis –** Please explain what it is about the impression/s which helped you to reach the outcome you did.

|  |
| --- |
| **Analysis** |
|  |

## Appendix 3 – ACE-V checklist

Unique reference number

Date of examination

Start time of examination

End time of examination

Materials used in examination.

If the friction ridge skin is consistent with finger or palm mark  
Which finger or area of the palm likely produced the mark?

If the friction ridge skin is part of a sequence, individual or multiple marks.

If there is a relationship to other marks which can aid with their decision

Location of the mark on the surface

Orientation of the mark

Quality of the mark

Quantity of the mark

What the matrix is

Volume of matrix

Presence of contaminants

Signs of movement

Amount of pressure

Substrate interference

Any overlapping

Any development technique effects

What level 1 features are present?

What level 2 features are present?

What other features are present?

Number of the level 2 features present

Confidence level assigned to features observed.

What feature rich area is selected for comparison.

Level 1 features annotated.

Level 2 features annotated.

Any other features annotated.

Result of examination at Analysis stage

Features in agreement

Explanations

Additional features recorded.

Reporting outcome at Evaluation stage

Identifier

Anatomical source

Copy of reference prints

## Appendix 4 – An example of a post-accreditation documentation – strategy 2 (tick sheet form)

Table

Description automatically generated

Diagram, schematic

Description automatically generated

Diagram

Description automatically generated with low confidence

## Appendix 5 – An example of a post-accreditation documentation – strategy 6 (tick sheet form)

Table, calendar

Description automatically generated

## Appendix 6 – An example of a post-accreditation documentation – strategy 4 (tick sheet form)

Diagram

Description automatically generated

## Appendix 7 – An example of a post-accreditation documentation – strategy 7 (free text / back of a photograph)

RF – Form 1

Arch, even pressure, good ridge detail and characteristics

Characteristics in agreement shown.

Date and signature

## Appendix 8 – An example of a post-accreditation documentation – strategy 1 (free text / back of a photograph)

Finger

Ident RF (suspect 1) – form used 10.

Date and signed.

Small standalone impression appears to be an arch impression with some approximation \ right slope. Ridge detail clear but appears wrong for colour. Characteristics visible – some pore structure also visible. Consider R hand and LF.

Compared to form 10 checked RF of plain impressions pattern visible, sufficient characteristics pricked to establish identification. Impression pricked wrong for colour.

## Appendix 9 – An example of a post-accreditation documentation – strategy 3 (free text / word document)

Text, letter

Description automatically generated

## Appendix 10 – An example of a post-accreditation – strategy 5 (free text / word document)

Text, letter

Description automatically generated

## Appendix 11 – Raw data of response rate for documentary suggestions and individual strategies

Image 1 / Grade 1

Table, calendar

Description automatically generated

Image 2 / Grade 2

Table, calendar

Description automatically generated

Image 3 / Grade 4

Table

Description automatically generated

Image 4 / Grade 1

Table

Description automatically generated

Image 5 / Grade 4

Graphical user interface, table

Description automatically generated

Image 6 / Grade 3

Table, calendar

Description automatically generated

Image 7 / Grade 2

Table, calendar

Description automatically generated

Image 8 / Grade 1

Graphical user interface, table, Excel, calendar

Description automatically generated

Letter of appreciation from Staffordshire Police for the work involving creation of a documentation strategy

Text, letter

Description automatically generated

This copy has been supplied on the understanding that it is copyright material and that no quotation from the thesis may be published without proper acknowledgement.

1. Best practice guidance refers to the “ACE-V checklist” created by the research team, after a literature search focusing on documentary suggestions for fingerprint examinations. [↑](#footnote-ref-1)
2. White-box study is an investigation of the internal logic and structure for example, understanding the decision making of an outcome by a fingerprint expert whereas, a black box study only tests the fundamental aspects and has no or little relevance with the internal logic and structure for example, the variance within experts’ reporting outcomes. [↑](#footnote-ref-2)
3. For the purpose of this research, a ‘correct outcome’ refers to successfully following the notes of an original expert to match the related mark and documentation. This is the measurement of effectiveness in this study. [↑](#footnote-ref-3)
4. Generic information is a subcategory of the ACE-V checklist. It refers to documentary suggestions standard information such as, unique reference number, date of examination, materials used in the examination. In addition, impression information such as anatomical source, single or multiple marks and relationship between marks. [↑](#footnote-ref-4)
5. Factors affecting the mark is a subcategory of the ACE-V checklist. This is a subcategory of the ACE-V checklist. It refers to the documentary suggestions regarding quality of the impression and any potential distortion factors including location of the impression, orientation of the impression, the quality and quantity of impression, the matrix the impression is left in and a list of distortion factors (movement, substrate, overlapping). [↑](#footnote-ref-5)
6. Level of detail and annotations is a subcategory of the ACE-V checklist. It refers to documentary suggestions regarding level one, two and three detail and the presence of annotating impressions. [↑](#footnote-ref-6)