



Discrepancies in expert decision-making in forensic fingerprint examination



Virpi Mustonen^{a,*}, Kai Hakkarainen^{b,1}, Juha Tuunainen^{c,2}, Pasi Pohjola^{d,3}

^a National Bureau of Investigation, PO Box 285, Vantaa FI-01301, Finland

^b University of Helsinki, Faculty of Behavioural Sciences, Siltavuorenpenger 5 A, PO Box 9, Helsinki, FI-00014, Finland

^c University of Oulu, Oulu Business School, PO Box 4600, Oulu, FI-90014, Finland

^d National Institute of Health and Welfare, PO Box 30, Helsinki, FI-00271, Finland

ARTICLE INFO

Article history:

Received 24 January 2015

Received in revised form 10 July 2015

Accepted 16 July 2015

Available online 29 July 2015

Keywords:

Forensic

Fingerprint

Discrepancy

Judgment

Decision making

Professional vision

ABSTRACT

The purpose of the present study was to analyse professional fingerprint examiners' investigative practices in the context of discrepancy decisions concerning challenging latents during fingerprint analysis and identification. The participants were fingerprint experts from the Forensic Laboratory of the Finnish National Bureau of Investigation. The data were from five audio-recorded "discrepancy meetings" where two examiners were discussing the rationale and justification for their differing interpretations of challenging and distorted fingerprint evidence. The meetings were chaired by the quality manager of the fingerprint group, who also in the first author of this article. The research questions addressed were as follows: What does the examiner see in the latent fingerprints? What does the examiner actively do with the latents? How were decisions made during the investigative process? In accordance with Goodwin's professional vision framework, the results revealed how the participants used partial and limited information in making judgments about the difficult and distorted latents. The examiners appeared to be involved in active, constructive efforts, mentally, to repair poor latents by supplementing with missing information. They also highlighted various aspects of latents by colour coding as well as manipulated fingerprint images in several ways so as to make the significant patterns easier to recognize. Because the methods and practices of characterizing latents were only vaguely specified, the examiners used locally developed ad hoc practices to facilitate their investigations, ending up with different interpretations. It is concluded in the article that the fingerprint community in Finland should make strong efforts to develop the methods of fingerprint investigation and determine clearer criteria for decision making and documentation practices. Furthermore, the interpretations made by fingerprint experts should be made more transparent to the customers.

© 2015 Elsevier Ireland Ltd. All rights reserved.

1. Introduction

The examination of fingerprint evidence consists of a complex stepwise process in matching two fingerprints, one found in the crime scene and the other from a fingerprint register. There is a high expectation set for the fingerprint examiners to produce accurate results during the examination process. The examiners are also expected to document this process and present their reports in court, if asked. In public, the forensic service provider's

results are often taken for granted as correct and considered to represent "the truth". Stakes are thus very high because even a single fingerprint may in fact be crucial evidence on the basis of which someone is sent to prison for a lifetime. A case of fingerprint identification or non-identification therefore carries a great deal of weight as evidence, and its factuality is seldom questioned in court. Due to several US Supreme Court rulings, however, fingerprint examination has been questioned and challenged and forensic service providers have been forced to consider whether the results of fingerprint examination can be assumed, in reality, to represent an infallible, error-free culture of investigation (see Ref. [1]). Also due to embarrassing fingerprint misidentifications, such as the Mayfield case in the USA and the PC Shirley McKie case in the UK, forensic service providers have confronted challenging situations requiring them to upgrade the level of their activities (e.g., Ref. [2,3]).

* Corresponding author. Tel.: +358 40 7016666.

E-mail addresses: Virpi.Mustonen@poliisi.fi (V. Mustonen),

kai.hakkarainen@helsinki.fi (K. Hakkarainen),

juha.tuunainen@oulu.fi (J. Tuunainen), pasi.pohjola@thl.fi (P. Pohjola).

¹ Tel.: +358 50 4624 058.

² Tel.: +358 40 5507 263.

³ Tel.: +358 50 4129 572.

Thus new transparency, productivity, and quality requirements are forcing the providers of forensic science services to find new ways of improving their efficiency without lowering the reliability of the investigation. Ethnographic laboratory studies (e.g., Ref. [4]) have revealed that actual laboratory work is much more messy, opportunistic, and contingent than the normative perspectives on forensic methods or retrospective accounts produced by examiners and other practitioners indicate. The actual link between variation in fingerprint feature selection and the examiner's ultimate decision-making strategies in fingerprint identification, especially in complex cases, is not well known or widely understood. Fingerprint examination often relies on ridgeology that was claimed to offer a more "holistic" and "scientific" basis for the identification of fingerprints than the "old" method of counting special minutiae [5,6]. Ridgeology with the ACE-V (Analysis, Comparison, Evaluation and Verification) process does not, however, give the fingerprint examiner the tools for producing arguments concerning decisions; it provides only a frame of reference for assessing critical aspects of the fingerprints being examined [7]. Moreover, a lack of validated and transparent thresholds is a recognized weakness of the ridgeology approach and the ACE-V protocol [8].

The present study focuses on analysing the fingerprint examination process in Finland, which involves making decisions on ambiguous and distorted fingerprints, where discrepant interpretations have been given by two independent examiners. The examiners are expected to state their conclusions according to Forensic Laboratory rules. They have to be confident regarding their final conclusions, because of which the rules for reporting discrepancies in fingerprint identification in Finland usually follow the more conservative (weaker) conclusion, such as "no value" (i.e., not identifiable) or "not identified". Examiners try to avoid false positives and minimize errors. Also the effect of the forensic culture makes these ethically followed, safeguarded rules [9,10].

Thus, 'not identified' is not an error, it can be a pass or a false negative. In order to avoid passes or false positives, joint meetings which include the examiners in question are organized to discuss the discrepant analyses. Discrepant judgements need to be understood and explained in order to see what is causing the discordance. Fingerprint examination methods need to be improved. Before explaining the Finnish fingerprint examination system in detail, we will briefly review the international research on expertise in fingerprinting.

2. Expertise in fingerprint examination

Learning to do fingerprint analysis requires sophisticated expertise acquired through years of training and practical experience. Expertise may be defined as individual mastery of a well-organized body of usable knowledge that a person utilizes to focus selectively on the critical aspects of a complex problem, and, thereby, reaches an exceptionally high level of performance in his or her trained domain of activity [11]. Expert knowledge represents the cultural-historical evolution of the professional domain and is embodied in social practices of expert communities and networks [12,13]. Professional expertise is, to a large extent, based on tacit or implicit knowing; experts know more than they can tell and often cannot provide reliable verbal description of their reasoning process [14]. This is especially true in such visualization-rich domains of expertise as interpreting X-ray images or analysing fingerprints, where pattern recognition plays a crucial role. Goodwin's theory of professional vision assists in understanding and explaining experts' sophisticated visual competencies in recognizing complex patterns when comparing target fingerprints (latents found at the crime scene) with filed prints. Examiners' internalized professional knowledge provides psychological tools [15] for separating significant cues and signs from

non-significant background. Such visual capabilities allow experts to make well-justified inferences relying only on partial and limited information. Goodwin [16] categorized professionals' visual activity according to three practices; coding, highlighting, and producing and articulating material representations. Coding by marking minutiae in latents enables making AFIS database searches and comparisons with the suspect's fingerprints. By highlighting disturbances and colour coding ridge flows (i.e. details) or minutiae, the examiner makes his or her perceptions visible to other experts. Constructing external material representations involves manipulating fingerprint images in various ways by using AFIS and/or PhotoShop and also by producing written annotations which ground and justify one's interpretations.

The very same mechanisms that allow an expert to detect patterns relying on fragmented and limited information, may, however, lead to mistakes and error. Human reasoning is biased by so-called confirmation bias [17–19], i.e., a tendency to selectively attend to evidence that supports one's own expectations or hypotheses. Rather than a one-directional flow of information from outside to inside, visual perception is a constructive process driven by a participant's expectations [20]. Due to the tacit or implicit nature of expert knowing [21], the fingerprint examiner may not be aware of the impact of heuristics and biases on his or her reasoning and decision making. Earlier studies have indicated a wide variation of results and interpretations in fingerprint case work. Technical errors often involve erroneous exclusions, missed identifications, and inappropriate inconclusive or no value decisions [22–24]. Experience as a fingerprint examiner, the quality of the latent, and expectations regarding the case being investigated affect examination in challenging and ambiguous decision situations [25]. Fingerprint identification is more vulnerable to internal and external elements and biases when the difficulty of the task increases [23,26]. Consequently, there are frequent occasions in which fingerprint examiners disagree about the same latent.

3. Research setting: the Forensic Laboratory of the National Bureau of Investigation of Finland

The present study was carried out in the Forensic Laboratory of the Finnish National Bureau of Investigation (NBI), which operates under the Ministry of the Interior. The Forensic Laboratory is one of NBI's divisions and provides a wide variety of forensic services to the law enforcement officials in crime investigation, such as the analysis of fingerprints, drugs, secretion marks, DNA, textile fibres, hair, paint, glass, gunpowder, handwriting and chemicals. The laboratory is also responsible for developing the procedures, equipment, and quality of crime-scene investigation in Finland. The Forensic Laboratory follows the standard of general requirements for the competence of testing and calibration laboratories SFS-EN ISO/IEC 17025. In 2014, the Laboratory had 120 experts representing different specializations.

The present study is grounded on efforts to improve the quality of fingerprint investigation. Traditionally, a fingerprint examiner's work has been very individualistic in nature, the examiner choosing the case to be examined, analysing all latents found at a crime scene, and working alone with the case until a research report was produced. Because of the idiosyncratic and opaque nature of an examiner's local analytic and interpretative practices, it has been very hard to make this work more transparent and subject to quality assessment. The investigation process embodied a great deal of tacit knowledge and also, potentially, hidden confirmation biases [25]. The examiner's perceptions, interpretations, and rationales in making decisions were not systematically documented. Discrepancies in challenging latents in identification between examiners were not documented in any way.

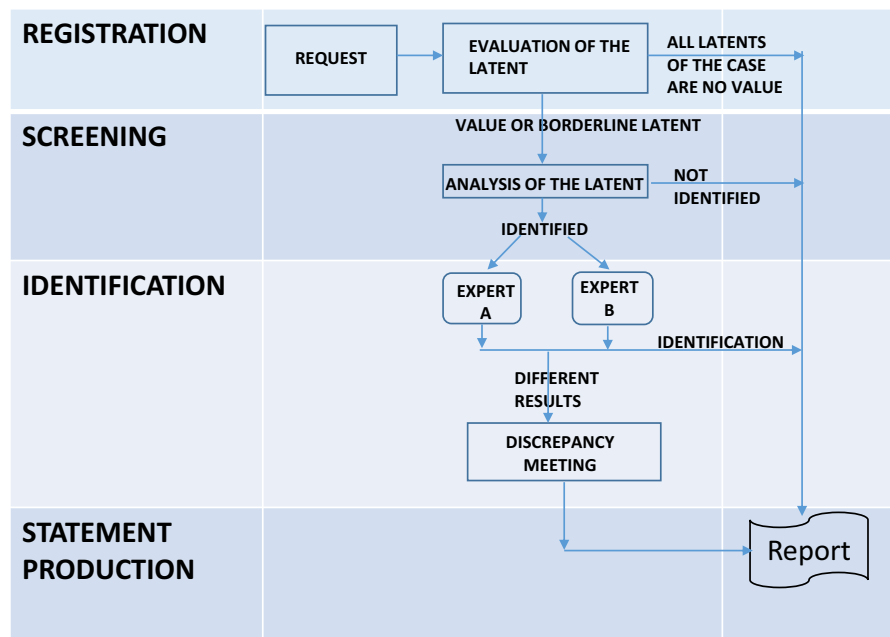


Fig. 1. Novel fingerprint examination process. There are 4–5 different fingerprint examiners involved in the process.

In 2010, the current investigative system was transformed in order to better meet quality requirements and make the fingerprint examination process more explicit, unbiased, and reasoned. The laboratory moved from analogical to digital fingerprint documentation data mediated by a Laboratory Information Management System (LIMS). The novel procedure involved independent examination and documentation of fingerprints by several examiners, according to a certain division of labour, instead of individually processing all latents related to a case. The new fingerprint examination process is divided into four different steps; registration, screening, identification and statement production (see Fig. 1). The previously mentioned ACE-V (Analysis, Comparison, Evaluation, and Verification) process used in the fingerprint domain shaped the new fingerprint examination process and provided an associated frame of reference.

The main data of the present study were from discrepancy meetings between examiners, together with the quality manager of the fingerprint group who also is the first author of this article. The former persons had independently analysed same latents and come to discrepant decisions regarding fingerprint identification¹.

The examination process starts with the *registration phase*. A competent fingerprint examiner records all the official information and makes an overall analysis and evaluation of all latents (fingerprints) in that particular case to the LIMS. If all the latents are categorised as “no value”, the examiner sends the case through the LIMS system to the statement production. If the latents are of “value for examination” (meaning value for comparison or identification), the examiner in the screening phase receive the latents for further analysis. During the *screening phase*, the fingerprint examiner follows the ACE process to analyse, compare and evaluate the information from the digital fingerprint (latent) provided by AFIS (Automated Fingerprint Identification System).

At the screening phase, the examiner analyses the latent and decides if it is of possible value for the examination process. After the decision to continue the process, the examiner compares the latent side-by-side to suspects’ fingerprints and legitimately handled persons’ (e.g., a person who has had permission to touch the source) fingerprints. If the latent matches with the fingerprints taken from a legitimately handled person, the examiner identifies it. If the latent matches with the suspect, the latent is sent to the identification phase. If the latent’s “owner” is still unknown, the latent continues to the AFIS database comparison process.

The comparison of fingerprints in the screening and identification phases relies on instruments provided by AFIS. Photoshop software is also recommended for making more precise annotations during the analysis process, such as annotating minutiae or ridge flows with GYRO colours where green stands for clear, good quality minutia, yellow for mediocre quality minutia, red for uncertain, poor quality and orange for minutia marked during the comparison stage (see Ref. [27]). The examiner’s analysis is always based on the images’ (i.e., fingerprints) general or detailed features, including the source of the print (i.e., a finger or a palm), its quality, pattern type, minutiae, possible noise, and distortions (see Ref. [28]) and so on.

In these phases, the latents are also analysed according to Level 1–3 features closely encompassed by ridgeology and related to the ACE-V procedure [29] used in the characterization of fingerprints: Level 1 includes the presence of the general ridge flow and pattern configuration; level 2 involves the presence of individual types and positions of formations, such as a ridge ending, bifurcation, dot, or combinations thereof; and level 3 all attributes of the accidental characteristics and intrinsic ridge formations [5,29], such as ridge path deviation, width, edge shape, pores, breaks, creases, scars and other permanent details affecting individual ridges as well as the proportion of and interrelationship between these minutiae. Examiners are also expected to document to the LIMS (Laboratory Information Management System) deposition pressure, Level 1–3 details, estimation of quality, occurrence of disturbances and so on.

After the comparison of the suspect’s and legitimately handled persons’ fingerprints, the examiner makes a database search to AFIS and compares and evaluates the latent to the registered

¹ Categories used in final decision-making in fingerprint examination in Finland in 2012: of no value = because of the lack of sufficient information in the latent, it cannot be used further in the examination process; an identification = two fingerprints that contain sufficient information so as to allow the decision according to which they share a common source of friction ridge skin; not identified = no fingerprint with same kind of friction ridge details has been found from the archives as the one under investigation.

fingerprint images in AFIS. The fingerprint examiner evaluates the fingerprints, and decides if there is a match between the registered person's fingerprint and the latent; if there is, the fingerprint is sent to the *identification phase*. At the identification phase, two separate examiners make an independent analysis and decision concerning whether or not the print can be regarded as an identification. After this, the final report is produced at the *statement production phase* and is sent electronically to the crime investigator. If the examiners still diverge in their opinions at this stage, the divergence is not revealed to the customer in the report, but the most conservative (lowest) category will be applied in the statement.

4. Methods

4.1. Participants and setting

The present study focused on analysing five audio-recorded “discrepancy meetings” in which two examiners (from the identification phase), who had come to different conclusions in the fingerprint examination process met together with the fingerprint quality manager (the first author of this article) to compare their interpretations and reconcile their differences. The number of the participating fingerprint examiners could have been three, if the fingerprint examiner from the screening phase also had been involved.

In all of the cases studied, the participating examiners had analysed the same latents found at the crime scene but had come to different conclusions regarding (a) the quality of the latents (whether they were of value for identification or not) and (b) identification. The discrepancy cases were selected from among hundreds of regular work flow cases of fingerprinting. Due to a deficiency in the Forensic Laboratory's reporting system, the statistics and total number of discrepancy cases in fingerprint investigation is not available.

The discrepancy meetings involved the participation of six fingerprint examiners altogether, some of whom took part in several meetings, thereby increasing the number of individual participants in the meetings to 10. The participants were professionals specialized in examining fingerprints: they worked in the Forensic Laboratory's fingerprint group together with a small number of qualified colleagues. They were very experienced, all of them having more than ten years of experience in fingerprint examination. There was an approximately equal number of male and female examiners. Three had Master's degrees and the others either had Bachelor's degrees or degrees from vocational colleges.

The first author, who has previously been a fingerprint examiner and forensic scientist and has been responsible for developing fingerprint examination practices at the Forensic Laboratory, chaired all the meetings. She did not take part in decision-making about the latents, but asked questions about the examination process, explored the participants' interpretations of latents as well as issues related to the development of fingerprinting methods. The aim of the discrepancy meetings was to examine reasons for divergent interpretations and associated decisions and discuss the examiners' perceptions and impressions of the latents. The aim was to resolve differences and reach a consensus conclusion. The examiners were encouraged to explain their thinking, actions, and associated rationale as well as reflect on the nature of their expertise. The discrepancy meetings were aimed at reflecting on the investigative process and collecting data about the examiners' decisions, judgements and ways of dealing with discrepancies. The aim was also to discuss the examiners' expertise and associated organizational practices.

4.2. Data analysis

The audio-recorded discrepancy meetings were transcribed word by word, which produced 59 manuscript pages altogether. The length of the meetings varied from 56 to 73 min, totalling altogether 5 h and 30 min of audio recordings. Those parts of the discrepancy meetings that addressed professional development in general and did not relate to the analysis of fingerprint cases were not transcribed. The transcribed data were analysed according to qualitative content analysis [30,31]. We read the transcribed discussions several times, trying iteratively to define central themes and recurring issues important for understanding and explaining fingerprint examiners' activities. With the help of ATLAS.ti qualitative data analysis software a large number of text segments were categorized representing three “families” of more specifically coded text fragments. The principal themes of discussion were: (1) interpreting fingerprint information, (2) manipulating images, and (3) making a decision. Several subcategories of content concerning each principal category were identified so that there were, altogether, eight more specified categories within the three families. Some initially considered categories were eliminated from the analysis because of ambiguous content or rare and unsystematic occurrence. Organizational quality reports and LIMS documentations (e.g., amount of minutiae or quality definitions) were used to support an interpretation of the discrepancy-meeting discussions. Quality reports comprised analysed fingerprint images with colour-coding and AFIS-screenshot documentations together with written descriptions of justifications of decisions regarding the identification process (Table 1).

5. Results

Various aspects of fingerprint analysis addressed in the discrepancy meetings will be examined in the next three sections. We will first address interpretation of fingerprint evidence; how examiners categorized latents and identified details. We will then move on to analyse various ways of manipulating fingerprints using different tools. Finally, we will describe the processes of making decisions.

5.1. Categorizing and analysing details

When reporting their findings about the latents to their colleagues in discrepancy meetings, the examiners often characterized fingerprints using very general terms. This characterization was based on the work done by the examiners at the beginning of the screening phase of the fingerprint examination process. At this phase, the examiners were directed by the Forensic Laboratory's guidelines to define the overall quality of the latents they had received through the Laboratory Information Management System (LIMS). The term ‘quality’ here refers to the visual clarity of the latent, an important precondition for the subsequent categorization. The analysis done during the screening phase and the coding of ‘features’ by means of latents was based on Ashbaugh's [29] methodology. According to this approach, known as *ridgeology*, the examiners were expected to analyse latents in terms of a three-level classification schema where Level 1 included the presence of the general ridge flow and pattern configuration (e.g., arch, loop and whorl); Level 2, the individual types and positions of formations visible in fingerprints (e.g., a ridge ending, bifurcation, dot, or combinations thereof); and Level 3, all the individual attributes of the ridges (e.g., ridge path deviation, width, edge shape, pores, breaks, creases, scars and other permanent details affecting individual ridges) as well as the proportion of and interrelationship between the foregoing minutiae. With Level 2 and 3 details, fingerprint examiners often also spoke about Galton

Table 1

Descriptions of main categories and sub-categories and frequency of coded text segments.

Main categories	Number of coded data excerpts	Sub-categories	Number of coded data excerpts	Description
Interpreting fingerprint information	160	Shaping outline	51	General features, but not specific definition of the feature, e.g., finger, palm print, toe, overall quality of the latent and noise, distortions, details belonging to levels 1–3
		Identifying traces	45	More detailed and specific information about the latent in question, e.g., “Galton” minutiae, pattern type, structure of the ridge flow, pores
		Identifying noise	49	Detailed information about distortions, disturbances or other abnormalities in the evidence or in the latent.
		Supplementing information	15	An examiner’s act of adding, picturing or modelling missing details of the latent with his/her own inferences
Manipulating images	44	Manipulating image	8	Technical treatment of the latent with the help of technical tools, such as AFIS or PhotoShop
		Marking or non-marking traces	36	Different forms of symbolic communication, e.g., colour coding.
Making decision	86	Making a judgement	69	An examiner’s specification of interpretation to assess the sufficiency of warranting a decision concerning identification (or making a compromise with another examiner during the discrepancy meeting)
		Making a decision	17	Changed or unchanged decision, reinforcement or abandonment of identification and rules of decision making

points or Galton details which refer to the various kinds of interruptions found in the ridges of friction skin, that is, ridge endings, ridge bifurcations, dots, short ridges and enclosures. Very often examiners referred to both of these types of detail by using the overarching term “special details” (i.e., minutiae). In what follows, we will first discuss the general classification of the latents, after which we will examine the detailed analysis of fingerprints.

According to the laboratory guidelines, the overall quality of the latents was to be classified as being generally either “poor”, “middle range” or “clear”, but no specific criteria were given for defining when a particular latent belonged to these categories. Because the categories thus remained undefined, the examiners used their visual competencies and reasoning to classify the latent prints. In addition to this, the examiners occasionally created categories of their own. One examiner, for instance, noted that she or he almost always defined latents as being “changeable” as they were usually clear in some parts but unclear in others. The other examiner said, in turn, that she or he called a fingerprint “normal” when “it was clear” and “relatively uniform in quality”, but that prints like that could almost always be characterized as “changeable”, too. In one case example, two examiners had looked at the very same latent and defined it differently, one characterizing it as “weak in quality” and another saying it was partially “quite clear”, thereby indicating the internal variance in the latent’s quality. The use of these kinds of terms was common in our data. It also suggests that the initial categorization of fingerprints was done by using ad hoc criteria that were only weakly connected to the general-level organizational standard. In another words, the general classes and the classification criteria had to be created (or recreated) simultaneously as the examiners looked at the prints on their computer screens (for “localization”, or redefinition, of general rules in practical action, see Ref. [32]).

Because the decision-making criteria were unstable, the initial categorization of latents sometimes differed from one examiner to another, a condition which was addressed in the discrepancy meetings. In the data excerpts cited below, two fingerprint examiners described why they had classified the same latent in different ways. In one case, for example, Examiner A thought the fingerprint was “weak” in quality and contained so little information that it should be regarded as “unidentifiable” (Data

excerpt 1), while Examiner B considered it “quite clear” and therefore found “fairly many” “special details” that paved the way for her/his conclusion that the print was indeed of value for identification (Data excerpt 2):

Data excerpt 1.

Examiner A: If I start from the overall image... First, there is a lot of noise, ridge flow is falling apart in many places. (...) The pattern type is a right loop or a double whorl. [The noise] on the right edge might confuse me, because here it goes like this... under the loop there is much noise. All of the edges are unclear, and then there are some Galton points. (...)

Chair: Did you decide that you would not compare [the latent with the ten-print card]? Did you analyse only this [print and conclude it was] unidentifiable?

Examiner A: Yes, I didn’t make a comparison. I highlighted those areas with red, there in the middle of the latent where the ridge flow is divided. Here, on the right side, there are a couple of minutiae, but connecting them here, to the left side, isn’t possible. There are so few that it wasn’t enough.

Chair: What are those that were so few?

Examiner A: Minutiae, or information.

Later during the same meeting, Examiner A further explained what she or he had meant by the lack of information: the latent had lacked distinctive features of its own that would have increased its “individuality”. Therefore, it earned the qualification of “a poor” and “unidentifiable” print: “That latent was poor in quality. It didn’t have individuality. There were these [special details] only: bifurcation, ending. (...) The ridge pattern and the area in the middle, which was visible, were basic. I would have required something more reliable.”

Interestingly, Examiner B decided to concentrate on the better areas of the latent. In doing so, she or he found “a considerable number” of identifiable special details from the fingerprint:

Data excerpt 2.

Chair: Well, what do you think?

Examiner B: I concluded it was identifiable. (...) Here, I can remember one minutia, which caught my eye. I have highlighted it with a pen in this drawing. When I started, of course, I looked at the pattern type, like you, and I concluded, at this stage, that it was probably a right loop (...). Just like you, [I regarded] the right edge [upper part] and the centre [lower part] as areas where I didn't even go. (...) I concentrated on the better areas and, yes, found quite a lot of minutiae (...).

As illustrated above, the examiners used their personal judgement to decide whether or not the given latent contained enough information for analysis to continue beyond the stage of initial fingerprint categorization. That is, if an examiner decided that the quality of the latent was “poor” and therefore “unidentifiable”, it would no longer be worked with; but if it was regarded as “identifiable”, the work could continue to the next phase, which was the detailed analysis. The question culminated in the drawing of boundaries within which one would consider it justifiable to make one interpretation rather than another, a task that, according to one examiner, was “really difficult”, because no clear rules existed “and could not exist”:

Data excerpt 3.

Examiner B: This is just defining one's own borders. I wouldn't decide with those minutiae that [the latent was] unidentifiable. You have five [details] which you are quite sure about. I would have continued with those criteria. If there is a ridge flow as clear as here... I very seldom decide the print is unidentifiable with few minutiae (...). I have to say that this wasn't [even] a borderline latent; I have marked and identified [even] worse. It is a problem for me if ridges are poor and scruffy. Here the ridges are so strong. What I first felt was that this was a usable latent. But it depends on so many things; how one sees it and where your personal border is. Nowadays, this goes according to one's feelings.

Examiner A: (...) the clear area was so small [and] the ridge flow broke in so many places. (...) There wasn't enough individuality.

Chair: I still challenge: What kind of individuality would you have needed to proceed to the next phase?

Examiner A: If there were more clear minutiae in a little area, or a short ridge, loop or something like that, or the strength of the ridge...

When it comes to the detailed analysis of the latents, the examiners appeared to use Ashbaugh's Levels 1–3 as psychological tools [15] to separate off significant cues and signs from their background. The criteria, however, did not provide examiners with stable operational meanings that would have unproblematically guided their practical action; rather, the application of the classification schema was part and parcel of the fingerprint analysis itself and had to be “localized” [32] to make it useful in any given situation. To put it another way, the examiners used their visual competencies to decide whether or not the given fingerprint exhibited qualities that belonged to a certain general category provided by the organizational documents. In the following case example, an examiner used her/his practical judgement and contextual information to decide that certain parts of the fingerprint should be classified as belonging in Level 1 and 2 details of Ashbaugh's schema rather than Level 3:

Data excerpt 4.

(...) One can see a clear contrast here at the edges of the ridges [and] between them. Yet, I did not mark it as Level 3, because I wasn't convinced. I wasn't convinced if this was the edge of the ridge or not. The pressure has been stronger here, in this area, as ridges [have] become much wider here, but I didn't mark it as Level 3.

Here, the examiner interlinked the details she or he had found in the latent with other identifiable minutiae present in the print, namely, the strength of the pressure that was used when leaving the fingerprint at the crime scene. When doing so, the examiner also compared ridges and their edges present in different parts of the print to decide whether the identified minutiae were to be classified as belonging to Level 3 details or not. In general, the examiners thus actively interpreted the information they were able to extract from the latents simultaneously as they complemented the information with their pre-existing knowledge about various factors that usually affected the fine characteristics of fingerprints². Because a great deal of this work appeared to rely on tacit professional knowledge of experienced examiners, it was not extensively described and reflected upon during the discrepancy meetings.

The situation in which the examiners ended up with variant decisions when analyzing the latents seemed at first to result from their differences in decision making. Yet this was only a part of the picture. Another important issue was the lack of an explicit method of investigation that would be easy to follow in any concrete situation. As we have tried to argue in this article, it is unlikely that such clear rules of method can ever be constructed so as to allow easy reproduction of analyses in such information-rich areas as fingerprint examination. It is therefore likely that in spite of all the efforts that have and will be made to develop investigative practices in fingerprints, there will always remain a certain amount of disagreement and variation in the work of experts, such as the ones analyzed here (for relevant analyses, see Refs. [7,18]).

In addition to applying general classifications of fingerprint details, the examiners were also involved in mentally repairing bad latents. This constructive perceptual activity was closely associated with the characteristics that hampered identification of relevant details in fingerprints. Such visual defects ranged from the movement of the finger to the effects caused by detection and visualization techniques used in fingerprint examination. Furthermore, some biological characteristics of the skin also created flaws that complicated the perception of details in fingerprints: pores, for instance, occasionally disrupted the flow of ridges so that they were seen as being discontinuous. In cases like this, the examiners typically reported the ridge flow as “having been broken”, “being messed up” or “fallen apart”, meaning that the ridges did not flow in an orderly manner as they were expected to do.

When complementing missing information, or mentally repairing visual distortions, the examiners used three specific methods of reconstructive inference: first, they assumed details that were not at all or were only barely seen in the images, second, they completed the distorted images with the help of logical reasoning, and, third, they imagined distortive flaws away from the latents. In all of these alternatives, the visual competence of the examiners helped them to identify significant details in imperfect or defective latents and repair those parts of the images that were

² Beyond identifying Galton or Ashbaugh's details, the examiners also referred to a wide range of other minutiae they had found when analysing the latents. These details were usually related to ridges and patterns they formed in the fingerprints. For instance, examiners reported in discrepancy meetings that they had “counted” the number of ridges between certain Galton details or “followed the flow of ridges” between them.

faulty. Arguably these experts did not simply misrepresent reality; these professional actions enhanced partial and fragmentary marks so as to make recognition of significant patterns and details possible. In the following, brief examples of these categories are given:

In the first example, an examiner presupposed certain fingerprint details that were only barely visible in the actual image. The examiner describes the way in which she or he supposedly saw the pores and the edges of the ridges in the fingerprint she or he was analysing:

Data excerpt 5.

Here, in the core area, are visible ridges. Supposedly, there are edges of the ridges, and the pores. It is difficult to say if the ridges are white or black...

The second form of reconstructive inference was logical reasoning. Relying on his or her visual competence, the examiner in this case filled in the missing parts of the image and proposed how observed patterns of ridge flow must continue in an orderly manner, although this was not visible in the actual print. For instance, the examiner explained to others the way in which she or he followed the line formed by ridges and came to the conclusion that there had to be two more lines and a special detail in the image, although they could not be seen:

Data excerpt 6.

Yet, here I did so that when I started to move from here, I followed these ridges. It has to be said that this ridge must be here in the middle. One ridge has to be here and there must be another one. Then I got from there to here, and realized that this has to be a minutia. Perhaps, it is not exactly here, but somewhere in this area.

Finally, the third form of reconstructive action was to imagine distortive flaws away from the image. In these kinds of cases, the examiners mentally deleted distortive parts of the latents so that the significant details could be better grasped:

Data excerpt 7.

Here, this [ridge] breaks all the time, but still it is ... I can see it as an unbroken line and it does not have any disturbance that would confuse me.

In summary, the initial categorization of fingerprints and identifying significant details in them was visually constructive activity that was aided by psychological tools that examiners had learned to use in a skilful manner during their socialization to the fingerprint work community. Although some parts of these tools were embedded in formal organizational, local guidelines, these directions were so vague that they did not provide examiners with detailed guides for practical decision making. The examiners, therefore, had to use a great deal of ad hoc reasoning and reconstructive visualization to manage the work tasks they were engaged in. The practical fingerprint analysis was not, however, mental activity only but involved a strong component of processing the images of fingerprints, such as highlighting selected areas in latents and colour coding their key details. The use of these external tools in fingerprint analysis is the topic we will next address.

5.2. The use of tools in fingerprint examination

Following the initial categorization and coding of fingerprints, the examiners utilized various technical devices to examine

fingerprint images (of latents) more closely. The technical devices for manipulating the images were used to extract more detailed information about the images, but also to colour-code and annotate the images. These external tools (PhotoShop and AFIS) acted as lenses through which the examiners constructed facts about the latents being analysed.

In the cases discussed in discrepancy meetings, the examined latents were in one way or another lacking full detail and quality. Because of that fact, identification of the minutiae and details concerning these latents was obviously challenging. Sometimes the latents had disturbances that made them either fuzzy and muddled or broken and partial. Such effects were caused, for instance, by dust, blood or finger movement. To construct representations or arguments from this kind of incomplete information of the latents, requires an ability to construct “objects of knowledge that become the insignia of a profession’s craft” (Ref. [16], 606).

The professional vision required for creating objects of knowledge became manifested through varied practices which were present in the discrepancy meetings. Examiners discussed their actions regarding latents, i.e., how they had enlarged the images, made them darker or lighter, changed the contrast, reversed colours, or otherwise made the image more illustrative for examination. There were, of course, practical limitations regarding what the examiners could do with these tools. The quality of an image limited the possibility of enlarging it or created technical noise which affected the examiners’ ability to analyse the latent. We say that this kind of highlighting, following Goodwin, is a professional practice that “makes specific phenomena in a complex perceptual field salient by marking them in some fashion” ([16], 606).

Another professional practice in Goodwin’s work was the production and articulation of material representations. Many of the discussions in the discrepancy meetings relied on this type of professional practice. Annotations also helped to share examiners’ perceptions and interpretations of the analysed latent with one another. Some examiners did not draft any documentation concerning identified minutiae or even used colour-coding in the images, but wrote a verbal report about how many minutiae he or she had seen when analysing the latent and comparing it to the ten-print card. Other image-manipulating tools, such as those provided by AFIS, were available, but they were not mentioned during the meetings.

The third professional practice in Goodwin’s work was coding, the transformation of phenomena observed into the objects of knowledge. The examiners marked traces involving colour-coded distortions or minutiae, marked minutia with arrows or other symbols, or printed AFIS screenshots. Colour-coding relied on the GYRO procedure explained in Section 3, but the Forensic Laboratory’s quality instructions did not require the use of colour marking in the analysis. The examiners used defined colours for marking certain, uncertain, and problematic minutiae and parts of the fingerprint features, such as bits of noise (Fig. 2a and b). Examiners marked green “annotations” to highlight a high level of confidence about existing features in the latent, yellow was used for a more uncertain medium level of confidence and red represented a great deal of uncertainty (see the red-coloured areas in Fig. 2a). Orange indicated in which phase the highlighted minutia had been noticed and marked.

The coding schemes, such as the four colours of the GYRO procedure, are a systematic professional practice of classification. Although the GYRO procedure can be understood as a coding scheme of abstract and universal cognition – a standard way of colour-coding the observed phenomena – it becomes evident that it is actually a situated and disparate constitution of knowledge [16]. The examiners’ practices of using the colour-coding of the

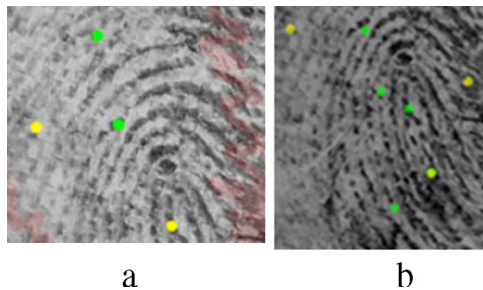


Fig. 2. (a) and (b) Two representations of the same latent differently colour-coded by two independent examiners. On the basis of NBI safety policy, figures are manipulated to preserve their anonymity. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

latent differed both in the number and quality of minutiae. In Fig. 2a, Examiner A had identified 5 minutiae, of which two were coded with green (certain) and three with yellow (uncertain). Examiner A had also used red for areas with disturbances. In Fig. 2b, Examiner B had colour-coded altogether 11 minutiae, of which five were with green (certain). Of the 11 minutiae identified by Examiner B, five were similar to the minutiae coded by Examiner A. Interestingly, the examiners' views regarding quality differed in three of the five shared minutiae. Only one minutia was colour-coded as certain (green) and only one uncertain (yellow) by both of the examiners. Fingerprints are information rich, hence examiners could naturally identify different minutiae; this is not a problem when the image quality is high but may lead to different interpretations in cases of poor latents, as illustrated in the next discussion excerpt.

Data excerpt 8.

Chair: Could you define or did you look how many minutiae there are?

Examiner A: Yes I looked. Two are green and three are yellow.

Chair: On what basis did you highlight green and yellow?

Examiner A: Well the greens, there are two of them; there is a bifurcation and then a ridge ending. It does not seem to be certain whether the yellow continues. Here is a thin ridge and a strong ridge that is a red flags (noise, distorted) areas in many places. Here are some very dot-shaped ridges. There is something strange here, because sometimes the ridges are ok, but then their thickness becomes totally different. I have written there that there is not enough information for identification.

[...]

Chair: What does B say?

Examiner B: [...] I have highlighted in this image 6 certain and 5 uncertain minutiae, but that one [gesturing], looking at it afterwards I could say it is also uncertain, more likely than certain.

Instead of merely being conventions about different colours to be used in marking traces in fingerprints, coding schemes also established the perspective of how the professional views the world. As the excerpt above shows, the schemes established were not entirely objective, but things that could be disagreed on

(to some extent). On the other hand, they also provided a distributed cognition structure for sharing and organizing action [16]. A coding scheme therefore provided a structure for negotiating the disagreements about latents in discrepancy meetings.

In the processes of coding the images, the examiners had to extract information from the disturbed latents through highlighting some relevant part of the image in order to make the information salient. Also, as with coding, the examiners had different ways of conducting this professional practice, such as the ways in which they manipulated images. Examiner A, for example, noticed in the beginning of the discrepancy meeting that image 1b was darker, and Examiner B addressed explicitly how she or he had manipulated the image and tried to extract information from it.

Data excerpt 9.

This image is a good example. I didn't bring it out, I made an enlargement of it, but then the problem was that the image pixelated.

Highlighting, as a professional practice, produced representations or what Lynch [33] has called the "externalized retina". It is a way of structuring the world of vision and scrutinizing it to reveal something. At the same time, it divided the image according to what was thought to be relevant to the professional activity and what was not. As with coding, the ways of executing this professional practice differed between examiners.

To summarize, the examiners processed the fingerprint images in various ways in order to enhance them and make minutiae easier to identify. They also coded the images by using the GYRO procedure so as to make a difference between the poor and good areas, and to highlight selected minutiae and details. Their diverging ways of manipulating the images appeared to affect, to some degree, their later interpretations and decisions. The manipulative process, that is, the path from the original to the enhanced image on which the final decision relied was not documented in any way. Although processing facilitated the analysis of the latents, there were problems related to digital image processing. During the discrepancy meetings, examiners shared concerns about this, for instance by saying that "this image is pixelated" or that image is "inaccurate". Fingerprint examiners considered these kinds of obstacles as everyday "technical noise" in their work.

Much of what was present at these discrepancy meetings relied on the practice of production and articulation of material representations. In the discrepancy meetings, the question was not so much about the production of material representations but was more linked with the two practices discussed above, coding and highlighting. In the meetings, the question was more about the articulation of these material representations, such as Fig. 2a and b. The figures were representations of what was especially relevant. In the discrepancy meetings, the focus was more on the examiners' articulation of what was there and what was relevant for each participant.

5.3. Patterns of fingerprint examiners' decision making

Fingerprint examination consists of a series of judgements in the matching process, and a fingerprint examiner must come to conclusions during this process. Decisions are based on previously described examiners' perceptions and interpretations when categorizing and analysing details and using different tools during examination. Examiners have to proceed with fingerprint examination according to the ACE-V protocol and they have to follow an organizationally defined examination process (see Fig. 1).

According to it, examiners must make decisions about the utility and identification of latents. The examiner must also determine whether or not the latent and the ten-print card found from the register are from the same source, that is, whether they are “sufficiently similar” with one another [17]. Examiners also have to assess the quality of the latent to continue from one phase to another in the examination process as exemplified in data excerpts 1–3. When an examiner has decided that the quality of the latent is “poor” and the print therefore “unidentifiable”, the analytic process comes to an end. Furthermore, decisions, based on examiners’ perceptions and interpretations, were supposed to be grounded on organizational rules, although they were quite vague for practical purposes as claimed earlier.

The quality system of the Forensic Laboratory dictated that the identification of the latent in problematic cases always followed the lower conclusion (“no value” or “not identified”). When such conclusions were made, examiners typically based their judgments and decisions on not having enough information; or, they claimed that the information present in the latent did not confirm a finding. They also reported that it was “easy to go according to the rules and for the lower (more conservative) conclusion”. Discussions regarding borderline latents indicated, however, the personal nature of the decision-making process and left the ultimate tacit knowledge hidden from others. The following excerpts show that the examiners had to adapt their decisions to the organizational rules concerning the final decision (identification, not identified, no value). In cases like this, examiners were generally aware of the discrepancy in information between the decisions and statements that were sent to the police, as will be illustrated below:

Data excerpt 10.

Examiner A: ... this possibility should be clear to the police and written down somewhere. If the result is a “no result”, there might still be strong evidence of identification, but we can’t be sure. Is it wrong to give such a piece of information to the police, but it is still a reality. We have no clear border.

(...)

Examiner B: ... in this case, I would have accepted something like “a hint of information” as a conclusion.

The examiners also indicated grounds for abandoning pieces of evidence when making decisions. The latents were often visually complex due to the poor quality with lots of noise and distortion, and the examiners usually ended up with “suitable” conclusions based on extraneous and case irrelevant contextual information, such as the type of crime, the amount of other evidence or the number of identifications in the studied case. Thus, the examiners appeared to have different decision thresholds based on contextual information and, therefore, made their decisions on the basis of varying reasons, for instance, time spent on analysing the latent, the number of other identifications in the case, or some another piece of background evidence. Data excerpts 11–14 illustrate the examiners’ contextual approach in justifying their decisions.

Data excerpts 11–14.

If this was the only latent in this case, I would have done more analysis, but in this situation I did a quick analysis.

(...) this is enough for me, it [the final decision] depends on Examiner A. However, there is no significance to the report. There were many other identifications in this case.

(...) it is very likely it is the suspect’s print because there are other identifications of him.

(...) in big cases, I admit, if I already have identifications, my threshold to exclude the latent becomes lower (...)

The analysis conducted in this article indicates that decision making was generally based on the minutiae, quality of the latent, evaluation of distortions, organizational rules, and also on contextual information. However, in documentation reports, decisions were usually justified by referring to the sufficiency of information regarding identification, the identification of a certain pattern type and ridge flow, and the presence of different kinds of minutiae. The actual documentation practices thus turned out to be highly somewhat inconsistent. Examiners routinely documented relevant aspects of fingerprint analysis into the LIMS system, but did so in heterogeneous, unsystematic, and, in many cases, highly fragmented ways. For instance, one examiner used only the image without any coding or written annotations whereas another applied GYRO colourings and wrote long explanations of her or his annotations. Most examiners, however, documented their work somewhere in between these two extremes. Furthermore, each examiner decided on the basis of his or her own experience what to document and how to express his or her observations. The factors which led to “no identification” or “no value” were only briefly described with comments like “definite comparison is not possible”, “the ten print card is inadequate” or “there are not sufficiently consistent minutiae to make an identification”. The conclusions examiners originally had ended up with concerning one and the same latent stretched from “identification” of the latent to “no identification”. However, consensus was always reached in the meetings as pre-required in the Forensic Laboratory’s quality system. According to it, the final decision was determined on the basis of the lower (more conservative) conclusion, meaning that if there were “identification” and “no identification”, the NBI’s final statement was “no identification”. All meetings analysed here ended up with the conclusion of “no value for identification”. During the discrepancy meetings, the possibility of having the category of “inconclusive” for a decision was discussed by the participants and relevant situations corresponding to this decision were identified. As a result of this, an initiative to alter the range of the Forensic Laboratory’s reported decisions was made and the categories of final decisions were altered. Nowadays it is possible to end up with the decision of inconclusive.

6. Discussion

The purpose of this study was to analyse professional fingerprint examiners’ investigative practices in the context of conflicting fingerprint examination processes. The aim was to find the ways in which the repeatability and accuracy of investigation in marginal, challenging fingerprint cases could be improved and increased. The data were collected from discrepancy meetings between examiners who had ended up with different independent decisions about the same challenging latents. The examiners’ written and annotated documentations supporting their decisions and the data from the LIMS system were also used. The data collected enabled us to observe, understand, and examine how examiners constructed their decisions in discrepant cases of investigation.

The results indicated how the examiners’ practical competencies allowed them to use partial and limited information in making judgments about the latents. The examiners used their professional tools and practices to interpret fingerprint information, manipulate images and guide their interpretation and decision

making. Their knowledge and experience of depositing fingerprints assisted in reconstructing and interpreting the latents they examined. Supplementing information represented examiners' constructive work with picturing and modelling latents' missing details using their personal inferences. Mental imagery appeared, in turn, to allow them to identify significant details in blurred, distorted, or otherwise messy latents. Rather than distorting reality, it enhanced partial and fragmentary marks in order to make recognition of significant patterns possible. Relying on their visual competencies, the examiners filled in the missing parts of the image and inferred, for instance, how observed patterns would likely continue, although they were not clearly visible as such. Accordingly, the examiners routinely reported to their colleagues that they had repaired the defective or otherwise problematic images by supplementing them with missing information. Although this instance of probable perception and interpretation had a minor value in the final decision making, it revealed the constructive nature of fingerprint examination as well as potential biases in the examiners' work. Such kinds of mental picturing and modelling in fingerprinting should be more researched in the future.

In the discrepancy meetings, the examiners reported to their peers that their opportunity to see and analyse essential details of fingerprints was often hampered by a multiplicity of intervening reasons and the effects of detection and visualization techniques, such as fingerprint lateral movement or dottiness due to a chemical agent. In order to overcome such challenges, the examiners manipulated digital images in various ways, for example, enlarging them, changing darkness, enhancing contrast, reversing colours, and so on. These image-processing practices were intended to make the significant patterns more easily recognizable. The examiners also marked good and poor quality areas in latents as well as identified minutiae by colour coding. Although divergent ways of manipulating (e.g., identifying different minutiae and details) in good quality latents does not affect the result of an investigation, it may have an important effect in cases of poor quality latents. As revealed by the analysed discrepancy meetings, the participants often manipulated the images differently, focused on different aspects of the latents, and, consequently, ended up with different interpretations.

The data at our disposal also showed, at certain times, that conclusions differed from one examiner to another. This is a typical and expected situation in complex investigative cases, and similar results have also been reported in other studies (e.g., Refs. [9,34–36]). Moreover, the use of informal characterizations that were interpreted differently by examiners and diverged from the norms of the Forensic Laboratory, was a common feature in our data. The data, for instance, revealed that the general characterization of fingerprints by the examiners was mostly done by using personal criteria that were only weakly connected to the overall organizational criteria and standards. Furthermore, the criteria used by the examiners were usually not made explicit in the discrepancy meetings. The prints were simply said to belong to one category or another, such as “no value” or “not identified”. Because of the limited guidelines for categorizing fingerprints, the examiners had to recreate the classification criteria as they analysed the prints. Quite often this also led to incongruity between individual interpretations. Because the categories used by the examiners were not particularly clear to them, especially in the analysis and the comparison phases, it appears desirable that the fingerprint laboratory's standard operating procedures should be revised.

Our study addressed discrepant decisions of problematic latents selected from among a larger number of fingerprints analysed. Although the data represented the actual investigative

practices of the involved examiners, there were certain methodological limitations we have to acknowledge. To begin with, we analysed only a small number (5) of discrepancy meetings involving six examiners, some of whom were involved in discrepant cases more than once. Fingerprint examiners were also sometimes reluctant to expose their activities to critical scrutiny by colleagues. However, the present participants engaged in rich interactions that provided relevant information of the rationale of their investigative activity. The meetings thus succeeded in creating a trusting atmosphere that encouraged the participants to talk openly about personal and collective challenges in fingerprint examination. A great deal of professional work appeared, however, to be tacit or implicit in nature. Due to the visual-perceptual nature of the knowledge in question, the examiners did not verbalize the process of their actual visual reasoning to any large extent, and mostly engaged in meta-level discussion of the process. In the analysed discrepancy meetings, the examiners typically reported to their peers only the number of minutiae they had observed in the latents, such as, “I saw nine relatively good minutiae in that crime scene print”. They did not talk much about specific aspects of the minutia in question, apparently assuming that the same things were easily visible to the other examiners as well.

7. Conclusion

The purpose of our study was to analyse fingerprint examination processes that involved making decisions on ambiguous, challenging fingerprints with discrepant interpretations by two independent examiners. The findings indicate that investigative practices in the forensic laboratory play a significant role in decision making and affect the reliability and validity of all stages of the criminal justice process. This investigation was carried out in the context of organization transformation that involved moving from the personal examination of fingerprints to a collectively distributed one. Newly established investigative procedures allowed us to identify cases where two independent examiners ended up with divergent interpretations. Our investigation indicated that Goodwin's framework of professional vision fits very well in examining the fingerprint examination process; the framework made various processes related to coding, highlighting, as well as document production, visible and available for reflection, allowing them to be subject to critical scrutiny. We believe that the results of this study may well be utilized to create developmental interventions to enhance other forensic areas, such as DNA, drugs, tool marks, shoe prints, bloodstains, handwriting, ear prints, bite marks and tyre marks. The theoretical framework for analysing professional vision in practice, therefore, appears to be an excellent tool for these areas. To some extent, this challenge is shared with the DNA complex mixed profile interpretations, with incomplete “peaks” in the profile [37].

Our results indicated how significant the fingerprint examiners' knowledge and experience were in making reliable and valid fingerprint analyses. It was possible to notice how the terminology and theory of fingerprint investigation shaped the examiners' everyday actions and discourse. Over a decade ago, fingerprint examiners did not have shared concepts to describe their investigative practices. There has been a huge advance in investigating various aspects of fingerprinting and identifying associated biases and challenges. The emergence of collectively shared conceptions, such as the Ashbaugh [29] framework, provides fingerprint examiners with a voice to address the various aspects of their investigative processes. It assists in explicating semi-automatic practices of pattern recognition, making the nature of their thinking and decision making more explicit and transparent. The examiners showed their personal abilities and

their ethical character in making identifications by following the rules despite personal disagreement with the decision of another examiner. This raised the question of whether or not these kinds of organizational rules restrict the construction of professional vision. In Finland, there were, during the time of the study, only three possible reported results of the fingerprint analysis: *no value, identification and not identified*. During the discrepancy meetings, examiners recognized the possibility of reporting an inconclusive decision. This is in contradiction with the present organizational decision rules and offers NBI an opportunity to reflect on its current organizational practices in the future. At the moment, 'inconclusive' decision is one of the Forensic Laboratory's reported results.

The discrepancy meetings illustrated the personal, organizational and technical factors (i.e., bits of "noise") that influenced examiners' work in various ways. Personal factors consisted of talk about work flow, tiredness of eyes, competence challenges, personal mistakes or other personal difficulties in handling latents. The organizational factors raised some ethical and moral questions, such as how the work should have been carried out, concerns about concentrating in an open-office environment, but also contradictions between different phases of the analytic process (screening and identification). These contradictions tended to be caused by the lack of common quality criteria and variation in perceptions and interpretations. Some of the challenges encountered by the examiners appeared to be caused by new digitally mediated tools and environments of fingerprinting.

In recent years, fingerprint examiners' work has become a technology-mediated process because the main bulk of fingerprint evidence material has changed from paper, ink, photos, and tapes to a digital form. While fingerprint examiners previously worked with concrete, physical evidence and instruments (such as a magnifying glass), currently they work, almost all the time, with digital instruments and images, and all work is done with the assistance of computers. In earlier times, physical photographs were inaccurate, too small, difficult to handle or new and better pictures were slow to get due to the physical, wet photograph process; in the present, digital-image period one sees similar kinds of "noise" in a different form. In decision making, these kinds of problems have to be taken into consideration, and the examiners' associated perceptions have to be better documented in the future. Observed discrepancies between traditional individual practices of fingerprint examination and the new way of doing collective work, reflected challenges in transforming the work processes.

As all forensic service providers know, fundamental methodological factors and human errors as well as errors caused by instrumentation and technology play a significant role in the fingerprint context [22,24,26,38]. There is concern that errors – which include bias – affect challenges, as well as apprehension about simple errors in documentation. In order to improve the reliability and validity of fingerprint examination, documentation practices should be transformed. Examiners are now expected to document their analyses, produce high-quality reports, and in quite a few cases, present them in court. Some of the examiners failed to document justifications for their decisions. When present, the documentation revealed how the examiners' perceptions and interpretations in discrepant latents drastically varied.

Implications of the main findings in Neumann et al. [35] report show the same kinds of results.

Simultaneously with the analysed discrepancy meetings, developmental seminars for all fingerprint examiners were organised by the first author. These seminars and discrepancy meetings were overlapping and influenced all developmental actions done in the fingerprint laboratory of NBI at the time, for instance, the further development of instructions for the documentation and methodological guidelines to be used in analysis. Despite this progress, drawbacks in the developmental efforts were encountered. Among these were issues concerning criteria to be used to categorize clarity and "analysability" of latent prints that arose from the discrepancy meetings. The NBI's fingerprint group did not find common criteria to be used in the analysis. An alternative approach that might provide NBI with a successful avenue to proceed with might be the use of clarity maps described by Hicklin et al. [39]. Providing examiners with tools to assess the clarity of friction ridges in these maps might help the organisation to make use of latent prints in situations where there are several identifications of the suspect.

Our analysis also suggests that the fingerprint community should work toward development of clearer criteria for decision making and define the corresponding documentary practices. Although the criteria can be explicated and improved, it may well not be possible, due to the complexity of fingerprint data, to develop rules that would cover all possible situations. It should therefore be considered whether the laboratory should move toward probability estimations, error-rate calculations or other measurement methods likely to improve the reliability and repeatability of fingerprint investigation. Statistics alone will not help, however; rather elaborate methods and criteria should be put in place for the selection of acceptable features that are used in probability calculations. Also the forensic laboratory should explore the possibilities offered by applying the methods of Bayesian reasoning. In short, forensic service providers do well to continue to do research, explore various decision-making models and promote technical development so as to reduce the variation of interpretations between examiners. The ultimate challenge is, therefore, to build a new community culture to elicit more reflective and communicative professional practices in order to provide a level of quality which meets the increasingly rigorous organizational and international requirements.

Acknowledgements

The first named author has had the main role and was responsible for collecting the extensive data, coding and analysing it. The other authors have participated as supervisors, discussing with and providing knowledge to the first author. The first author has been responsible for the article as a whole and the other authors have contributed by analysing their own parts of the data and by participating in designing the theory part of the article. All authors have approved the article. The first author has had financial support from the Finnish Education Fund in writing this article and the University of Helsinki has supported the article by providing the language revision service.

Appendix A

Descriptions of families and categories.

Families	Categories	Description
Interpreting fingerprint information	Shaping outline	This shaping outline is a very general category. General features were referred to in discrepancy meetings, but specific definitions of a feature were not given. These general features were referred to: anatomical orientation, presence of Levels 1, 2 and 3, substrate, development medium, deposition pressure, lateral movement, rotational movement, quality of the latent, ridge flow, presence of pattern, distortions, core and delta, large or little quantity of ridges between core and delta and tolerance
	Identifying traces	More detailed and specific information about the latent in question e.g. structure of the ridge flow, calculation of ridges between minutiae, "Galton" minutiae, pattern type, scars, warts, pores, incipient ridges and dots, creases and special combinations and specific shapes of minutiae
	Identifying noise	All referred, detailed distortions in latent e.g. pattern type is broken down or divided, ridge is dotted or discontinuous, double pressures, caused by matrix, dust, blood, latents' edges
	Supplementing information	Expert supplements information by picturing and modelling the latent's missing details with his/her own inference
Manipulating pictures	Manipulating picture	Latent technical treatment with AFIS or Photoshop e.g. darken, zoom, print, colour-code, add arrows, use lasso tool, change contrast or flip ridges from black to white
	Marking or non-marking traces	Forms of symbolic communication e.g. colour-coding distortions or minutiae, AFIS screenshots, marking minutiae with arrows or other symbols
Making a decision	Making a judgement	Expert specifies interpretations, adequacy of evidence, compromise talk or other case relevant/irrelevant information
	Making a decision	Changed or unchanged decision, reinforcement or abandonment of identification and rules of decision making

References

- [1] M. Lynch, S. Cole, Science and technology studies on trial: dilemmas of expertise, *Soc. Stud. Sci.* 35 (2) (2005) 269–311.
- [2] M.J. Saks, J.J. Koehler, The coming paradigm shift in forensic identification science, *Sci. N. Ser.* 309 (5736) (2005) 892–895.
- [3] National Academy of Sciences (NAS), Executive summary of the national academies of science reports, strengthening forensic science in the United States: a path forward, *Forensic Sci. Policy Manage.* 1 (2009) 106–122.
- [4] B. Latour, S. Woolgar, *Laboratory Life*, Princeton University Press, Princeton, NJ, 1979.
- [5] A. Anthonioz, N. Egli, C. Champod, C. Neumann, R. Puch-Solis, A. Bromage-Griffiths, Level 3 details and their role in fingerprint identification: a survey among practitioners, *J. Forensic Ident.* 58 (5) (2008) 562–589.
- [6] S.A. Cole, *Suspect Identities: A History of Fingerprinting and Criminal Identification*, Harvard University Press, Cambridge, MA, 2001.
- [7] G. Langenburg, A Critical Analysis and Study of the ACE-V Process, Université de Lausanne, Lausanne, 2012 (Thèse de Doctorat en Science Forensique).
- [8] National Institute of Justice (NIJ), The Fingerprint Sourcebook. Resource Document, May, 2014 <https://www.ncjrs.gov/pdffiles1/nij/225320.pdf>.
- [9] B.T. Ulery, R.A. Hicklin, J. Buscaglia, M.A. Roberts, Accuracy and reliability of forensic latent fingerprint decisions? *Proc. Natl. Acad. Sci.* 108 (19) (2011) 7733–7738.
- [10] G. Langenburg, A performance study of the ACE-V process, *J. Forensic Ident.* 59 (2009) 219–257.
- [11] K.A. Ericsson, The influence of experience and deliberate practice on the development of superior expert performance, in: K.A. Ericsson, N. Charinness, P. Felto-vich, R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance*, Cambridge University Press, Cambridge, MA, 2006, pp. 683–704.
- [12] K. Hakkarainen, T. Palonen, S. Paavola, E. Lehtinen, *Communities of Networked Expertise: Professional and Educational Perspectives*, Elsevier, Oxford, 2004.
- [13] J. Lave, E. Wenger, *Situated Learning: Legitimate Peripheral Participation*, Cambridge University Press, Cambridge, MA, 1991.
- [14] M. Polanyi, *The Tacit Dimension*, Peter Smith, Gloucester, MA, 1966 (Reprinted 1983).
- [15] L.S. Vygotsky, *Mind in society*, in: *The Development of Higher Psychological Processes*, Harvard University Press, USA, 1978.
- [16] C. Goodwin, Professional vision, *Am. Anthropol.* 96 (3) (1994) 606–633.
- [17] S.M. Kassir, I.E. Dror, J. Kukucka, The forensic confirmation bias: problems, perspectives, and proposed solutions, *J. Appl. Res. Mem. Cogn.* 2 (2013) 42–52.
- [18] I.E. Dror, C. Champod, G. Langenburg, D. Charlton, H. Hunt, R. Rosenthal, Cognitive issues in fingerprint analysis: inter- and intra-expert consistency and the effect of a 'target' comparison, *Forensic Sci. Int.* 208 (2011) 10L 17.
- [19] I.E. Dror, R. Rosenthal, Meta-analytically quantifying the reliability and biasability of forensic experts, *J. Forensic Sci.* 53 (4) (2008) 900–903, <http://dx.doi.org/10.1111/j.1556-4029.2008.00762.x>.
- [20] R. Gregory, *The Intelligent Eye*, Weidenfeld and Nicolson, London, UK, 1970.
- [21] R.C. Schank, *Dynamic Memory Revisited*, Cambridge University Press, USA, 1999.
- [22] I.E. Dror, D. Charlton, Why experts make errors, *J. Forensic Ident.* 56 (4) (2006) 600–616.
- [23] J.P. Black, Is there a need for 100% verification (review) of latent print examination conclusions? *J. Forensic Ident.* 62 (1) (2012) 80–100.
- [24] J.M. Kruger, D. Dunning, Unskilled and unaware of it: how difficulties in recognizing one's own incompetence lead to inflated self-assessments, *J. Pers. Soc. Psychol.* 77 (1999) 1121–1134.
- [25] I.E. Dror, A. Péron, S.-L. Hind, D. Charlton, When emotions get the better of us: the effect on contextual top-down processing on matching fingerprints, *Appl. Cogn. Psychol.* 19 (6) (2005) 799–809.
- [26] B. Schiffer, C. Champod, The potential (negative) influence of observational biases at the analysis stage of fingerprint individualisation, *Forensic Sci. Int.* 167 (2007) 116–120.
- [27] G. Langenburg, C. Champod, The GYRO system—a recommended approach to more transparent documentation, *J. Forensic Ident.* 61 (4) (2011) 373–384.
- [28] A.V. Maceo, Qualitative assessment of skin deformation: a pilot study? *J. Forensic Ident.* 59 (4) (2009) 390–440.
- [29] D. Ashbaugh, *Quantitative–Qualitative Friction Ridge Analysis: An Introduction to Basic and Advanced Ridgeology*, CRC Press LLC, New York, NY, 1999.
- [30] J. Saldaña, *The Coding Manual for Qualitative Researcher*, SAGE, Los Angeles, CA, 2012.
- [31] K. Krippendorff, *Content Analysis: An Introduction to Its Methodology*, SAGE, Los Angeles, CA, 2013.
- [32] B. Gregg, Using legal Rules in an indeterminate world. Overcoming the limitations of jurisprudence, *Polit. Theory* 27 (3) (1999) 357–378.
- [33] M. Lynch, The externalized retina: selection and mathematization in the visual documentation of objects in the life sciences, *Hum. Stud.* 11 (2/3) (1988) 201–234.
- [34] G. Langenburg, C. Champod, T. Genessay, Informing the judgments of fingerprint analysts using quality metric and statistical assessment tools, *Forensic Sci. Int.* 219 (1–3) (2012) 183–198 (10).
- [35] C. Neumann, C. Champod, M. Yoo, T. Genessay, G. Langenburg, Improving the understanding and the reliability of the concept of sufficiency in friction ridge examination, in: *Resource Document, National Institute of Justice—Office of Justice Program, Fundamental Research to Improve Understanding of the Accuracy, Reliability, and Measurement Validity of Forensic Science Disciplines*, 2013 <https://www.ncjrs.gov/pdffiles1/nij/grants/244231.pdf>.
- [36] B.T. Ulery, R.A. Hicklin, J. Buscaglia, M.A. Roberts, Repeatability and reproducibility of decisions by latent fingerprint examiners, *PLoS ONE* 7 (3) (2012) e32800.
- [37] I.E. Dror, G. Hampikian, Subjectivity and bias in forensic DNA mixture interpretation, *Sci. Justice* 51 (2011) 204–208.
- [38] G. Langenburg, C. Champod, P. Wertheim, Testing for potential contextual bias effects during the verification stage of the ACE-V methodology when conducting fingerprint comparisons, *J. Forensic Sci.* 54 (3) (2009) 571–581.
- [39] R.A. Hicklin, J. Buscaglia, M.A. Roberts, Assessing the clarity of friction ridge impressions, *Forensic Sci. Int.* 226 (1–3) (2013) 106–117 (PubMed PMID: 23313600).