Measuring behaviours during a computer-assisted consensus approach in fingermark examination

Bernd Göttker

Faculty of Behavioural, Management and Social Sciences, University of Twente

201400214: Conflict, Risk and Safety

Peter de Vries & Steven Watson

December 2, 2024

Contents

Abstract

The results of fingermark examination are used as evidence in the justice system and are generally viewed as scientific and robust. However, as human experts are the main instrument of analysis, human error can occur because of subjectivity and flaws in the ACE-V methodology. Therefore, a computer assisted consensus approach is tested. This means participants are instructed to use the software PiAnoS to support their decision-making process and later in establishing a consensus. The PiAnoS software can track which features of a fingerprint have been marked for analysis and comparison, hence supporting the discussion on those features during the consensus approach.

To account for potential pitfalls of the consensus approach, 23 participants consisting of fingermark examiners and students all with different levels of expertise engaged in group discussions on ten marks of differing complexity, with surveys administered before and after. The initial survey served as a baseline for the variables, psychological safety and stress. The subsequent survey captured negative group behaviours, psychological safety, and stress. Psychological safety is assumed to be a prerequisite for a successful discussion in a team. The work environment of fingerprint examiners includes factors that enhance stress. This could be disruptive to effective decision making, therefore stress experienced during the consensus approach is explored.

Results indicate higher psychological safety and lower stress during the computer assisted consensus approach compared to baseline measures. Negative group behaviours occurred at an acceptable frequency. Hence, the computer-assisted consensus approach is recommended for future deployment in fingerprint examination, albeit requiring further testing in the workplace.

1.Introduction

In 1998, Stephan Cowans was wrongfully convicted for the armed assault and attempted murder of a police officer, largely due to false testimony by fingerprint analysts and erroneous eyewitness identifications. Later it was found that the key evidence, a thumbprint on a water mug, was misidentified by Boston Police Department fingerprint analysts Dennis LeBlanc and Rosemary McLaughlin. In 2004, DNA testing excluded Cowans as the source of biological evidence, leading to his acquittal. Initially, Cowans had been sentenced to 35 to 50 years in prison, but the New England Innocence Project's efforts and DNA testing ultimately proved his innocence after 6 years (Warden & Aikins, 2014).

The occurrence of human error by forensic analysts is equally harmful for the stability of the criminal justice system, but much more difficult to detect. In general, forensic analysis techniques are commonly regarded as error prone by courts and the public (Ribeiro et al., 2019). Despite this, expert witness testimonies based on the results of DNA analysis and fingermark analysis are used as incriminating evidence by courts (Edmond, 2022). However, feature comparison analysis disciplines such as fingermark analysis, are frequently questioned in terms of their scientific validity and reliability (Edmond, 2022). Nonetheless, judges and lawyers often lack the knowledge on scientific methodologies and are sometimes oblivious to those concerns in forensic analysis (Leonetti, 2024). To better understand how these issues can arise, it is essential to examine the procedural steps involved in fingermark examination.

1.1 Procedure of fingermark examination

The process of fingermark analysis starts by capturing fingermarks, also called latent prints, found at a crime scene (Ezegbogu & Omede, 2022). A fingermark is defined as “Recovered traces left by unprotected fingers in uncontrolled conditions” (Meuwly, 2014). The origin of the fingermark is usually unknown. While a fingerprint would be the inked imprint copied from the finger’s dermal ridges (Meuwly, 2014). The main objective of fingermark analysis is to find features in both latent print and the exemplar fingerprint that can be utilised for a comparison (Malhotra et al., 2021).

The features of a fingerprint are divided in three categories (Meuwly, 2014). In the first category are the general patterns of a fingerprint that can be an arch, a loop or a whirl. Second category are minutiae which are divided into ridge ending, bifurcation and dot. Third level details are ridge edges, the shape of minutiae and the shape and position of the pores (Meuwly, 2014). A mark is considered complex if the fingerprint is small, unclear, distorted or has few features for comparison (Ulery et al., 2011). It is also possible that marks overlap

with each other, or that they emerge on a surface that is ill-fitted for the collection process thus leading to a distorted mark (Ulery et al., 2011).

A suspect that is linked to a crime, gets his fingerprint captured by the police and then the fingermarks previously found, are compared to the fingerprints of the suspect. This is usually done via a structured approach, referred to as the ACE-V methodology, which is an acronym for Analysis-Comparison-Evaluation-Verification (Ezegbogu & Omede, 2022). However, fingermark analysis is not standardised across the world. Therefore, different law enforcement agencies do have varying techniques for their analysis (Ezegbogu & Omede, 2022). Although varying in their approaches, most law enforcement agencies use the ACE-V approach or a method that resembles the ACE-V approach.

1.2 ACE-V approach

During the first step, fingermark and fingerprint are analysed and core features of both are marked. In the comparison step the previously marked features are compared. This leads to the evaluation phase, where a decision is made. Either the mark is an identification (indicating that the mark and print share the same origin), an exclusion (the mark and print do not share the same origin), or inconclusive (indicating insufficient features to reach a decision). The verification step is added in some cases, where another examiner follows the steps of ACE a second time without knowing the outcome of the previous analysis (Ezegbogu & Omede, 2022).

But the ACE-V approach is inherently subjective, meaning that the expert's personal judgement plays a significant role during the analysis (Sikorski, 2022). Fingerprint examiners must determine whether specific differences between two prints is enough to argue for an exclusion or whether similarities between two prints justifies an identification. Nonetheless, a consistent methodology to arrive at such conclusions is missing, for example how many similarities are needed to decide for a match, or which minutiae are documented for the comparison of prints (Ulery et al., 2016).

1.3 Contextual information

At first sight, the ACE-V methodology appears to be an objective method for the analysis of fingermarks. However, issues are introduced when the human brain is the main instrument of analysis (Sikorski, 2022). Dror and Charlton (2006) found that fingerprint examiners can be influenced by extraneous contextual information to reach a different conclusion on a mark they already examined in the past. There are various forms of

extraneous contextual information. For example, if the suspect has already confessed to the crime, or if the results of a preceding analysis like DNA is known to the examiner (Stoel et al., 2015). When the results of one line of evidence are known, it has the potential to bias the following forensic analyses and render them as erroneous. This effect is termed "Biasing Snowball Effect" (Dror & Stoel, 2014). The contextual information used in the study was either "suspect confessed to the crime" when the print was an exclusion or "the suspect was in police custody at the time of the crime" when the print was in fact an identification (Dror & Charlton, 2006). The fingerprint examiners were influenced by the previously received information and therefore made the incorrect decision. This is interesting since they made the correct decision in the past when analysing the same prints without extraneous information present. However, this effect was more pronounced when analysing complex fingerprints (Dror & Charlton, 2006).

1.4 Subjectivity, bias and variability

The study by Dror and Charlton (2006) exemplified that contextual information can bias the judgement of fingerprint examiners. In addition, Cooper and Meterko (2019) found in their meta-analysis that fingerprint examiners are susceptible to confirmation bias. Confirmation bias is defined as "seeking or interpreting of evidence in ways that are partial to existing beliefs, expectations, or a hypothesis in hand" (Nickerson, 1998). In the study by Dror and Charlton (2006), examiners expected that the suspect was either innocent or guilty depending on their randomly assigned group, thus distorting their judgements towards the decision they were primed to adopt. Fingerprint examination is based on human interpretation rather than objective measures. Most of the time fingerprint examiners do replicate each other's decision. Nonetheless, in some cases two examiners can reach a different conclusion when analysing the same fingerprints (Hicklin et al., 2020). Hence, variability in judgements between examiners does exist.

In addition, decision making in fingerprint examination happens with uncertainty (Georgiou et al., 2020). This uncertainty stems from the nature of forensic investigations. Inevitably, forensic science is applied to reconstruct the circumstances of a crime that happened in the past (Taroni & Biedermann, 2015). But a ground truth in fingerprint examination is non-existent, meaning that the correct decision cannot be known (Haber & Haber, 2014). Since some degree of subjectivity of examiners will be given, the risk is increased that incorrectly interpreted finger marks could amplify miscarriages of justice.

1.5 Expert decision making and top-down processing

Subjectivity is not the only precursor for the occurrence of confirmation bias and subsequently human error in fingermark examination. The way how experts process information also has the potential to result in confirmation bias (Zapf & Dror, 2017). Incoming information is either processed bottom-up or top-down (Stoel et al., 2015). Bottom-up processing happens when no preconceived information about the stimulus is present before analysing said stimulus. When information about a stimulus is already known by the examiner, it is processed top-down, meaning that preconceived knowledge about a stimulus influences how it is perceived. Hereby, large amounts of information can be processed quickly and automatically (Zapf & Dror, 2017). Top-down cognitive processing is fundamental to human intelligence and is related to having reached expertise in a domain (Stoel et al., 2015). Expertise in this study is defined as years of experience in the field. However, top-down processing can be counterproductive as it can lead to issues like overlooking important information, tunnel vision and cognitive bias (Stoel et al., 2015). These mental shortcuts used while processing information can be misleading.

1.6 Workplace stress

Besides the human factors related to decision making in fingerprint examination, other factors need to be taken into consideration. First, the work environment of forensic analysts is subjected to factors that increase the occurrence of perceived stress (Ditrich, 2015). Those factors often arise from the organisational structure of forensic laboratories and include time pressure, expectations to reach results and budget controls to name a few (Dror, 2020).

In 2020, Phillips-Wren and Adya identified four stressors that can impair a decision-making process. Namely, the four factors are information overload, time pressure, complexity and uncertainty with perceived time pressure. The detrimental impacts of stress on decision quality arise when cognitive resources are redirected toward stress management (Phillips-Wren & Adya, 2020). In a study conducted by Pabst et al. (2013) found that a peak of the stress hormone cortisol was related to impaired decision-making. Therefore, erroneous judgements of fingerprint examiners could be partly caused by the perceived stress they are exposed to. Especially when taking into consideration that three out of four stressors can easily arise in fingerprint examination. Those are time pressure, uncertainty and in some cases complex marks (Dror, 2020; Georgiou et al., 2020; Ulery et al., 2011). Regulating perceived stress in high stakes decision making teams can be influenced positively through team psychological safety (Hebles et al., 2022).

1.7 Psychological safety and teamwork

In 1999, Edmondson posited that psychological safety is better understood as a team-level variable. She defined it as the "shared belief held by members of a team that the team is safe for interpersonal risk taking" (Edmondson, 1999). Moreover, the concept of psychological safety refers to an individual's perception of feeling at ease to express and utilise their authentic self without the apprehension of facing adverse consequences to their self-image, status, or career (Newman et al., 2017). If team psychological safety is low, different levels of expertise negatively affects team performance (Martins et al., 2012). On the other hand, if team psychological safety is high, "increased knowledge sharing, engagement, creativity, innovation, and ultimately performance," is present (Newman et al., 2017). Psychological safety contributes to alleviating communication difficulties by fostering an environment where individuals feel free to openly express their thoughts and emotions. In addition, when team members have the freedom to express their disagreement openly, there is a greater likelihood of effective cooperation and collaboration (Cottrel, 2023).

1.8 Consensus as solution

Consensus is defined as a mutual agreement between group members about a course of action or decision (Palomares et al., 2014; Moore & O'Doherty, 2013). When comparing decision rules with each other, consensus produced decisions are higher in quality than a majority vote rule (Miller, 1989). The task of examining a mark to a print in a group and reaching consensus, resembles the structure of group decision-making. In group decision-making, two or more experts must find a solution to a problem from a set of alternatives (Palomares et al., 2014). In the context of the study, the experts are fingerprint examiners and have to decide if during the analysis stage a mark is of value for identification, exclusion or inconclusive. During the comparison and evaluation stage if the mark and print lead to an identification, exclusion, or an inconclusive result. During the process of reaching consensus, experts are able to adjust their opinions, so they are closer to the collective opinion of the group (Palomares et al., 2014).

Moreover, two categories of behaviours are identified that could hinder the occurrence of consensus. First, overconfident behaviours of decision-makers pertain to exhibiting excessive confidence in their opinions during the decision-making process, resulting in a reluctance to compromise or accept divergent viewpoints (Dong et al., 2021). This overconfident behaviour often arises from a decision maker´s overly optimistic

assessments of their own abilities, knowledge, or judgments, sometimes overlooking the existence of uncertainty and insufficient information (Dong et al., 2021). Second, manipulative behaviour in decision-makers often originates from a motivation to gain personal or specific benefits, or to steer decision outcomes to align with one's individual preferences (Dong et al., 2021). In conclusion, these behaviours can lead group members to be easily influenced and might lead to an erroneous decision.

1.9 Groupthink and conformity

Groups that actively pursue agreement in their decision-making processes may prioritise group harmony to an extent that hinders constructive criticism (Kellermanns et al., 2011). Another way to grasp this issue is through the lens of conformity. Research about conformity has a long history with the research conducted by Asch in 1951 being the most famous example. His experiment deals with how an individual reacts to group pressure in a simple decision-making task. In his experimental setup, one participant was put in a group of mock participants and each participant had to compare lines with each other and publicly tell their opinion. The mock participants were instructed to give a false answer to see if this influenced the opinion of the participant (Asch, 1951). The participants who yielded to group pressure were in the minority, however the influence a majority has over the individual has been proven to exist. In response to these results, the majority effect was derived, which occurs when the beliefs held by the majority in a group prevail over the beliefs of the minority (Asch, 1951). This effect was most pronounced when a majority of three was present but also a majority of two people could lead to distorted judgement (Asch, 1951). When individuals adopt the majority's stance despite privately disagreeing, they conform, which may hinder the establishment of a valid consensus.

The successful implementation of a consensus approach among fingerprint examiners can be distorted by several factors. First, the issue of groupthink must be discussed. Groupthink can be explained as a process that discourages the questioning of the prevailing perspective. Individuals who dissent from the dominant perspective are pressured by group members to establish a consensus (Størseth et al., 2014). As soon as groupthink becomes evident peer pressure, pressure from authorities and pressure to reach consensus are the main behaviours that lead to poor and uninformed decision-making (Solomon, 2006). Since groupthink may lead to poor decision-making, an encouragement to show dissent might prove to be valuable. Dissent here is useful as opinions will be improved through criticism and a

response to that criticism (Solomon, 2006). One factor that could facilitate dissent would be psychological safety.

1.10 The current research

In summary, the field of forensic fingerprint examination is prone to erroneous decision making, which is an indirect result of human subjectivity and in turn confirmation bias. However, unreliable methods like ACE-V and the introduction of extraneous contextual information to an examiner hinder objective decision-making (Sikorski, 2022; Dror & Charlton, 2006). To mitigate these issues, the introduction of a computer-assisted consensus approach among fingerprint examiners is tested. It is assumed that establishing a consensus has a positive impact on decision-making.

The participants were instructed to individually examine ten pairs of fingerprints and marks, while keeping track of their decision in the PiAnoS software. Afterwards, participants will fill out a questionnaire. This phase is called the pre consensus phase. In the second part of the experiment, participants are taking part in a workshop and have to talk about their individual findings from the pre consensus phase. The goal is to establish a consensus on the ten fingerprints pairs. After that participants are instructed to fill out a second questionnaire. The second phase of the experiment is called the post consensus phase. However, establishing a consensus may involve pitfalls like conformity and groupthink, manipulation, and overconfidence and seniority. Especially in an environment where different levels of expertise are present, this issue can be intensified (Martins et al., 2012).

In addition, the perceived stress fingerprint examiners are exposed to could also hinder effective decision-making (Phillips-Wren & Adya, 2020). Lastly, the concept of psychological safety is introduced, because it is assumed to be a precursor for a successful implementation of consensus and may mitigate issues such as group think. This leads us to the research question: “What impact does the implementation of a consensus approach in fingerprint examination have on psychological safety and stress?”.

Based on the available research the following hypotheses have been constructed:

H1: The average psychological safety levels are higher in the post-consensus phase compared to the psychological safety levels in the pre-consensus phase.

H2: Expertise has a negative moderating effect on psychological safety levels in the post consensus condition.

H3: Workplace stress levels are higher compared to the stress levels in the post-consensus phase.

This research is intended to be explorative. The idea is to assess tools like PiAnoS software, the consensus roles or the guided checklist in terms of their usefulness for practitioners in the field. In addition, it will also be explored if and how consensus is already employed in fingerprint examination. Lastly, the frequency of behaviours regarded as disruptive for consensus will be assessed.

2. Methods

2.1 Design

This study employed a pre-test/post-test quasi-experimental design. The independent variable (IV) was the **consensus approach** with two levels: **consensus applied (yes)** and **consensus not applied (no)**. The dependent variables (DVs) were **stress** and **psychological safety**, both of which were measured before and after the consensus approach. Additionally, **expertise** was hypothesised to moderate the relationship between the consensus approach and psychological safety.

2.2 Participants

All participants were either employed as friction ridge examiners or were students of forensic science. A non-probability sampling method was used to acquire participants. In total, 23 participants took part in the experiment. However, 12 participants were excluded from the repeated measures analysis because it was not possible to identify them across pre and post conditions, leaving a sample of 11 participants for the primary analyses. They were not identifiable because they did not input their identification numbers correctly into the questionnaires.

Nevertheless, demographic data were collected from 17 participants since five participants did not fill out the demographic questions. Therefore, their information is missing. Fifteen participants were from the Netherlands and two were from Scotland. The Scottish examiners only filled in the post-consensus questionnaire, because they already completed the consensus approach prior to the experiment as part of a pilot study. Most participants were under the age of 35, with 26.7% falling into the under 25 category and 26.7% falling into the 25-34 age group. Participants aged 35 and above represented the smaller portion of the sample. Most participants identified as female, comprising 73.3% (12) of the total sample, while male participants represented a smaller proportion with 26.67% (5). In terms of educational background, the majority had a bachelor’s degree (7), followed by

high school diploma (5). The participants had varying degrees of expertise in the field of friction ridge analysis.

2.3 Procedure

Pre consensus

This study took place during a workshop with fingermark examiners that was specifically designed for this research. Before the workshop, the latent print examiners were asked to perform ACE on ten pairs of fingermarks and fingerprints of varying complexity without being informed about the correct decision. The verification stage of ACE-V was not applied, because the examiners had to work individually. During the analysis stage of ACE, they decided on the value for identification, value for exclusion only and no value for the specific areas of the prints. In the comparison and evaluation stages, examiners decided if it was an exclusion, identification, or inconclusive result. If they decided for inconclusive, examiners were asked to give a direction for their decision, either towards identification or towards exclusion. While analysing the mark, every decision was documented using the PiAnoS software that will be explained later in more detail. After the examiners had completed their individual examinations, they were asked to complete the pre-consensus questionnaire measuring stress levels and the psychological safety scores in the workplace.

Workshop

When all participants completed this phase, a two-day focused workshop was conducted. The workshop consists of two phases, first an educational phase and later the consensus phase. In the educational phase, five forums were presented as a PowerPoint presentation. First, forensic judgments dealt with the processes active while applying the ACE approach. The second forum was about human factors in forensic science. Participants learned about subjectivity in human judgement when applying the ACE-V approach and how this subjectivity is a limitation of friction ridge analysis. The third forum was about debiasing strategies. In the fourth forum, participants learned about the characteristics of consensus, how it is applied in the forensic context, and what kinds of behaviours can negatively or positively affect consensus. Lastly, in the fifth forum it was explained how to use the PiAnoS software but also limitations of the software were pointed out.

Consensus tools

After the forums were completed, the participants met in groups of three that were constructed beforehand. At their workstation, they were given the consensus cards and the

guided checklist. Four different consensus cards could be used by the participants to facilitate consensus in a non-verbal manner in the case that verbal communication comes to a halt. The cards had different colours and shapes. First, the green circle could be used to signal to the group members that one wants to say something during the discussion and could be used to show agreement during the call for consensus. Next, the yellow square could be used to indicate a direct response to a previous statement during the discussion. During the call for consensus, this card could be used to indicate a reservation for the judgement made by the group, but it does not mean that one opposes the decision. The orange arrow was able to be used when a group member wants to show that he does not oppose the decision or opinion, but he also does not support the decision either. This card has no usage during the call for consensus. Lastly, the red triangle could be used when the discussion goes off-topic, the time limit was reached, or when other rule breaches occurred. During the call for consensus, the red card was used to show opposition to the judgement made by their group, but it still advocates that consensus had been reached. When used correctly, the consensus cards should facilitate group discussion.

The PiAnoS software was developed by the University of Lausanne (Appendix D). PiAnoS is an acronym for "Picture Annotation system," and it was developed for fingerprint analysts to document their reasoning on a given mark. During the analysis, the examiner could decide between three different outcomes for a judgement on a fingerprint mark: identification, exclusion or inconclusive, for prints that did not have enough data to proceed with the analysis. Three functions of PiAnoS were used by the participants to inform their decision-making. First, ESLR which stands for expected score-based likelihood ratio and is used to search for potential identifications of fingerprints (Stoney et al., 2020).

The guided checklist was provided at each workstation to ensure that participants followed a clear method when performing the consensus approach and using PiAnoS. When correctly followed, the list guided the participants in the usage of PiAnoS and its resources. The participants were instructed to tick a box when they completed a step in the analysis or indicate why they have not done so. This list closely resembles the ACE-method and is divided into analysis, comparison and evaluation (see Appendix C).

Post consensus

While the participants were discussing their decisions on latent prints, three different sets of roles were assigned to them. The role of the timekeeper was to make sure that the group discussion moves along the given timeframe of one hour per comparison. The mission of the facilitator was to make sure that the group sticks to the task and encouraged the group to use the consensus cards. Another task of this role was to make sure that all items on the consensus checklist have been met and ticked off. Lastly, the role of the driver was concerned with annotating and recording the judgments of the group. Additionally, the driver also has to record the consensus results and ensure the ESLR and SLR functions are selected and sent. The participants discussed the ten pairs of fingermarks they previously analysed alone, within their respective groups. Each individual pair was discussed within a timeframe of one hour. After each pair was discussed and consensus had either been reached or not, the participants were given the post-consensus questionnaire. When the post-questionnaire measuring stress levels and the psychological safety scores during the consensus approach was completed by each participant, the experiment was finished.

2.4 Materials

2.4.1 Demographic questionnaire. In the beginning of the questionnaire the demographics of the participants were assessed. These included age, gender, educational background, and expertise. The demographic questionnaire had twelve items. Here expertise was assessed by the question “Level of Expertise” where participants choose between Novice, Beginner, 1-star, 2-star and 3-star. The Scottish fingerprint examiners do not use the star system so for them the options were “Novice”, “Beginner” and “Expert”. All scottish examiners who choose “Expert” were regarded as “3-star” and the categories “Novice” and “Beginner” were regarded as equal. To get a detailed assessment of expertise, participants answered seven follow up questions about their work routine. For example, “Experience with Automated Fingerprint Identification System (AFIS)?” or “Type of Training received”

2.4.2 Complex marks and consensus at work. This part of the survey contained questions relating to the work routine and work environment of friction ridge examiners. Participants were asked to indicate if and how they resort to consensus as a form of decision-making in their daily work routine. In addition, the frequency of analysing complex marks was of interest here. All items were measured on a Likert scale ranging from 1-5, 1=Strongly disagree, unless indicated otherwise. Example items are “How often do you engage in formal

discussions or meetings about a mark with your peers?" and "How often do you think you come across a complex mark during your work?"

2.4.3 Perceived Occupational Stress Scale. The perceived occupational stress scale is an instrument that measures the perceived stress level in the context of the workplace (Marcatto et al., 2022). The variable stress was constructed by calculating the mean score of all items from the occupational stress questionnaire. The higher the mean score of one participant was, the higher their stress levels were. This scale consists of seven items which were scored using a 5-point Likert scale. In its original version, this scale has four items, but three items were added that were considered useful in the context of this study. One example item is "Thinking about my work makes me feel tense." This questionnaire was distributed at two points. First, in the pre-consensus questionnaire where it was used as a baseline measure and during the post-consensus questionnaire. A few items in the post-consensus questionnaire were rephrased so they make sense in the context of the consensus approach. Lastly, this scale had high internal consistency in the pre-consensus phase ($\alpha = .88$) as well as in the post-consensus phase ($\alpha = .89$).

2.4.4 Psychological safety questionnaire. The psychological safety questionnaire was used to measure "team psychological safety". This questionnaire is a modified version from the Edmondson (1999) questionnaire measuring team psychological safety. The eleven items were measured on a Likert-scale ranging from 1-5, 1=strongly disagree, 5=strongly agree. The variable psychological safety was constructed by calculating the mean score of all items from the psychological safety questionnaire. The higher the mean score, the higher the psychological safety level of participants. Items from this questionnaire include "If I admit to an error or mistake, I will not face retaliation or criticism in my workplace." or "My peers welcome my ideas and give them time and attention". This questionnaire was distributed at two points. First, in the pre-consensus questionnaire where it was used as a baseline measure and second during the post-consensus questionnaire. A few items in the post-consensus questionnaire were rephrased so they make sense in the context of the consensus approach. For example, changes included "My peers welcomed my ideas and gave them time and attention" or "If I admitted to an error or mistake, I did not face retaliation or criticism from my peers ". For the psychological safety questionnaire in the pre-consensus phase, a Cronbach alpha of .81 was reported. In the post-consensus phase, a Cronbach alpha of .78 was found.

2.4.5 Consensus tools questionnaire. This questionnaire was used to capture the feelings of participants about the tools they use during the consensus approach. In addition, their usefulness was assessed. The tools that were assessed are the consensus cards, the guided checklist, PiAnoS and consensus roles. Most of the nine items were measured on a Likert-scale ranging from 1-5, but some questions allow for multiple answers. Example items are “I would recommend the PiAnoS tool to my supervisor for facilitating a consensus” and “If available, for which of the following purposes would you consider using the PiAnoS tool in the future? (Select all that apply) “.

2.4.6 Negative consensus behaviours questionnaire. The negative consensus behaviours questionnaire was used to capture the occurrence and frequency of behaviours that could obstruct consensus from being achieved. This questionnaire was constructed by the researcher by taking into consideration antecedents of groupthink and behaviours that could arise in a strict hierarchy and a difference of expertise. It consisted of six items that were measured on a Likert-scale ranging from 1=Never and 5 =Always. Example items are “How often did you change your opinion on your [features annotated/suitability decision/my features paired/my source level opinion] because of pressure from group members?” and “How often did the most experienced person in the group lead the discussion?”

2.4.7 Pre consensus questionnaire. The pre-consensus questionnaire was distributed to the participants before they engaged in the consensus phase. This questionnaire consisted of the demographic questionnaire, the consensus at work questionnaire, the perceived occupational stress scale and the psychological safety questionnaire that were described above (see Appendix A)

2.4.8 Post consensus questionnaire. The post-consensus questionnaire was distributed to the participants after they engaged in the consensus phase. This questionnaire consisted of the consensus at work questionnaire, the negative consensus behaviours questionnaire, the consensus tools questionnaire, the perceived occupational stress scale and the psychological safety questionnaire that were described above (see Appendix B)

2.5 Data analysis

The data is analysed using SPSS 26. For testing the hypotheses, a repeated measures anova is used. One with psychological safety as the dependent variable, expertise as moderator variable and the two measurement points of psychological safety in the pre and

post conditions as the independent variable. The other repeated measures anova is done with stress as the dependent variable and the two measurement points of stress in the pre and post conditions as the independent variable. The variable expertise is constructed by splitting the six options the participants had into two categories. The first category was expertise low and consists of “Novice”, “Beginner” and “1-star” which is contrast coded as -1. The second category was expertise high and consists of “2-star”, “3-Star” and “Expert” which is contrast coded as 1. Contrast coding is applied to perform a hierarchical regression analysis. In addition, the interaction term for this analysis is computed by multiplying expertise (1, -1) with consensus (1,0) in this case consensus (1) means that the consensus approach was applied, and consensus (0) means that the consensus approach was not applied yet.

3. Results

3.1 Descriptive statistics

Table 1 shows the demographic characteristics of all participants. It also represents characteristics which are important to describe the participants in the context of their work field

Table 1

Demographics

Variable	Attribute	N
Age in years	Under 25	4
	25-34	4
	35-44	3
	45-54	4
	55-64	2
Gender	Male	5
	Female	12
Level of education	High school diploma	5
	Some college, no degree	1
	Associate degree	3
	Bachelor's degree	7
	Master's degree	1
Working as fingerprint analysts	Less than a year	2
	1-5 years	4
	6-10 years	4
	16-20 years	2
	More than 20 years	2
Level of expertise	Novice	4
	Beginner	1
	2-star	4
	3-star	8
Primary work setting	Law enforcement agency	8
	Academic institution	3

Government (non-law enforcement)	1
Other	4

3.2 Correlation table

Table 2 shows the correlations between the variables psychological safety, stress and expertise.

Correlation Table

Variable	*M*	*SD*	1	2	3	4	5	6
1.PsysafetyPost	3.92	.26		.16	.23	-.1	.22	-.22
2. StressPost	2.23	.64			-.72*	.5	.4	-.4
3.PsysafetyPre	3.47	.39				-.8**	-.56	.56
4. StressPre	2.7	.35					.4	.4
5.ExpertiseLow								-1
6.ExpertiseHigh								

*Correlation is significant at the 0.05 level (2-tailed).

**Correlation is significant at the 0.01 level (2-tailed).

3.2 Psychological safety

A repeated measures ANOVA was conducted to examine the effects of consensus and expertise on psychological safety. The analysis revealed a significant main effect of consensus (F (1, 9) = 9.05, p = .01), indicating that participants during the consensus condition (M = 3.92, SE = 0.11) reported higher psychological safety than during the non-consensus condition (M = 3.47, SE = 0.15).

Additionally, there was a significant interaction effect between consensus and expertise (F (1, 9) = 7.36, p = .024). In the low expertise group, participants reported higher psychological safety in the consensus condition (M = 4.03, SE = 0.12) compared to the non-consensus condition (M = 3.47, SE = 0.15). The high expertise group showed no significant difference between the two conditions (M = 3.89, SE = 0.14 for consensus vs. M = 3.92, SE = 0.11 for non-consensus). A significant difference between the baseline levels of psychological safety and psychological safety in the consensus approach has been found. In detail, the scores of psychological safety were higher in the consensus approach than compared to the

psychological safety scores in the workplace. Therefore H1: “The average psychological safety levels are higher in the post-consensus phase compared to the psychological safety levels in the pre-consensus phase” can be accepted.

In the low expertise group, psychological safety levels increased significantly in the consensus condition when compared to the non-consensus condition. In the high expertise group, expertise had no significant effect on psychological safety. Thus H2: “Expertise has a negative moderating effect on psychological safety levels in the post consensus condition” can be rejected.

3.3 Stress

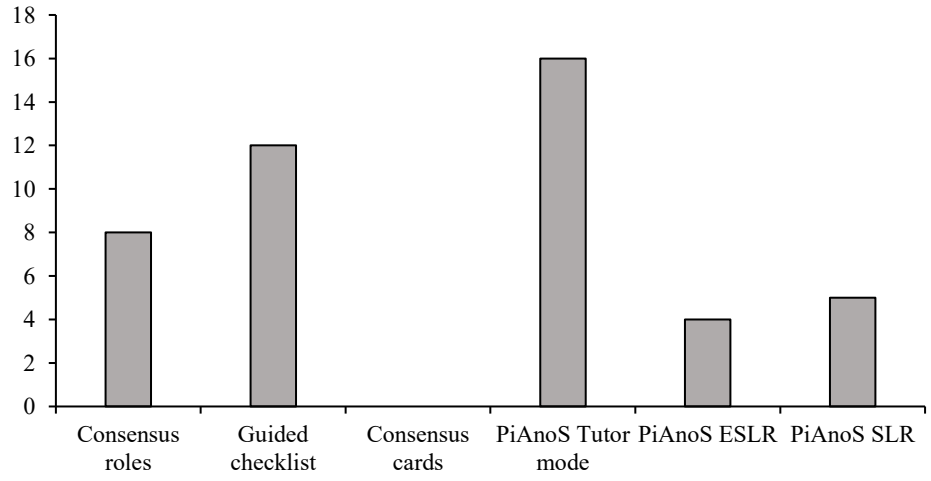
A repeated measures ANOVA was conducted to examine the effect of the consensus phase on stress levels. The results indicated a significant main effect of stress, F (1, 10) = 7.18, $p < .01$. Hence a decrease in average stress levels during the consensus approach is observed. Therefore H3: “Workplace stress levels are higher compared to the stress levels in the post-consensus phase” can be accepted.

3.4 Consensus tools

This figure shows which of the tools used in the consensus approach were regarded as helpful in guiding the decision-making process. The “PiAnoS Tutor mode” was most often described as helpful (n=16). After this the “Guided checklist” was considered as helpful (n=12). Lastly, the three “Consensus roles” were regarded as helpful (n=8).

Figure 1

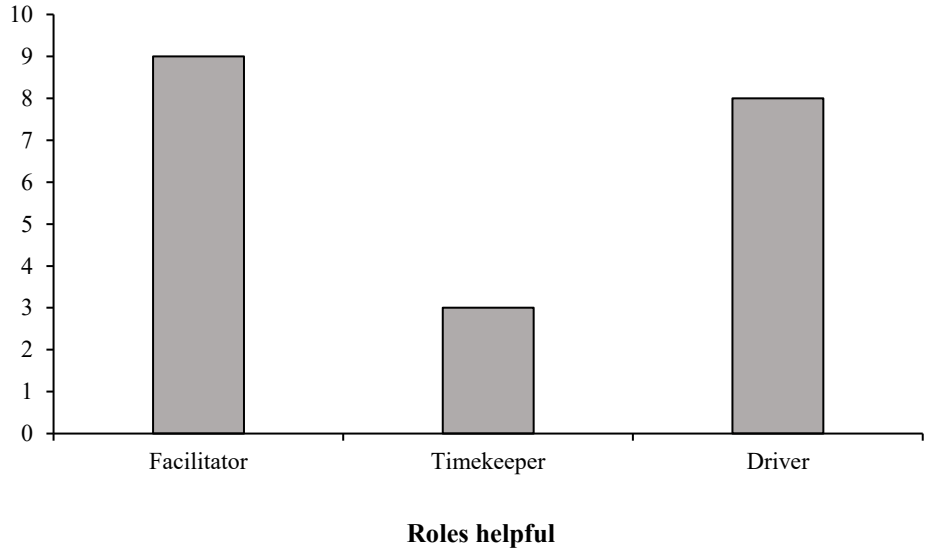
Consensus tools helpful



This figure shows the specific roles during the consensus approach and if they were regarded as useful. The role “facilitator” (n=9) was most often listed as helpful, “Driver” closely follows (n=8). The least favoured role was “Timekeeper” (n=3).

Figure 2

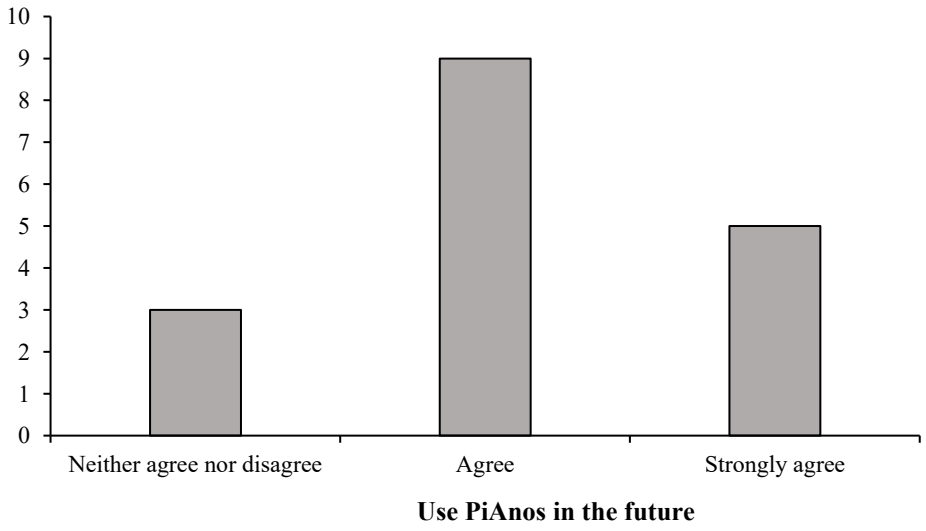
Which role was helpful during the consensus approach?



This figure shows if the PiAnoS tool would find usage in the future. Here, most participants “Agree” (n=9), followed by “Strongly agree” (n=5) to use the PiAnoS tool in the future. The option “Neither agree nor disagree” is present three times.

Figure 3

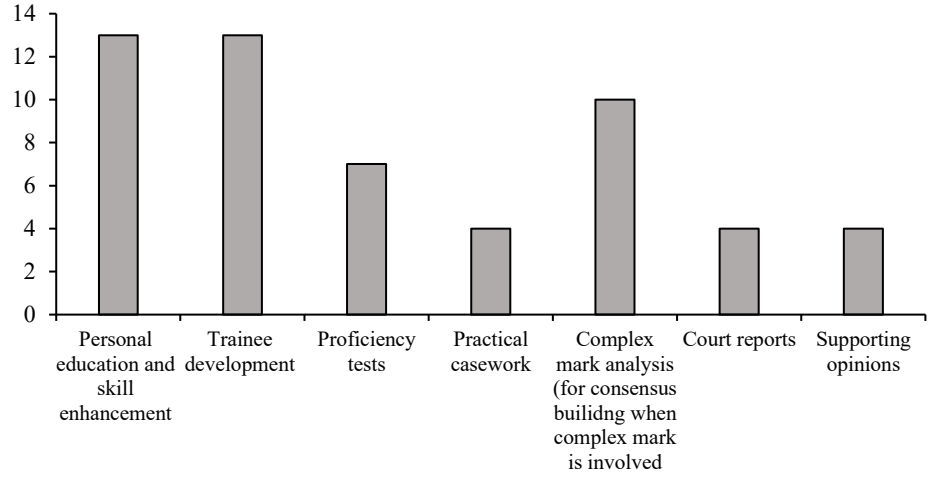
Use PiAnoS as a tool in the future



This figure shows the possible purposes that the PiAnoS software could be used for in the future. The two options “Personal education and skill enhancement “ and “Trainee development” are most often reported (n=13). After this, the option “Complex mark analysis” was most often indicated (n=10).

Figure 4

Possible purposes of PiAnoS

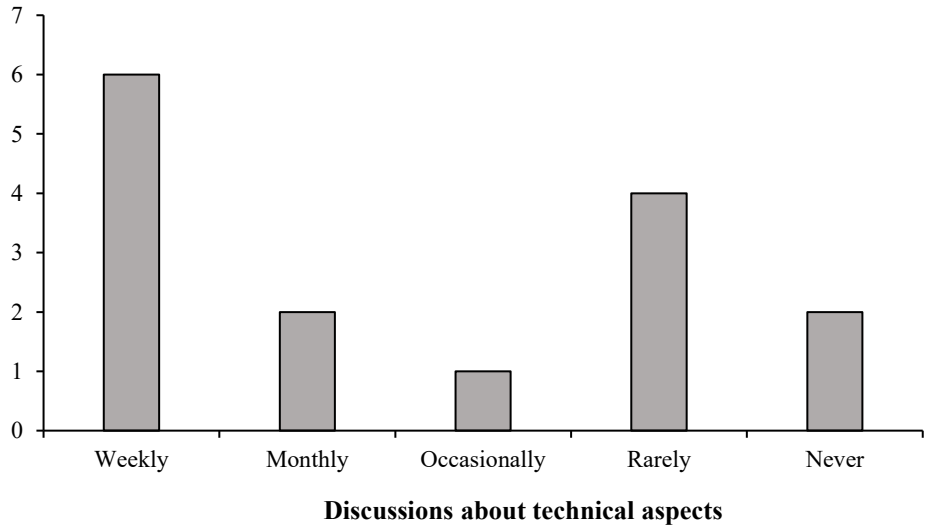


3.4 Consensus at work

This figure shows how frequently the participants engage in discussions about the technical aspects of fingerprint analysis. The most frequent answer is “Weekly” (n=6), followed by “Rarely” (n=4). The options “Monthly” and “Never” were both reported two times.

Figure 5

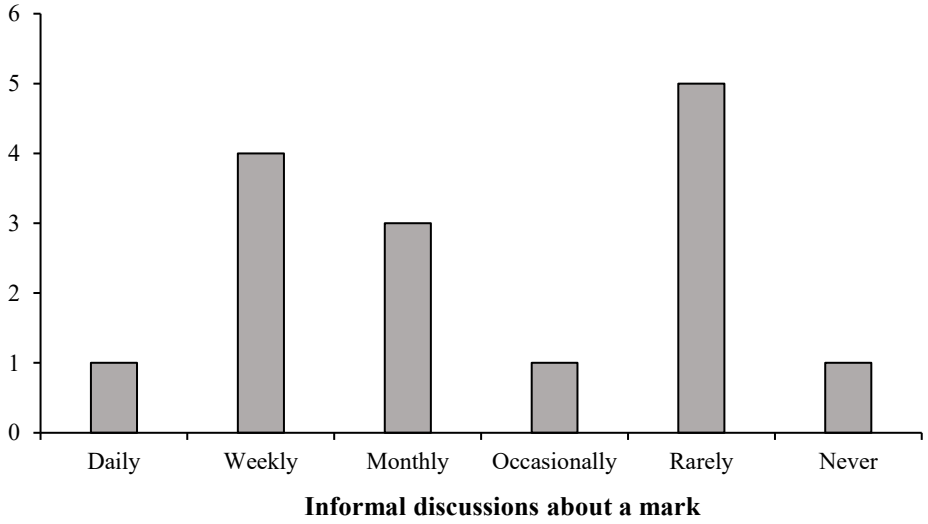
Discussion about technical aspects of fingerprint analysis



This figure shows how often participants engage in informal discussion about a mark with their peers. Here the majority engage “Rarely” (n=5) in informal discussions, followed by “Weekly” (n=4) and “Monthly” (n=3). The options “Never”, “Occasionally” and “Daily” were each indicated by one participant.

Figure 6

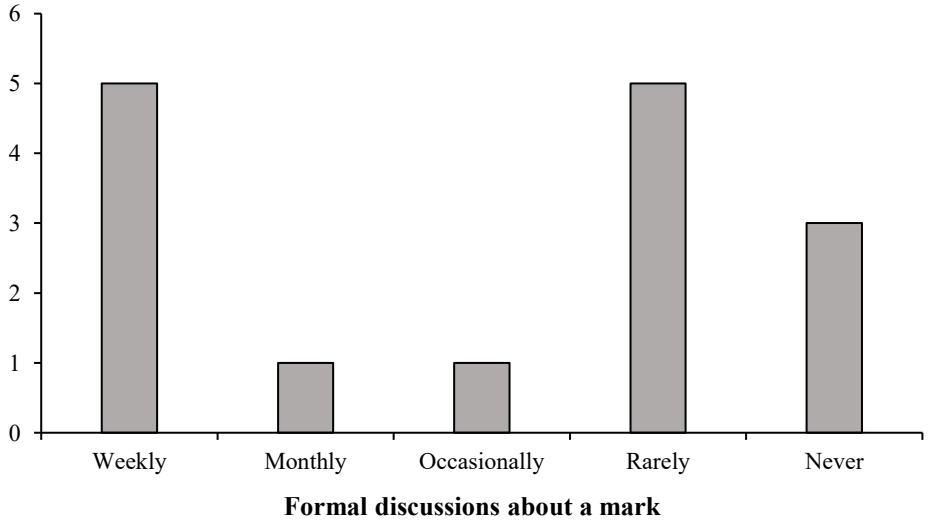
Informal discussions about a mark



This figure shows how frequently participants engage in formal discussions about a mark with their peers during their work. Both options “Weekly” and “Rarely” are present five times. This is followed by the option “Never” with three participants giving that answer. Lastly, “Monthly” and “Occasionally” were both indicated one time.

Figure 7

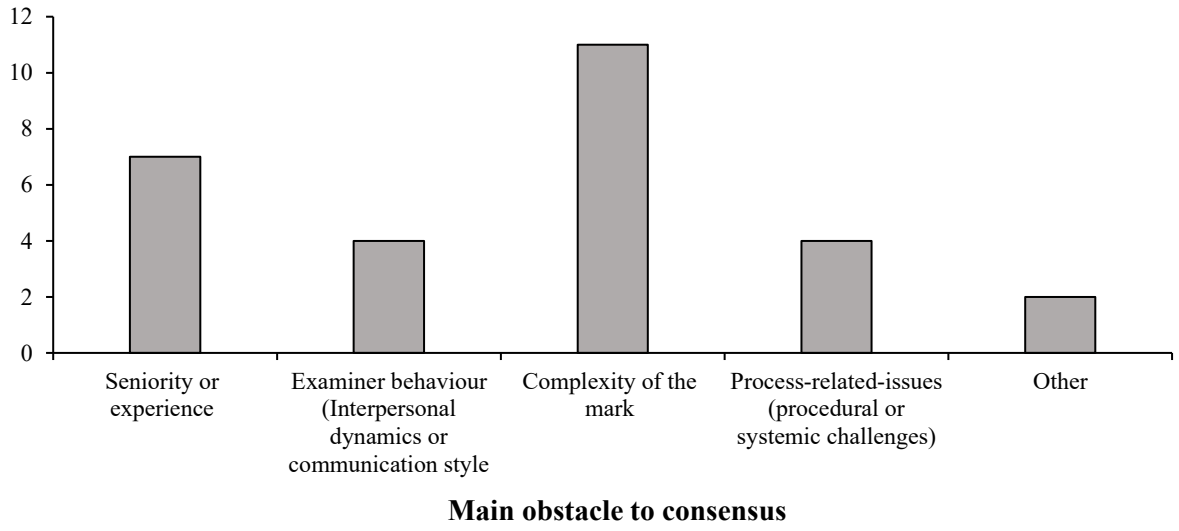
Formal discussions about a mark



This figure shows what the participants regard as the main obstacles achieving consensus. The biggest challenge to consensus is the "Complexity of the mark" (n=11). Next, the "Seniority or experience" of other examiners (n=7) has been identified as an obstacle to consensus.

Figure 8

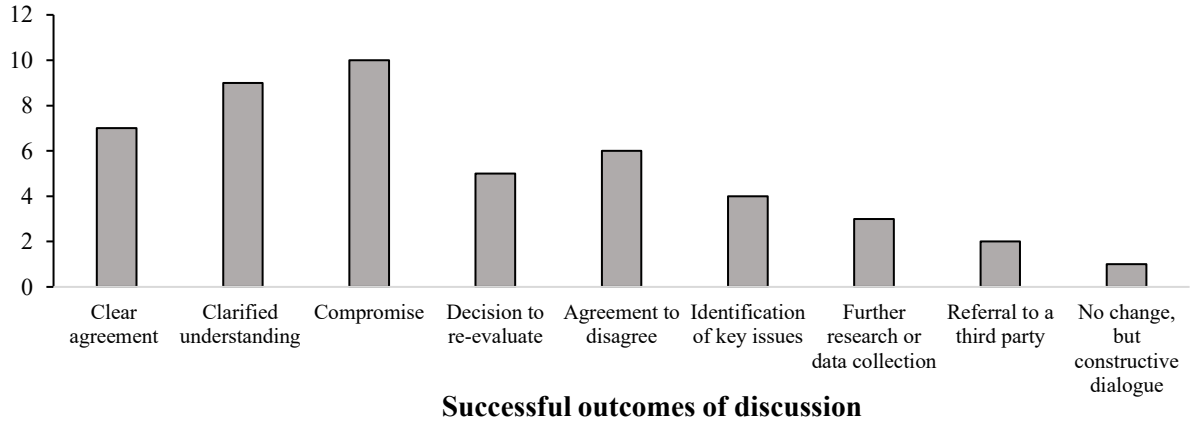
Main obstacle to consensus



This figure shows what the participants regard as successful outcomes when discussing a mark with their peers. The three highest ranking answers are “Compromise” (n=10), “Clarified understanding” (n=9) and “Clear agreement” (n=7) between examiners. After these ranks “Agreement to disagree” (n=6), “Decision to re-evaluate” (n=5) and the “Identification of key issues” (n=4). The least favored option was “No change, but constructive dialogue” (n=1).

Figure 9

Successful outcomes of discussion

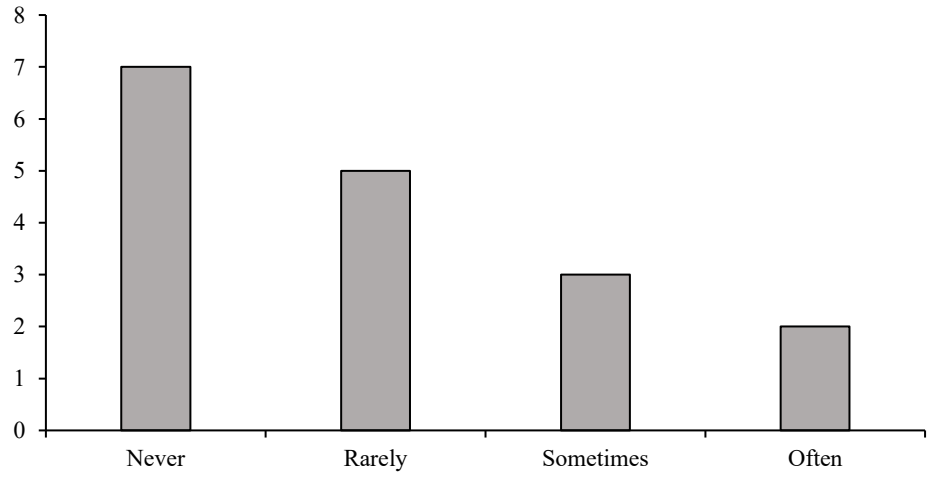


3.5 Negative consensus behaviours

This figure shows how often the participants changed their opinion due to pressure from the group during the consensus approach. The option “Never” is most prominent with seven answers. This is followed by “Rarely” with five answers.

Figure 10

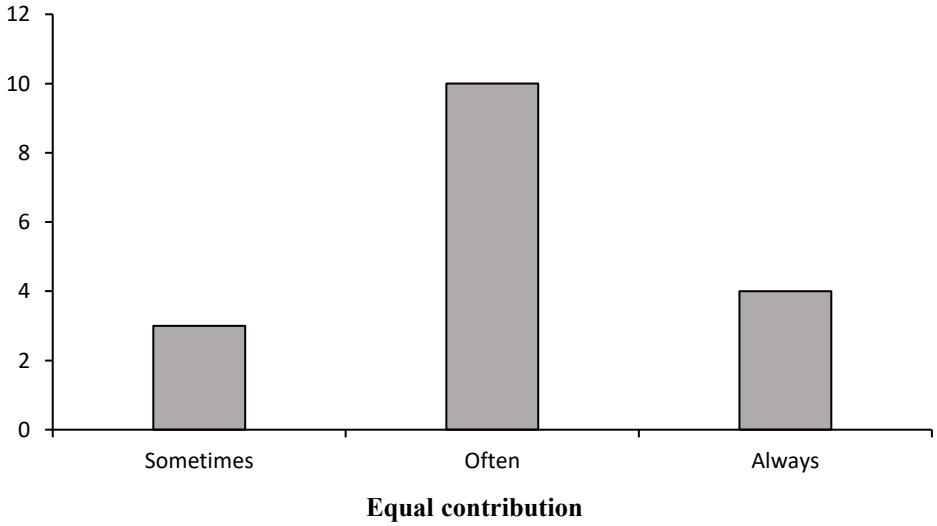
Changing opinion due to pressure



This figure shows if everyone contributed equally during the consensus approach. In total ten participants reported that equal contribution happened often, four participants reported that equal contribution was always the case. Lastly, three people answered that equal contribution happened sometimes.

Figure 11

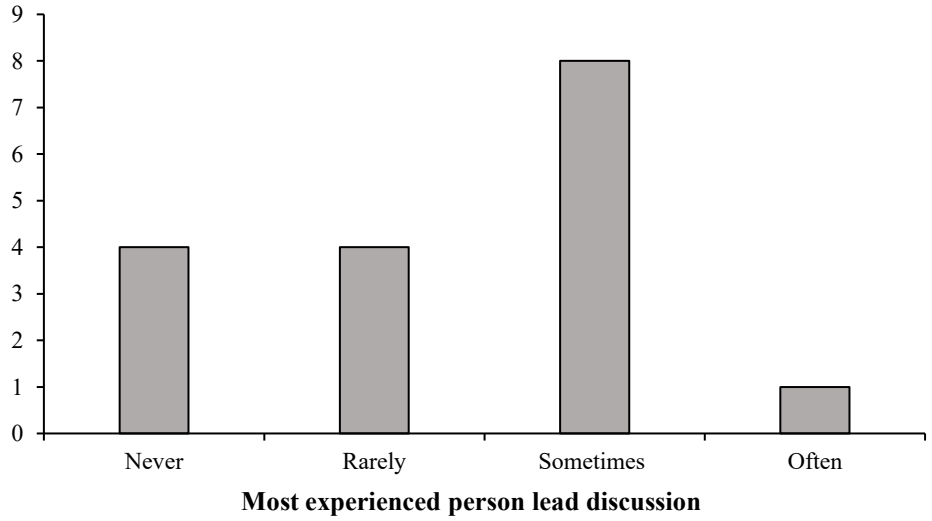
Equal contribution



This figure shows how frequently the most experienced person leads the discussion during the consensus approach. The majority indicated that “Sometimes” the most experienced person leads the discussion. This is followed by “Never” and “Rarely”, which were both answered four times.

Figure 12

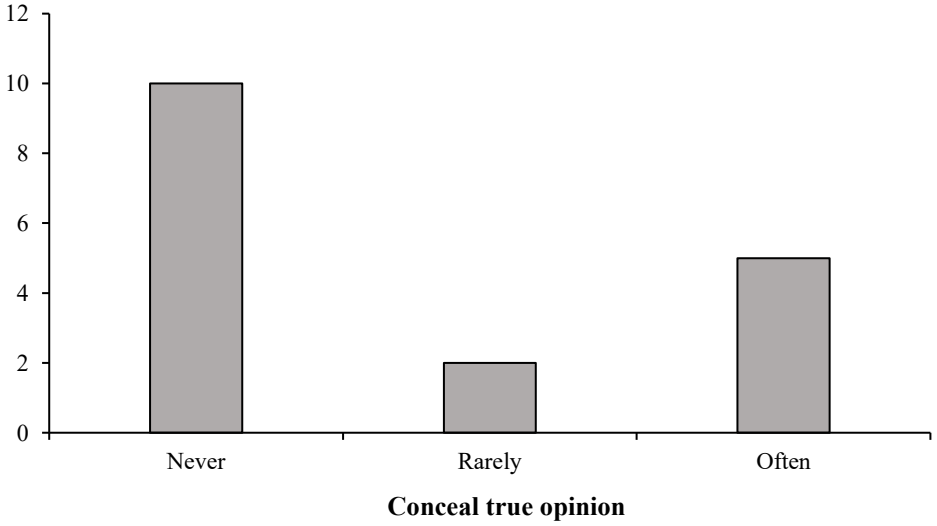
Most experienced person lead the discussion



This figure shows if participants concealed their true opinion when it differed from that of their consensus group. The majority of participants reported that they “Never” (n=10) concealed their true opinion. Followed by “Often” which was reported five times and “Rarely” which was reported two times.

Figure 13

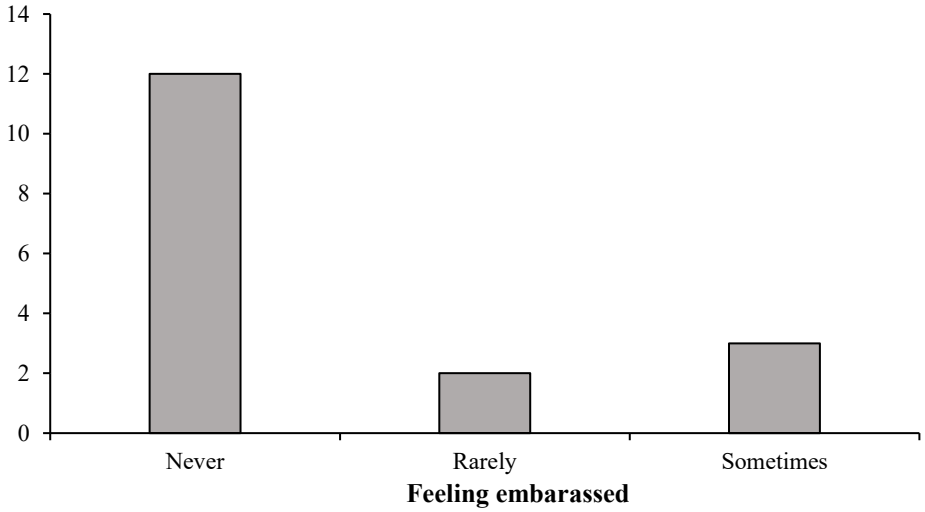
Conceal true opinion



This figure shows how frequently participants felt embarrassed during the consensus approach when their opinions differed from that of the majority. Here most participants answered “Never” (n =12). This is followed by “Sometimes” with a count of three and “Rarely” with a count of two.

Figure 14

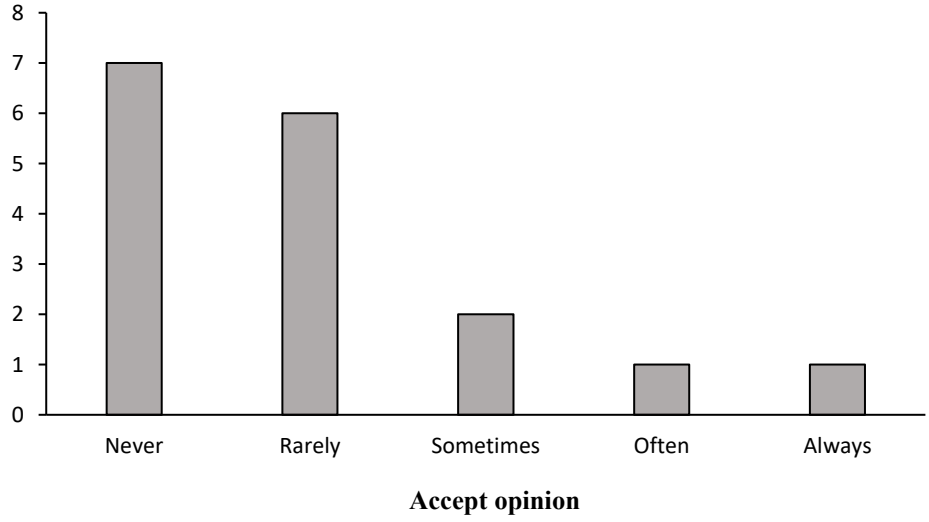
Feeling embarrassed



This figure shows how frequently the participants accepted an opinion even if they did not share the same opinion on a mark during the consensus approach. The majority answered “Never” (n=7), followed by “Rarely” (n=6).

Figure 15

Accept opinion



4.Discussion

This study aimed to assess if a consensus approach is an applicable method to increase the reliability of fingerprint examination. To test the consensus approach, the research question "What impact does the implementation of a consensus approach in fingerprint examination have on psychological safety and stress?" was constructed. Overall, it was evident that psychological safety was higher in the post consensus phase and that stress could be effectively reduced.

4.1 Psychological safety

A significant difference between the baseline levels of psychological safety and psychological safety in the consensus approach has been found. In detail, the scores of psychological safety were higher in the consensus approach than compared to the psychological safety scores in the workplace. As high levels of psychological safety support knowledge sharing, creativity and performance, it is a promising sign that the consensus approach led to an increase of psychological safety (Newman et al., 2017). However, these findings have to be reviewed with caution. Since the participants were clearly instructed to keep a respectful atmosphere during discussion and were informed that this was not a performance study. Contrary to the safe environment of the study, the work environment of fingerprint analysts is different. Other confounding factors like stress, fear of making an error and hierarchy should not be disregarded when comparing both conditions. Hence, an occurrence of the Hawthorne effect could serve as another explanation for this result (Sedgwick & Greenwood, 2015). The Hawthrone effect was first observed during productivity studies in the Hawthorne factory between 1924 to 1933 and describes a change in the behaviour of participants when observed (Muldoon & Zoller, 2020; Paradis & Sutkin, 2016).

Furthermore, low levels of psychological safety can lead to worse team performance in cases where different levels of expertise are present in a group (Martins et al., 2012). Therefore, the effect of different levels of expertise on psychological safety was also investigated. In the low expertise group, psychological safety levels increased significantly in the consensus condition when compared to the non-consensus condition. In the high expertise group, expertise had no significant effect on psychological safety.

4.2 Stress

The variable stress was measured in the context of the workplace of fingerprint examiners and during the consensus approach. After the results of both measurements were compared, it can be concluded that there was a decrease in stress during the consensus approach. As stress can potentially impair the quality of decision making it was investigated whether there are differences in stress levels (Pabst et al., 2013). The finding that there was a slight decrease in stress levels constitutes to the effectiveness of the consensus approach. However, the same consideration as for the variable psychological safety must be addressed. Since an occurrence of the Hawthorne effect cannot be ruled out (Sedgwick & Greenwood, 2015). Although, stress levels were significantly lower in the post condition compared to the pre-condition. Nonetheless, at the same time stress levels in the workplace were only slightly higher than during the consensus approach. This was surprising because the profession of fingerprint examination is associated with time pressure and a high caseload, therefore stress was hypothesised to be much higher in the pre-condition (Almazrouei et al., 2021).

In addition, unfamiliarity with the consensus approach and different levels of expertise were also hypothesised to lead to increased stress levels. However, there is still a significant difference between the two measures. Hence, the employment of the consensus approach did lower stress levels and may in turn lead to better decision quality of the participants. In a study conducted by Hebles et al. (2022), a positive effect of psychological safety on stress has been found, therefore it can also be argued that stress in the post condition was in part lower because of high psychological safety levels.

4.3 Consensus tools

In this study, participants were given a set of tools to support their decision making. PiAnoS was one of these tools, but the usefulness of single functions from this software were also investigated. Here, the PiAnoS tutor mode was considered the most helpful tool in reaching a consensus. The other functions LogLR, ESLR and SLE did not reach this level of positive reception among the participants. Despite this, many participants gave their approval to use the PiAnoS software in the future. In addition, a majority indicated that PiAnoS would be beneficial for “personal education and skill enhancement”, “trainee development” and “consensus building” during complex mark analysis.

The guided checklist also received a positive reception among participants as well as the consensus roles Facilitator and Driver. However, the role of Timekeeper was disregarded by most participants with some of them calling it obsolete, since time was not an issue.

Lastly, the consensus cards were also perceived as obsolete and did not find much use. In conclusion, the positive reception of PiAnoS, the guided checklist and roles gives an indication that fingerprint examiners prefer a guided form of consensus. Moreover, PiAnoS could find employment for training new fingerprint analysts and is shown to be an attractive tool for skilled fingerprint analysts.

4.4 Consensus at work

The following section revolves around how and if the participants resort to consensus in their workplace. For example, the majority does engage in technical discussion about fingerprint examination on a weekly basis. Both formal and informal discussions about specific marks happen rarely. This finding indicates that there exists an interest in applying consensus in the workplace, however in reality it is seldom done in a guided and constructive manner. When consensus is applied the biggest issue in reaching it is the complexity of the mark and the seniority of other examiners.

4.5 Negative consensus behaviours

Certain behaviours in intergroup situations were hypothesised to have a negative impact on achieving consensus. The fact that participants reported seniority and experience and interpersonal dynamics to be an issue in establishing a consensus, indicated that these issues could persist in the consensus approach. First, changing one's opinion due to pressure from group members rarely or never happened. In addition, equal contribution occurred in most of the groups. This is a promising finding as this gives a first indication that behaviours being antecedents of groupthink were almost not present during the consensus approach. In some cases, the most experienced person leads the discussion, however, this can result in two different extremes. First, the most experienced person could lead the discussion authoritatively and would not respond to criticism which would also be a form of overconfident and manipulative behaviour and therefore be disruptive to consensus building (Dong et al., 2021). It could have also been the case that the most experienced did not lead the discussion with an iron hand but rather had a supportive role. This could be the case in groups where the levels of expertise were significantly different as students were also taking part in this study.

When their opinion differed from that of the group most participants did not decide to conceal their opinion, which is additional evidence that antecedents of groupthink either did not occur or only occurred on a small scale. In addition, the high psychological safety levels

can be attributed to this finding. Feeling embarrassed when an opinion differed from the rest of the group, also only occurred on a small scale. This finding can also be connected to the high levels of psychological safety. Accepting the opinion of the group when it was contrary to one's own opinion, occurred during the consensus approach. However, almost half of the respondents said they never accepted an opinion when it differed from their own. At first sight, this seems to be a negative outcome of consensus, but participants always had the option to not reach a full consensus with the group. So, in the instance that consensus had not been reached it is likely that respondents did not accept the opinion of others. All in all, negative consensus behaviours did occur during the consensus approach, but only on a small scale. In general, the findings suggest that groupthink tendencies had no negative effects on consensus and the participants were able to freely voice their opinions without fear of judgement from their groups.

4.6 Limitations

Although this research gives valuable insights into the applicability of a new method for fingerprint examination, it also has its downsides. From the previously 23 participants, only the responses from seventeen participants could be analyzed. For the other 5 important demographics were missing. In addition, a portion of the participants did not fill in their identification number on both questionnaires, which made it impossible to identify them across the pre- and post-questionnaire which was crucial for the repeated measures anova. Since this skewed the data, it would be useful to prevent this in the future. One way to do this is by installing a feature into the questionnaire which makes the answer field mandatory to be filled out. For example, in the software qualtrics, an answer can be "forced" (Response Requirements & Validation, 2024). If left blank, the field will pop up in a red colour and a message will appear that this is still missing and is required to be filled out. Without filling it out, the questionnaire cannot be continued.

Regarding the size of the sample, it remains difficult to obtain a larger sample size as the population for this research is quite small. Fingerprint examiner is a specific category and field of work. Furthermore, the goal of this research is not to compare the results to a general population, rather it is used to explore the usage of a consensus approach in the future.

As previously stated, the setup of the study and the emphasis on respectful communication may have biased the results towards higher psychological safety and lower stress. The last point of criticism is the scale used to identify if and in what frequency negative group behaviours occurred. This scale was constructed by the researcher with

possible negative group behaviours in mind, as there were no pre-existing questionnaires. Since this scale has been used for the first time, validity and test-retest reliability cannot be guaranteed. It cannot be confirmed that this scale measured what it was intended to measure. Using a pre-existing questionnaire might have been more accurate. However, self-constructing this scale had the advantage that it was tailored to the context of the consensus approach.

4.7 Future research

To directly take inspiration from the limitations, future research regarding the consensus approach should more directly focus on negative group behaviours and apply a more robust measure to identify its occurrence. Before completely applying this approach in the work field of fingerprint examination, additional studies are required to verify its usefulness. As time is of the essence during fingerprint examination and many examiners suffer under time constraints, it should be investigated how the consensus approach influences the workflow of examiners. It takes time to discuss opinions and find a consensus, and that could be time that examiners are not able to sacrifice. Therefore, a time effective method of the consensus approach should be investigated. As stated previously, the Hawthrone effect could have occurred in this study. Hence, it is important to test this approach in the workplace of fingerprint analysts without the safe environment of a study. Only then a clear recommendation of the consensus approach can be given. One way this could be done is through an audit. This means an external auditor visits the facility and controls if certain processes are executed adequately.

4.8 Conclusion

Both psychological safety and stress profit from the employment of a consensus approach and subsequently erroneous decision making in fingerprint examination can be further mitigated against. Furthermore, the tools provided to participants were successful in facilitating a consensus. The PiAnoS tool was especially considered helpful by the participants. Not only does it enhance the establishment of a consensus it is also useful in the day-to-day work of fingerprint examination since it helps with documentation and in turn a more robust decision-making process. In addition, using PiAnoS for training purposes of future fingerprint analysts is recommended. It should be considered to apply a consensus approach during criminal investigations. But as already mentioned, future research is needed to make sure how to make use of consensus most effectively without running into the risk of

factors like seniority and groupthink. All in all, the implementation of a consensus approach is seen as a useful intervention to improve the decision quality of fingerprint examiners.

Reference list

Almazrouei, M. A., Morgan, R. M., & Dror, I. E. (2021). Stress and support in the workplace: The Perspective of Forensic Examiners. *Forensic Science International: Mind and Law*, *2*, 100059. https://doi.org/10.1016/j.fsiml.2021.100059

Asch, S.E. (1951). Effects of Group Pressure on the Modification and Distortion of Judgments. In Guetzknow, H., Ed., Groups, Leadership and Men, Pittsburgh, PA, Carnegie Press, 177-190.

Bigun, J. (2009). Fingerprint features. Encyclopedia of Biometrics, 465–473. https://doi.org/10.1007/978-0-387-73003-5_50

Cooper, G. S., & Meterko, V. (2019). Cognitive Bias Research in forensic science: A systematic review. Forensic Science International, 297, 35–46. https://doi.org/10.1016/j.forsciint.2019.01.016

Cottrell, S. R. D. (2023). Effective team decision-making : exploring the role of psychological safety. (Doctoral dissertation, University of Canterbury]. http://dx.doi.org/10.26021/14265.

Ditrich, H. (2015). Cognitive fallacies and criminal investigations. Science & Justice, 55(2), 155–159. https://doi.org/10.1016/j.scijus.2014.12.007

Dong, Y., Zha, Q., Zhang, H., & Herrera, F. (2021). Consensus reaching and strategic manipulation in group decision making with trust relationships. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 51(10), 6304–6318. https://doi.org/10.1109/tsmc.2019.2961752

Dror, I. E., & Charlton, D. (2006). Why experts make errors. Journal of Forensic Identification, 56(4)

Dror, I. E., Charlton, D., & Péron, A. E. (2006). Contextual information renders experts vulnerable to making erroneous identifications. Forensic Science International, 156(1), 74–78. https://doi.org/10.1016/j.forsciint.2005.10.017

Dror, I. E., & Stoel, R. D. (2014). Cognitive forensics: Human cognition, contextual information, and Bias. Encyclopedia of Criminology and Criminal Justice, 353–363. https://doi.org/10.1007/978-1-4614-5690-2_147

Edmond, G. (2022). Latent justice? A review of adversarial challenges to fingerprint evidence. Science & Justice, 62(1), 21–29. https://doi.org/10.1016/j.scijus.2021.10.006

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. Administrative Science Quarterly, 44(2), 350–383. https://doi.org/10.2307/2666999

Ezegbogu, M. O., & Omede, P. I.-O. (2022). The admissibility of fingerprint evidence: An African perspective. Canadian Society of Forensic Science Journal, 56(1), 23–41. https://doi.org/10.1080/00085030.2022.2068404

Georgiou, N., Morgan, R. M., & French, J. C. (2020). Conceptualising, evaluating and communicating uncertainty in forensic science: Identifying commonly used tools through an interdisciplinary configurative review. Science & Justice, 60(4), 313–336. https://doi.org/10.1016/j.scijus.2020.04.002

Haber, R. N., & Haber, L. (2014). Experimental results of fingerprint comparison validity and reliability: A review and Critical Analysis. Science & Justice, 54(5), 375–389. https://doi.org/10.1016/j.scijus.2013.08.007

Hebles, M., Trincado-Munoz, F., & Ortega, K. (2022). Stress and turnover intentions within healthcare teams: The mediating role of Psychological Safety, and the moderating effect of covid-19 worry and supervisor support. Frontiers in Psychology, 12. https://doi.org/10.3389/fpsyg.2021.758438

Hicklin, R. A., Ulery, B. T., Ausdemore, M., & Buscaglia, J. (2020). Why do latent fingerprint examiners differ in their conclusions? Forensic Science International, 316, 110542. https://doi.org/10.1016/j.forsciint.2020.110542

Kellermanns, F. W., Walter, J., Floyd, S. W., Lechner, C., & Shaw, J. C. (2011). To agree or not to agree? A meta-analytical review of strategic consensus and Organizational Performance. Journal of Business Research, 64(2), 126–133. https://doi.org/10.1016/j.jbusres.2010.02.004

Leonetti, C. (2024). Ensuring the reliability of evidence in the New Zealand criminal courts: The Admissibility of Forensic Science. Common Law World Review, 53(4), 197–222. https://doi.org/10.1177/14737795241237799

Malhotra, A., Sankaran, A., Vatsa, M., Singh, R., Morris, K. B., & Noore, A. (2021). Understanding ace-V latent fingerprint examination process via eye-gaze analysis.

IEEE Transactions on Biometrics, Behavior, and Identity Science, 3(1), 44–58. https://doi.org/10.1109/tbiom.2020.3027144

Marcatto, F., Di Blas, L., Luis, O., Festa, S., & Ferrante, D. (2022). The perceived occupational stress scale. European Journal of Psychological Assessment, 38(4), 293–306. https://doi.org/10.1027/1015-5759/a000677

Martins, L. L., Schilpzand, M. C., Kirkman, B. L., Ivanaj, S., & Ivanaj, V. (2012). A contingency view of the effects of cognitive diversity on Team Performance. Small Group Research, 44(2), 96–126. https://doi.org/10.1177/1046496412466921

Meuwly, D. (2014). Forensic use of fingerprints and fingermarks. Encyclopedia of Biometrics, 1–15. https://doi.org/10.1007/978-3-642-27733-7_181-3

Miller, C. E. (1989). The social psychological effects of group decision rules. In P. B. Paulus (Ed.), Psychology of group influence (2nd ed., pp. 327–355). Lawrence Erlbaum Associates, Inc.

Moore, A., & O'Doherty, K. (2013). Deliberative voting: Clarifying consent in a consensus process. Journal of Political Philosophy, 22(3), 302–319. https://doi.org/10.1111/jopp.12028

Muldoon, J., & Zoller, Y. (2020). Contested paths: A meta-analytic review of the Hawthorne Studies Literature. Handbook of Research on Management and Organizational History. https://doi.org/10.4337/9781788118491.00010

Newman, A., Donohue, R., & Eva, N. (2017). Psychological safety: A systematic review of the literature. Human Resource Management Review, 27(3), 521–535. https://doi.org/10.1016/j.hrmr.2017.01.001

Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. Review of General Psychology, 2(2), 175–220. https://doi.org/10.1037/1089-2680.2.2.175

Pabst, S., Brand, M., & Wolf, O. T. (2013). Stress and decision making: A few minutes make all the difference. Behavioural Brain Research, 250, 39–45. https://doi.org/10.1016/j.bbr.2013.04.046

Palomares, I., Martinez, L., & Herrera, F. (2014). A consensus model to detect and manage noncooperative behaviors in large-scale group decision making. IEEE Transactions on Fuzzy Systems, 22(3), 516–530. https://doi.org/10.1109/tfuzz.2013.2262769

Paradis, E., & Sutkin, G. (2016). Beyond a good story: From Hawthorne effect to reactivity in Health Professions Education Research. Medical Education, 51(1), 31–39. https://doi.org/10.1111/medu.13122

Phillips-Wren, G., & Adya, M. (2020). Decision making under stress: The role of information overload, time pressure, complexity, and uncertainty. Journal of Decision Systems, 29(1), 213–225. https://doi.org/10.1080/12460125.2020.1768680

Response Requirements & Validation. Qualtrics XM: The Leading Experience Management Software. (2024, November 7). https://www.qualtrics.com/support/survey-platform/survey-module/editing-questions/validation/

Ribeiro, G., Tangen, J. M., & McKimmie, B. M. (2019). Beliefs about error rates and human judgment in forensic science. Forensic Science International, 297, 138–147. https://doi.org/10.1016/j.forsciint.2019.01.034

Sedgwick, P., & Greenwood, N. (2015). Understanding the hawthorne effect. BMJ. https://doi.org/10.1136/bmj.h4672

Sikorski, M. (2022). Is forensic science in crisis? Synthese, 200(3). https://doi.org/10.1007/s11229-022-03685-z

Solomon, M. (2006). Groupthink versus The Wisdom of Crowds: The social epistemology of deliberation and dissent. The Southern Journal of Philosophy, 44(S1), 28–42. https://doi.org/10.1111/j.2041-6962.2006.tb00028.x

Stoel, R. D., Berger, C. E., Kerkhoff, W., Mattijssen, E. J., & Dror, I. E. (2015). Minimizing contextual bias in forensic casework. Forensic Science and the Administration of Justice: Critical Issues and Directions, 67–86. https://doi.org/10.4135/9781483368740.n5

Stoney, D. A., De Donno, M., Champod, C., Wertheim, P. A., & Stoney, P. L. (2020). Occurrence and associative value of non-identifiable fingermarks. Forensic Science International, 309, 110219. https://doi.org/10.1016/j.forsciint.2020.110219

Størseth, F., Hauge, S., & Tinmannsvik, R. K. (2014). Safety barriers: Organizational potential and forces of psychology. Journal of Loss Prevention in the Process Industries, 31, 50–55. https://doi.org/10.1016/j.jlp.2014.06.006

Taroni, F., & Biedermann, A. (2015). Uncertainty in forensic science: experts, probabilities and Bayes' theorem. Italian Journal of Applied Statistics, 27(2), 129-144.

Ulery, B. T., Hicklin, R. A., Buscaglia, J., & Roberts, M. A. (2011). Accuracy and reliability of forensic latent fingerprint decisions. Proceedings of the National Academy of Sciences, 108(19), 7733–7738. https://doi.org/10.1073/pnas.1018707108

Ulery, B. T., Hicklin, R. A., Roberts, M. A., & Buscaglia, J. (2016). Interexaminer variation of Minutia markup on latent fingerprints. Forensic Science International, 264, 89–99. https://doi.org/10.1016/j.forsciint.2016.03.014

Warden, R., & Aikins, M. (2014, April 27). Stephan Cowans. The National Registry of Exonerations. https://www.law.umich.edu/special/exoneration/Pages/casedetail.aspx?caseid=312

Zapf, P. A., & Dror, I. E. (2017). Understanding and mitigating bias in forensic evaluation: Lessons from forensic science. International Journal of Forensic Mental Health, 16(3), 227–238. https://doi.org/10.1080/14999013.2017.1317302

Appendices

Appendix A

Pre consensus questionnaire

Pre consensus survey

Start of Block: Start

Welcome
Dear Participant, welcome and thank you for your participation in the survey study, "Computer-Assisted Consensus Approach among Forensic Examiners," led by Bernd Göttker from the Faculty of Behavioural, Management, and Social Sciences at the University of Twente.

The purpose of this survey is to explore consensus behaviours in the work environment of fingerprint examination. It is assumed that a consensus approach could enhance the reliability of the fingerprint comparison process. In addition, the effect of stress on decision-making will be explored in this context. Your contribution is estimated to take approximately 30 minutes and will significantly contribute to advancing the field of fingerprint examination. As a participant, you will begin by answering a demographic questionnaire, followed by questions about complex marks, your engagement in discussions and consensus building with your co-workers, and an examination of the working atmosphere in your team. The survey concludes with a short questionnaire about perceived stress.

Your involvement in this study is entirely voluntary. You have the freedom to withdraw at any stage or to skip any questions you prefer not to answer.

Ethical Approval: This study strictly adheres to ethical guidelines and has received full approval from the BMS Ethical Committee of the University of Twente. Ethical request number: 231359

In compliance with GDPR, your participation is confidential, and your responses will remain anonymous, ensuring that the results cannot be linked back to you.

Contact Information: For any further information or questions, feel free to contact Bernd Göttker at b.goettker@student.utwente.nl.

We greatly appreciate your time and valuable input in this research. Your insights are crucial to the success of our study. Thank you once again for your participation.

End of Block: Start

Start of Block: Demographic

D The first section of the survey gathers general demographic data, including age, gender, work experience, and educational background

---

Page Break

Q1 Please enter your PiAnoS ID

_____

Q2 How old are you?

- ○ Under 25 (4)
- ○ 25-34 (5)
- ○ 35-44 (6)
- ○ 45-54 (7)
- ○ 55-64 (8)
- ○ 65 and above (9)
- ○ Prefer not to say (10)

Q3 Gender Identity

- ○ Male (4)
- ○ Female (5)
- ○ Non-binary / third gender (6)
- ○ Prefer not to say (7)
- ○ Prefer to self-describe (8)

_____

Q4 Highest Level of Education

- ○ High school diploma or equivalent (1)
- ○ Some college, no degree (2)
- ○ Associate degree (3)
- ○ Bachelor´s degree (4)
- ○ Master´s degree (5)
- ○ Doctoral or professional degree (6)
- ○ Prefer not to say (7)

Q5 How long have you been working as a fingerprint analyst?

- ○ Less than 1 year (1)
- ○ 1-5 years (2)
- ○ 6-10 years (3)
- ○ 11-15 years (4)
- ○ 16-20 (5)
- ○ More than 20 years (6)

Q6 Primary work setting

- ○ Law enforcement agency (1)
- ○ Private sector (2)
- ○ Academic institution (3)
- ○ Government (non-law enforcement) (4)
- ○ Freelance/Independent (5)
- ○ Other (6) _____

Q7 Certifications in Fingerprint Analysis

- ○ Certified Latent Print Examiner (1)
- ○ Certified Tenprint Examiner (2)
- ○ Other certification (please specify) (3)

_____

- ○ No certifications (4)

Q8 Type of Training received

- ○ Formal university/college education in forensic science (1)
- ○ Specialized training programs (e.g. specific to fingerprint expertise) (2)
- ○ On-the-job training (3)
- ○ Other (please specify) (4)

_____

Q9 Experience with different Fingerprint Analysis Techniques

- ○ Latent print analysis (1)
- ○ Tenprint (rolled prints) analysis (2)
- ○ AFIS (3)
- ○ Crime scene processing for fingermarks (4)
- ○ Other (please specify) (5)

_____

Q10 Experience with Automated Fingerprint Identification System (AFIS)

- ○ No experience (1)
- ○ Limited experience (2)
- ○ Moderate experience (3)
- ○ Extensive experience (4)

Q11 Participation in proficiency testing

- ○ Regularly participate in proficiency testing (1)
- ○ Occasionally participate in proficiency testing (2)
- ○ Never participated (3)
- ○ Not applicable to my role (4)

Q12 Primary Role in the Field of Fingerprint Analysis

- ○ Latent print analysis (1)
- ○ Tenprint (rolled prints) analysis (2)
- ○ AFIS (3)
- ○ Crime scene processing for fingermarks (4)
- ○ Not applicable (5)

Q13 Level of Expertise

- ○ Novice (1)
- ○ Beginner (2)
- ○ 1-star (3)
- ○ 2-star (4)
- ○ 3-star (5)

Page Break

Complex marks The next set of questions relates to complex marks.

Page Break

Q14 How would you define a complex mark? (for example number of minutiae and clarity of ridges)

_____

Q15 How often do you think you come across a complex mark during your work?

- ○ Daily (1)
- ○ Weekly (2)
- ○ Monthly (3)
- ○ Occasionally (less frequent than monthly) (4)
- ○ Rarely (only a few times a year) (5)
- ○ Never (7)

Q16 How often do you have discussions about the technical aspects of fingerprint analysis with your peers?

- ○ Daily (1)
- ○ Weekly (2)
- ○ Monthly (4)
- ○ Occasionally (less frequent than monthly) (5)
- ○ Rarely (only a few times a year) (6)
- ○ Never (7)

End of Block: Demographic

Start of Block: Consensus

C The next set of questions aims to explore various aspects of consensus discussions within your work environment. We are interested in understanding: (1) The frequency of these discussions, (2) the nature of these discussions, whether they are formal (involving structured procedures) or informal (casual discussions among peers), (3) your personal perceptions and experiences regarding the process and outcomes of reaching consensus. Your responses will help us gain valuable insights into the dynamics of consensus-building in your professional setting

End of Block: Consensus

Start of Block: Consensus

Q17 How often do you engage in informal discussions about a mark with your peers?

- Daily (1)
- Weekly (2)
- Monthly (3)
- Occasionally (less frequent than monthly) (4)
- Rarely (only a few times a year) (5)
- Never (6)

Q18 How often do you engage in formal discussions or meetings about a mark with your peers?

- ○ Daily (1)
- ○ Weekly (2)
- ○ Monthly (3)
- ○ Occasionally (less frequent than monthly) (4)
- ○ Rarely (only a few times a year) (5)
- ○ Never (6)

Q19 When engaged in informal discussions about a mark, how often do you reach a consensus with your peers?

- ○ Never (1)
- ○ Seldom (2)
- ○ Don´t know (3)
- ○ Often (4)
- ○ Always (5)

Q20 How often do you resort to a formal consensus (multiple procedure) when there's a difference in opinion about a mark?

- ○ Never (1)
- ○ Seldom (2)
- ○ Don´t know (3)
- ○ Often (4)
- ○ Always (5)

Q21 In your experience, how often does the formal consensus procedure result in an agreed-upon decision on the source of the mark?

- ○ Never (1)
- ○ Seldom (2)
- ○ Don´t know (3)
- ○ Often (4)
- ○ Always (5)

Q22 When consensus is reached (either informally or through formal procedures), how often is the decision implemented as discussed?

- ○ Never (1)
- ○ Seldom (2)
- ○ Don´t know (3)
- ○ Often (4)
- ○ Always (5)

Q23 How often do you find that the outcome of a consensus discussion (whether formal or informal) is satisfactory to all parties involved?

- ○ Never (1)
- ○ Seldom (2)
- ○ Don´t know (3)
- ○ Often (4)
- ○ Always (5)

Q24 Please indicate your level of agreement with the following statement: 'Reaching a consensus when there is a difference of opinion is important.'

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q25 In your opinion, what are the main obstacles to achieving consensus in fingerprint analysis discussions? (Please select all that apply)

- ☐ Seniority or experience of the other expert (Differences in levels of experience or authority) (1)
- ☐ Examiner behaviour (Interpersonal dynamics or communication styles) (2)
- ☐ The complexity of the mark (Intricacies or ambiguities in the fingerprint) (3)
- ☐ Process-related-issues (procedural or systemic challenges) (4)
- ☐ Other (please specify) (5)

_____

Q26 When entering a discussion about a mark, what is your primary goal? Please select the option that best represents your main objective

- ○ Reaching a consensus on the analysis (1)
- ○ Understanding different viewpoints and interpretations (2)
- ○ Sharing and advocating for my viewpoint (3)
- ○ Ensuring the accuracy and reliability of the analysis (4)
- ○ Collaborating to enhance overall team understanding (5)
- ○ Other (Please specify) (6)

_____

Q27 What would you consider a successful outcome from a discussion about a mark?

- [ ] Clear Agreement: All parties reach a unanimous agreement on the mark. (1)
- [ ] Clarified Understanding: Even if no agreement is reached, all parties have a clearer understanding of each other's perspectives. (2)
- [ ] Compromise: A compromise is made that most parties can accept, even if it's not their first choice. (3)
- [ ] Decision to Re-evaluate: A decision to revisit or re-evaluate the mark at a later time. (4)
- [ ] Agreement to Disagree: A mutual understanding that consensus cannot be reached, but with respect for differing opinions. (5)
- [ ] Identification of Key Issues: Identifying the key issues or points of contention for further investigation or discussion. (6)
- [ ] Further Research or Data Collection: Recognizing the need for additional data or research to inform the decision. (7)
- [ ] Referral to a Third Party: Deciding to refer the matter to a neutral third party or expert for resolution (8)
- [ ] No Change, but Constructive Dialogue: The mark remains unchanged, but the discussion was constructive and enlightening. (9)
- [ ] other (10) _____

End of Block: Consensus

Start of Block: psy safety

p The next set of questions are about your experiences in the workplace and how your coworkers interact with you.

End of Block: psy safety

Start of Block: psy safety

Q28 In this team, it is easy to discuss difficult issues and problems.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q29 If I admit to an error or mistake, I will not face retaliation or criticism in my workplace.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q30 It is easy to ask members of this team for help.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q31 I feel safe offering new ideas, even if they are not fully-formed plans.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q32 In this team, people are accepted for being different.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q33 My peers welcome my ideas and give them time and attention.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q34 My peers value the contributions of others´.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q35 People talk about mistakes and ways to improve and learn from them.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Page Break

Q36 We take time to find new ways to improve our team´s work processes.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q37 My peers are encouraged to raise concerns they have about team plans or decisions.

- Strongly disagree (1)
- Disagree (2)
- Neither agree nor disagree (3)
- Agree (4)
- Strongly agree (5)

Q38 We try to discover our underlying assumptions and seek counterarguments about issues under discussion.

- Strongly disagree (1)
- Disagree (2)
- Neither agree nor disagree (4)
- Agree (5)
- Strongly agree (6)

End of Block: psy safety

Start of Block: workplace stress

w The next set of questions are about your perception of stress in the work place.

End of Block: workplace stress

Start of Block: workplace stress

Q39 My work is stressful.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q40 Thinking about my work makes me feel tense.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q41 At work I feel under pressure.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q42 My work has negative effects on my health.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q43 I feel stress when I reach a different opinion to my peers

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q44 I change my opinion to reduce my stress

- Strongly disagree (1)
- Disagree (2)
- Neither agree nor disagree (3)
- Agree (4)
- Strongly agree (5)

Q45 I feel stress at the thought of making an error

- Strongly disagree (1)
- Disagree (2)
- Neither agree nor disagree (3)
- Agree (4)
- Strongly agree (5)

End of Block: workplace stress

Appendix B

Post consensus questionnaire

Post consensus survey

Start of Block: Start

Welcome

Dear Participant, welcome to the second part of our survey "Computer-Assisted Consensus Approach among Forensic Examiners,". This segment is designed to explore your experiences while seeking consensus within your group. First, you will be asked about the tools you utilized during the consensus phase. After that, we will assess behaviors that may have occurred during the consensus approach, followed by questions about the working atmosphere during this process. The survey concludes by measuring your perceived stress throughout the consensus approach. Your contribution is estimated to take approximately 30 minutes.

Please note that the questions related to working atmosphere and stress may resemble those from the first part of the survey; however, they specifically apply to your experiences during the consensus approach and should not be confused with your broader work environment.

Your involvement in this study is entirely voluntary. You have the freedom to withdraw at any stage or to skip any questions you prefer not to answer.

This study strictly adheres to ethical guidelines and has received full approval from the BMS Ethical Committee of the University of Twente. Ethical request number : 231359

In compliance with GDPR, your participation is confidential, and your responses will remain anonymous, ensuring that the results cannot be linked back to you.

Contact Information: For any further information or questions, feel free to contact Bernd Göttker at b.goettker@student.utwente.nl.

We greatly appreciate your time and valuable input in this research. Your insights are crucial to the success of our study. Thank you once again for your participation.

End of Block: Start

Start of Block: General information

Q1 Please enter your PiAnoS ID

_____

Q2 Have you completed the pre-consensus survey?

- ○ Yes (1)
- ○ No (2)

End of Block: General information

CM The next set of questions relates to complex marks.

Page Break

Q15 How would you define a complex mark? (for example, number of minutiae and clarity of ridges)

_____

Q16 How often do you think you come across a complex mark during your work?

- ○ Daily (1)
- ○ Weekly (2)
- ○ Monthly (3)
- ○ Occasionally (less frequent than monthly) (4)
- ○ Rarely (only a few times a year) (5)
- ○ Never (7)

Q17 How often do you have discussions about the technical aspects of fingerprint analysis with your peers?

- ○ Daily (1)
- ○ Weekly (2)
- ○ Monthly (4)
- ○ Occasionally (less frequent than monthly) (5)
- ○ Rarely (only a few times a year) (6)
- ○ Never (7)

End of Block: Demographic

Start of Block: Consensus tools

CT The next set of questions are about the tools you were provided during the computer-assisted consensus approach. These tools are the guided checklist, your assigned role (Facilitator, Timekeeper, Driver), the coloured consensus consent and dissent cards, the PiAnoS Tutor mode (to view individual markups), and the PiAnoS score-based likelihood ration (SLR).

End of Block: Consensus tools

Start of Block: Consensus tools

Q18 Did you find the consensus tools provided during the computer-assisted consensus approach helpful in reaching a consensus?

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly Agree (5)

Page Break

Q19 During the computer-assisted consensus approach which of the tools were most helpful in reaching a consensus ? (you may select more than one)

Consensus Roles (Facilitator, Time-keeper, Driver (annotater)) (1)

☐ • Facilitator (F): Guide the group, ensuring focus on tasks and use of consensus cards. Confirm completion of the consensus checklist (7)

☐ • Timekeeper (T): Monitor session time. Alert the group at 30 minutes to allocate time for analysis, comparison, and evaluation (9)

☐ • Driver (D): Manage the PiAnoS tool. Duties include annotating cores, recording judgments in case notes, documenting consensus results, and handling the ESLR/SLR function (8)

☐ Guided checklist (2)

☐ PiAnoS Tutor Mode (Tutor mode allows you to visualise multiple examiner examinations once) (3)

☐ PiAnoS expected score-based likelihood ratio (ESLR) tool (to help inform the consensus suitability conclusion (mark holds value/ mark is no value) (10)

☐ PiAnoS SLR tool (to help inform the consensus source conclusion (support for same source/ different source) (11)

☐ Consensus cards (green (wish to speak/ agree), yellow (direct response/reservations),orange Standaside),red Procedural point/objection)) (4)

☐ The provided tools were equally helpful (5)

☐ None of the tools were helpful (6)

Q20 I followed my assigned role during the computer-assisted consensus approach.

- ○ Rarely or Never - I rarely or never followed my assigned role (1)
- ○ Sometimes - I sometimes did not follow my assigned role (2)
- ○ Occasionally - I occasionally followed my assigned role (3)
- ○ Usually - I usually followed my assigned role (4)
- ○ Always - I always followed my assigned role (5)

Q21 Did you use the consensus cards during the computer-assisted consensus approach?

- ○ Never (1)
- ○ Sometimes (2)
- ○ Occasionally (3)
- ○ Usually (4)
- ○ Always (6)

Q22 I would recommend the PiAnoS tool to my supervisor for facilitating a consensus.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q23 I used the LogLR function of PiAnoS to inform my groups decision making.

- ○ Never (1)
- ○ Sometimes (2)
- ○ Occasionally (3)
- ○ Usually (4)
- ○ Always (5)

Q24 I would use the PiAnoS tool in the future if it was an option.

- ○ Strongly Disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q25 If available, for which of the following purposes would you consider using the PiAnoS tool in the future? (Select all that apply)

- ☐ Personal Education and Skill Enhancement: To enhance my own knowledge and skills. (1)
- ☐ Trainee Development: To assist in the training and development of new trainees. (2)
- ☐ Proficiency Tests: For documenting my proficiency tests. (3)
- ☐ Practical Casework: To aid in regular casework activities (4)
- ☐ Complex Mark Analysis: For consensus-building in procedures involving complex marks. (5)
- ☐ Court Reports: To use as a tool in preparing and presenting court reports. (6)
- ☐ Supporting Opinions: To provide additional support and validation for my professional opinions. (7)

Q26 I would consider using the PiAnoS ESLR (Expected Scorebased Likelihood Ratio) and SLR (Sourcebased Likelihood Ratio) functions in the future for the following purposes:

(Select all that apply)

☐ • Suitability Decisions: Using the ESLR to inform my decisions regarding the r value of a mark (1)

☐ • Source Evaluation: Using the SLR to inform my opinion (same source/ different source) (2)

☐ • Non-Usage: I would not use these functions in my work. (3)

☐ • Dislike: I do not like using these functions for personal or professional reasons. (4)

☐ • Lack of Understanding: I do not fully understand these functions or how they would be applied in my work. (5)

☐ • [Additional Option if needed]: [Describe the additional context or application where the tool might be used] (6)

_____

End of Block: Consensus tools

Start of Block: Consensus Behaviours

CB The next set of questions are about the behaviours from you and your group members during the computer-assisted consensus approach.

End of Block: Consensus Behaviours

Start of Block: Consensus Behaviours

Q27 How often did you change your opinion on your [features annotated/suitability decision/my features paired/my source level opinion] because of pressure from group members?

- Never (1) - I never changed my opinion due to group pressure. (1)
- Rarely (2) - I seldom changed my opinion because of group pressure. (2)
- Sometimes (3) - Occasionally, group pressure influenced my opinion change. (3)
- Often (4) - Frequently, my opinion was swayed by group pressure. (6)
- Always (5) - I always changed my opinion when there was pressure from the group. (4)

Q28 How frequently did everyone contribute their opinions equally during the discussion?

- Never (1) - No one ever contributed their opinions equally. (1)
- Rarely (2) - It was rare for everyone to contribute equally. (2)
- Sometimes (3) - Sometimes everyone contributed equally, but not always. (3)
- Often (4) - Often, everyone contributed their opinions equally. (4)
- Always (5) - Everyone always contributed their opinions equally. (5)

Q29 How often did the most experienced person in the group lead the discussion?

- Never (1) - The most experienced person never led the discussion. (1)
- Rarely (2) - It was rare for the most experienced person to lead the discussion. (2)
- Sometimes (3) - The most experienced person sometimes led the discussion. (3)
- Often (4) - The most experienced person often led the discussion. (4)
- Always (5) - The most experienced person always led the discussion. (5)

Q30 How often did you conceal your true opinion when it differed from that of your group members?

- Never (1) - I never concealed my true opinion when it differed from the group's. (1)
- Rarely (2) - I rarely concealed my opinion under these circumstances. (2)
- Sometimes (3) - Sometimes I concealed my opinion when it differed from the group's. (4)
- Often (4) - I often concealed my opinion in such situations. (7)
- Always (5) - I always concealed my true opinion when it differed from the group's. (6)

Q31 How often did you feel embarrassed when your opinion differed from the group's opinion?

- Never (1) - I never felt embarrassed when my opinion differed. (1)
- Rarely (2) - I rarely felt embarrassed under these circumstances. (2)
- Sometimes (3) - Sometimes I felt embarrassed when my opinion differed. (3)
- Often (4) - I often felt embarrassed when my opinion was different. (4)
- Always (5) - I always felt embarrassed whenever my opinion differed from the group's. (5)

Q32 How often did you accept an opinion from your peers even though you silently disagreed?

- ○ Never (1) - I never accepted a peer's opinion when I disagreed. (1)
- ○ Rarely (2) - I rarely accepted opinions when I silently disagreed. (2)
- ○ Sometimes (3) - Sometimes I accepted opinions despite my silent disagreement. (3)
- ○ Often (4) - I often accepted peer opinions even though I disagreed. (4)
- ○ Always (5) - I always accepted the opinions of peers, regardless of my personal disagreement. (5)

End of Block: Consensus Behaviours

Start of Block: psy safety

PS The next set of questions are about the interaction with your group members during the computer-assisted consensus approach.

End of Block: psy safety

Start of Block: psy safety

Q33 It was easy to discuss complex marks and differing opinions with my peers.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q34 If I admitted to an error or mistake, I did not face retaliation or criticism from my peers.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q35 It was easy to ask members of this team for help.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q36 I felt safe offering new ideas, even if they were not fully formed plans.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q37 In this team, people were accepted for being different.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q38 My peers welcomed my ideas and gave them time and attention.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q39 My peers valued the contributions of others′.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q40 People talked about mistakes and ways to improve and learn from them.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Page Break

Q41 We took time to find new ways to improve our team´s work processes.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q42 My peers were encouraged to raise concerns they had about team plans or decisions.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q43 We tried to discover our underlying assumptions and seek counterarguments about issues under discussion.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (4)
- ○ Agree (5)
- ○ Strongly agree (6)

End of Block: psy safety

Start of Block: Stress during consensus

SdC The following questions are about your perceived stress during the computer-assisted consensus approach. Please note that we are specifically interested in the stress you experienced during group deliberations, not your general work-related stress.

End of Block: Stress during consensus

Start of Block: Stress during consensus

Q44 My work was stressful.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q45 Thinking about my work made me feel tense.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q46 At work I felt under pressure.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q47 My work had negative effects on my health.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q48 I felt stress when I reached a different opinion to my peers.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q49 I changed my opinion to reduce my stress.

- ○ Strongly disagree (1)
- ○ Disagree (2)
- ○ Neither agree nor disagree (3)
- ○ Agree (4)
- ○ Strongly agree (5)

Q50 I felt stress at the thought of making an error.

- Strongly disagree (1)
- Disagree (2)
- Neither agree nor disagree (3)
- Agree (4)
- Strongly agree (5)

End of Block: Stress during consensus

Start of Block: Feedback

Q51 This section is your space to share your thoughts openly. Whether you enjoyed certain aspects of the consensus approach or had reservations, we want to hear it all. Your insights are incredibly valuable to us. If you don't have any specific feedback, feel free to skip this question.

_____

End of Block: Feedback

Start of Block: Thank you

Thx Dear Participants, I want to express my sincere appreciation for your active involvement in our survey. This study is landmark and attitudes and behaviours between forensic examiners during a consensus discussions remain unexplored. Your responses are invaluable

and have the potential to influence the future of consensus approaches between examiners in forensic operational environments.

Kind regards,
Bernd Göttker

End of Block: Thank you

Appendix C

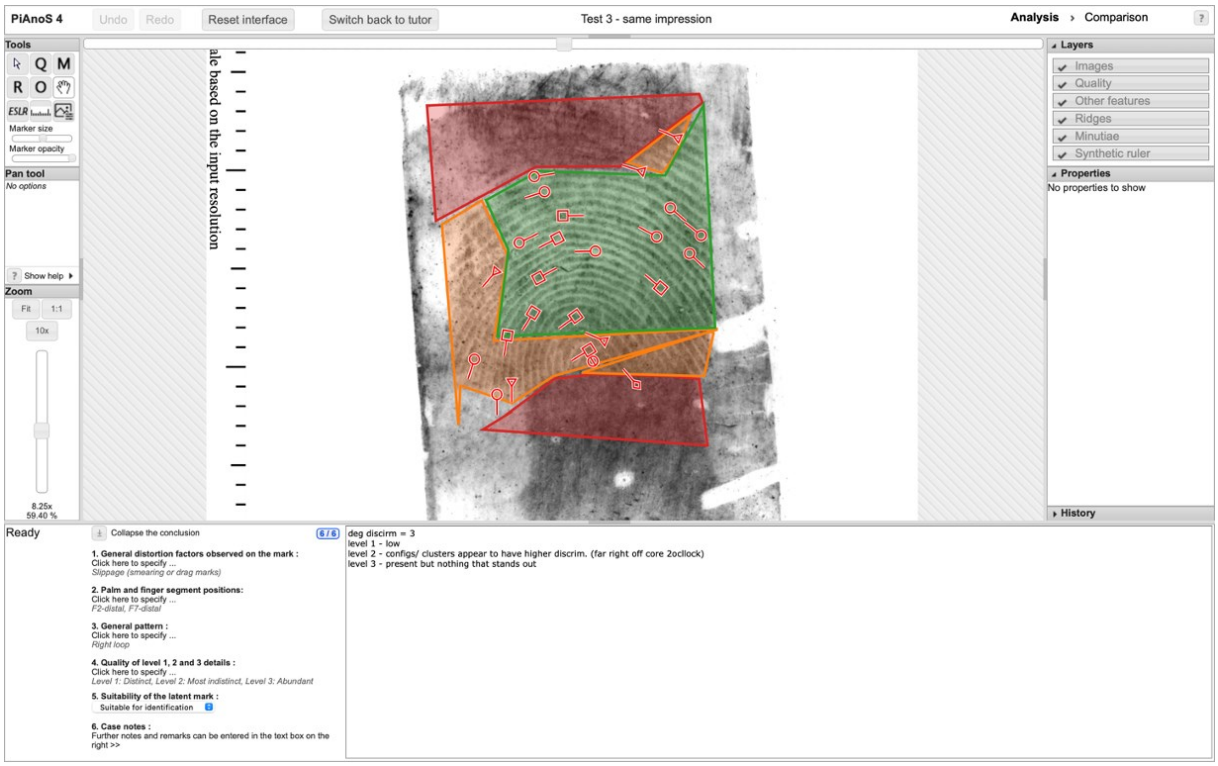
Guided checklist

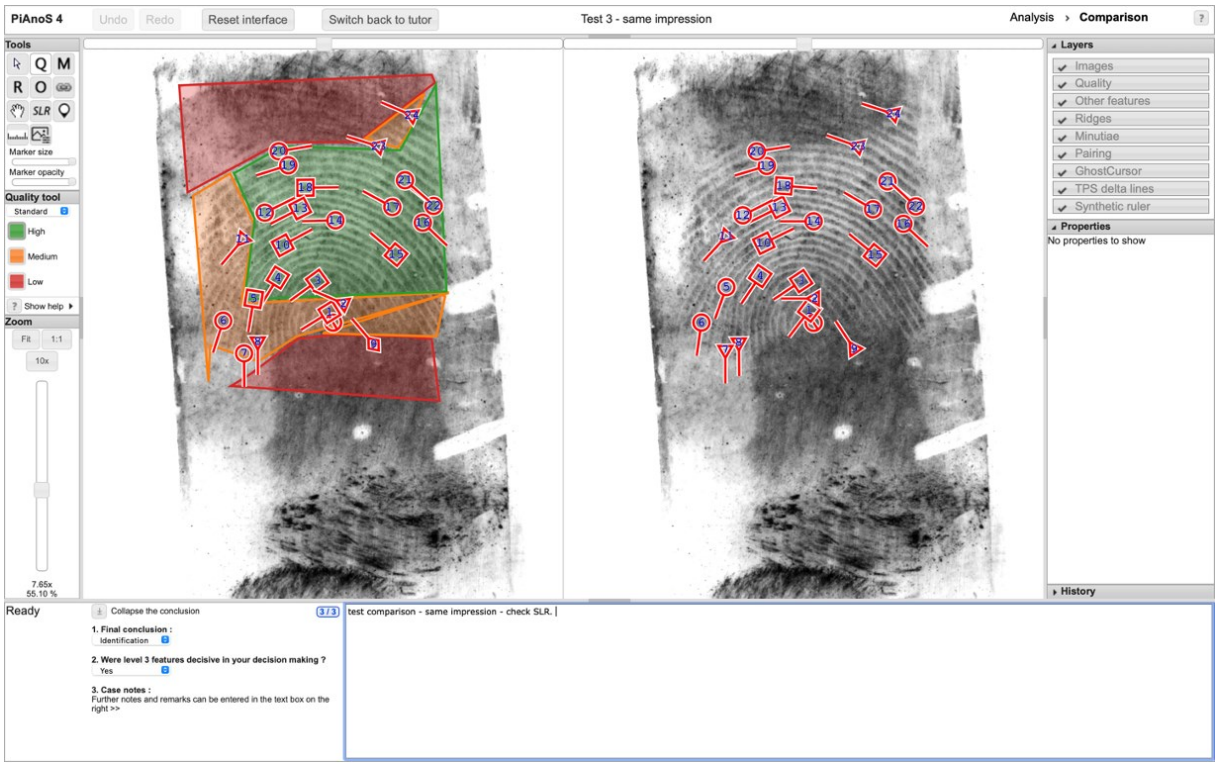
COMPUTER-ASSISTED CONSENSUS CHECKLIST – TO COMPLETE EACH EXERCISE			
PLEASE RECORD YOUR FOCUSED CONSENSUS SESSION, EXERCISE AND NAMES HERE: (e.g., Session IV, ACE_03, CA_V, AND YOUR NAMES)			
TIME START:			
Prior to beginning your consensus ACE exercise:			
Please tick the appropriate boxes	**Yes**	**No**	**If no, why?**
We are aware of our roles during this exercise (see consensus group matrix) **(Facilitator, Timekeeper, Driver)**	◻	◻	
IMPORTANT: PiAnoS TUTOR MODE ANALYSIS			
Using a separate login and webpage, we have successfully logged in to **TUTOR MODE** using the additional provided login (i.e., Ctutor01; Ctutor02; Ctutor03)	◻	◻	
TUTOR MODE ANALYSIS (see instructional slides) We have…			
1. Selected 'Analysis	◻	◻	
2. Selected the 3 examiners for this focused session and unselected the rest	◻	◻	
3. Explored our variability both individually and simultaneously in quality assessment, feature-markup, and core placement to assist us in annotating the most reliable features and core placement during this consensus	◻	◻	
ANALYSIS: During our consensus analysis we have completed/ used the following:			
1. Quality Tool	◻	◻	
2. Minutiae Tool (L2D)	◻	◻	
3. Estimated and annotated the core on the mark	◻	◻	
4. Answered all 6 questions 'expand the conclusion'	◻	◻	
5. Clicked on the ESLR function and clicked **'SEND' and 'GET'**	◻	◻	
6. **Download the PDF and view** – does the graph inform your suitability judgment? (Evidential score to right in red curve? Or left in blue curve)	◻	◻	
IN CASE NOTES			
7. Recorded our **degree of discriminability [1 to 5]**	◻	◻	
8. Recorded the outcome of your consensus	◻	◻	
(i.e., Full consensus, consensus minus one, no agreement reached and why)			
Record here if you forgot:			
9. Recorded your consent and dissent options	◻	◻	
Record here if you forgot:			
10. Recorded any major issues or disagreements	◻	◻	
Record here if you forgot:			
11. SAVED AND SUBMIT	◻	◻	
YOU HAVE NOW COMPLETED YOUR ANALYSIS			

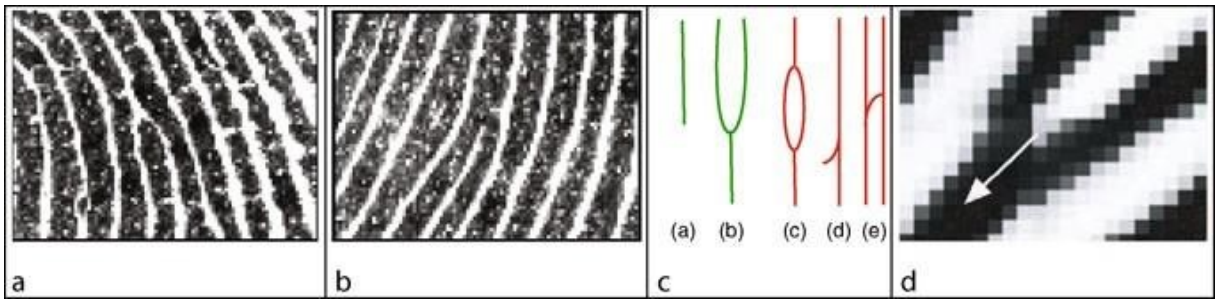
IMPORTANT: PiAnoS TUTOR MODE COMPARISON			
Please tick the appropriate boxes	**Yes**	**No**	**If no, why?**
TUTOR MODE COMPARISON (see instructional slides) We have…			
Using a separate login and web page, we have successfully logged in to TUTOR MODE using the additional provided login (i.e., Ctutor01; Ctutor02; Ctutor03)	◻	◻	
1. Selected 'Comparison'	◻	◻	
2. Selected the 3 examiners for this focused session and unselected the rest	◻	◻	
3. Explored our variability in minutiae (paired/difference), core placement to assist us in annotating the most reliable features and core placement during this consensus	◻	◻	
COMPARISON: During our consensus comparison we have completed/ used the following tools:			
1. Estimated and annotated the core on the print	◻	◻	
2. Minutiae tool (L2D)	◻	◻	
3. Pairing Tool (**minimum 4 minutiae** must be paired even for an exclusion)	◻	◻	
IN CASE NOTES			
Recorded our **degree of similarity** [1 to 5]	◻	◻	
Record here if you forgot:			
Recorded the outcome of your consensus (e.g., pairing)			
(i.e., Full consensus, consensus minus one, no agreement reached and why)	◻	◻	
Record here if you forgot:			
Recorded your consent and dissent options	◻	◻	
Record here if you forgot:			
EVALUATION: During our consensus evaluation we have completed/ used the following tools:			
1. Clicked on the SLR function AND clicked 'SEND' and 'GET"	◻	◻	
2. Download the PDF and view – does the graph inform your source OPINION? (Evidential score to right in red curve = support for ID. Left in blue curve = support for exclusion)	◻	◻	
3. Answered all questions in 'expand the conclusion'	◻	◻	
IN CASE NOTES			
4.Recorded our **degree of support** [1 to 5]	◻	◻	
Record here if you forgot:			
5. Recorded the outcome of your consensus opinion			
(i.e., Full consensus, consensus minus one, no agreement reached and why)	◻	◻	
Record here if you forgot:			
6.Recorded your consent and dissent options	◻	◻	
Record here if you forgot:			
TIME END:			

Appendix D

Example fingerprints and PiAnoS interface







. (**a**) a ridge termination engulfed in a valley bifurcation; (**b**) vice-versa. (**c**) basic ridge types in green (termination, bifurcation) and derived types in red (lake, spur, crossover) (**d**) the direction of a minutia exemplified at a ridge-bifurcation. (Bigun, 2009)

(**a**) Arch, (**b**) Tented Arch, (**c**) Left Loop, (**d**) Right Loop, (**e**) Whorl, (**f**) Twin Loop. (Bigun, 2009)

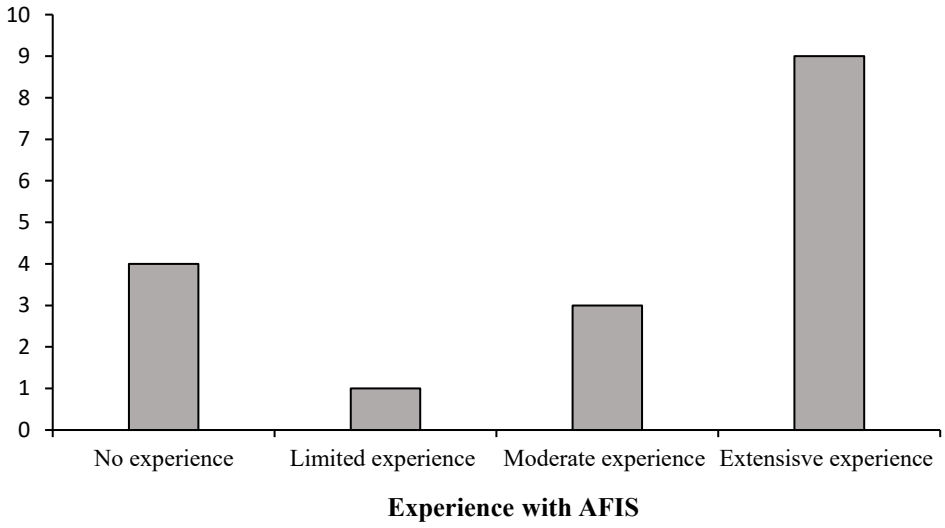
Appendix E

Tables

3.2 Fingerprint examination work details

Figure 1

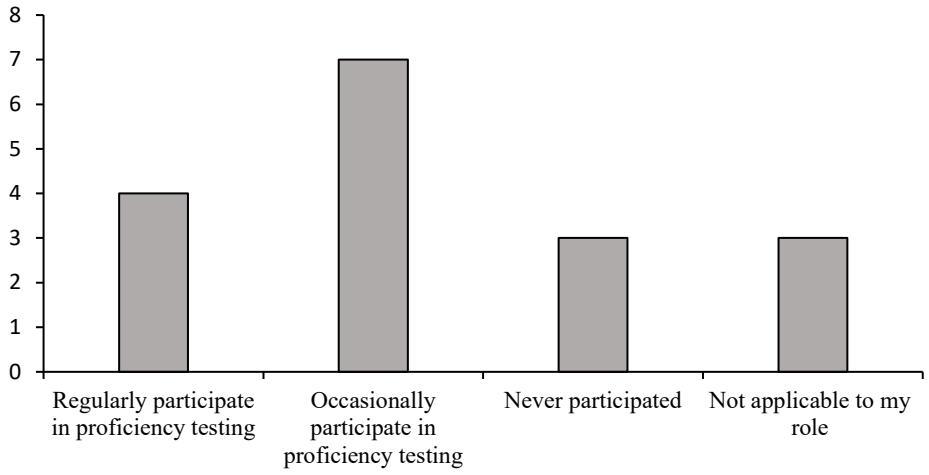
Experience with AFIS



The majority of participants indicated an “Extensive experience” (n=9) with AFIS, next four participants have “No experience” with AFIS and three participants have a moderate experience.

Figure 2

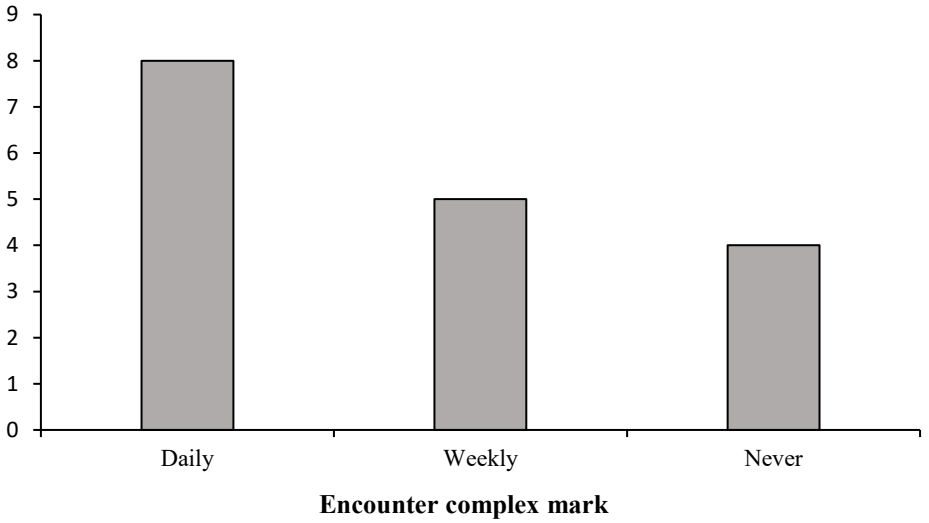
Participation in proficiency testing



This figure shows how frequently the participants engage in proficiency testing. The majority is participating "Occasionally" (n=7) in proficiency testing followed by "Regularly" participation (n=4). The rest either "Never participated" (n=3) or proficiency testing is "Not applicable to their role" (n=3) in fingerprint analysis.

Figure 3

Encounter a complex mark



This figure shows how frequently the participants encounter a complex mark in their work setting as a fingerprint analyst. Most participants encounter complex mark "Daily" (n=8), followed by participants that encounter them "Weekly" (n=5). Lastly, four participants "Never" encounter such marks.

Appendix F

During the preparation of this work, I (Bernd Göttker) used Chat Gpt to paraphrase statements I wrote on my own and to check for grammatical errors. After using this tool/service, we thoroughly reviewed and edited the content as needed, taking full responsibility for the final outcome.