Applied Cognitive Psychology

WILEY

RESEARCH ARTICLE

Capturing Fingerprint Expertise With Protocol Analysis

Brooklyn J. Corbett | Jason M. Tangen

School of Psychology, The University of Queensland, St Lucia, Queensland, Australia

Correspondence: Brooklyn J. Corbett (b.corbett@uq.edu.au)

Received: 5 October 2023 | **Revised:** 5 October 2024 | **Accepted:** 4 November 2024

Funding: This work was supported by Australian Research Council, LP170100086.

Keywords: cognition | cognitive processes | decision-making | expertise | fingerprint examination | verbal protocol analysis

ABSTRACT

In this study, we used think-aloud protocols to compare how 44 fingerprint examiners and 44 novices analysed prints. Through this qualitative approach, we discovered several notable differences. Experts focused on local ridge characteristics and minute details of each print, while novices concentrated more on global features. Additionally, experts demonstrated adaptability in dynamically planning their systematic approach and integrated specialised knowledge of causal factors affecting prints. In contrast, the novices relied more on general intuition. Furthermore, experts consistently displayed critical thinking and metacognition, carefully weighing the reliability of each identifying feature before making conclusions. However, there was variation in the precise evaluation approaches and conclusion thresholds among experts. Overall, these findings reveal the substantial complexity, adaptability and domain knowledge enhancing expert performance in fingerprint analysis. We discuss implications including balancing training of intuitive and analytical reasoning, implementing more detailed documentation, incorporating falsification practices and driving statistical advancements to strengthen evidence evaluation.

1 | Introduction

Becoming an expert is not just about knowing more; it is about thinking differently. This qualitative shift in cognition redefines the way experts engage with their domain. For example, novice chess players see individual pieces and moves, while grandmasters perceive interconnected patterns enabling longer-term planning (Gobet and Simon 1996). Standard cognitive methods like think-aloud protocols, latency measures, error and eye-tracking analyses can be used to trace the mental processes that underpin an expert's advantage (e.g., Ericsson and Simon 1993; Gegenfurtner, Lehtinen, and Säljö 2011; Schriver et al. 2008). Here, we use a think-aloud method to investigate the expertise of fingerprint examiners, whose job it is to determine the source of crime-scene prints.

1.1 | Fingerprint Examination

Fingerprint analysis, a key discipline in forensic science, involves human examiners comparing fingerprints with determine if they come from the same or different sources. This usually involves comparing a latent print from a crime scene—often incomplete or of low quality—with a full, rolled print from a suspect. This process is guided by the ACE-V framework, where analysts *Analyse* the crime-scene print for sufficient information, *Compare* it with the suspect's print, *Evaluate* their analysis and have another expert *Verify* the findings for accuracy (Ashbaugh 1999; Champod et al. 2016).

However, the ACE-V framework has been criticised for lacking specificity, leading to variability among examiners

(Haber and Haber 2007; Mnookin 2008; National Research Council 2009). There is no consensus on the features for comparison, nor clear criteria for determining if the evidence is sufficient to support a conclusion. To address these concerns and standardise practices, various organisations have developed guidelines. In Australia, the Australia New Zealand Policing Advisory Agency (ANZPAA) and the National Institute of Forensic Science (NIFS) provide guidelines that outline best practices for fingerprint examination, including requirements for documentation, peer review processes and quality assurance measures (e.g., ANZPAA 2017, 2019). Despite these efforts, implementation can still vary among agencies and jurisdictions within the country (Edmond et al. 2016). Similar variability exists in other countries, such as the United States, where the Scientific Working Group on Friction Ridge Analysis, Study and Technology (SWGFAST) provides standards, but adherence is not uniform across the country (SWGFAST 2012a, 2012b). This lack of uniformity is reflected in research showing that examiners' judgements about the adequacy of information and the number of identifying features required for a decision can vary significantly (Ulery et al. 2014, 2015, 2016). In the absence of quantified rubrics, examiners rely on their training and experience to subjectively assess the evidence.

This subjectivity does not necessarily compromise the process—examiners have shown genuine expertise (see Towler et al. 2018 for a review)—but it does reduce transparency, making it harder to fully understand expert decision-making. Indeed, professional fingerprint examiners perform exceedingly well compared with untrained novices, even under time constraints, with visual noise and highly similar nonmatches (Tangen, Thompson, and McCarthy 2011; Thompson and Tangen 2014; Thompson, Tangen, and McCarthy 2014; Vogelsang, Palmeri, and Busey 2017). They can detect prints left by different fingers of the same person despite no overlapping features (Searston and Tangen 2017a) and perform better in domain-specific visual search tasks compared with novices (Robson, Tangen, and Searston 2021). However, as with any judgement-based method, fingerprint analysis is not infallible and errors do occur (Cole 2005). Researchers have found many ways contextual information can influence these examiners' judgements (see Kassin, Dror, and Kukucka 2013 for a review). Given the significant legal implications and the potential for error in forensic examination, it is important to better understand the decision-making process of examiners to develop stronger, more reliable and innovative protocols for forensic identification.

1.2 | The Nature of Expertise

Expertise is typically characterised by the ability to perform a task quickly and accurately (Kahneman and Klein 2009). In clinical diagnostic thinking, for instance, decisions are often based on the rapid and unconscious recognition of patterns seen in past situations, a process known as non-analytic processing (Norman and Brooks 1997). Extensive experience in a specific field allows experts to identify meaningful patterns and regularities (Brooks 1978, 2005; Goldstone 1998; Kellman and Garrigan 2009; Norman, Young, and Brooks 2007). Over time, they amass a large repository of domain-specific exemplars, which refines their perception and enables rapid pattern recognition and intuitive judgements (Brooks 1978, 2005; Norman, Young, and Brooks 2007). Compared with novices, fingerprint examiners can more accurately distinguish prints at a glance, indicating the use of non-analytic processing (Busey and Vanderkolk 2005; Thompson and Tangen 2014). However, their matching accuracy improves when given more time to decide, suggesting that slow, analytic processing also contributes to their expertise (Thompson and Tangen 2014). Both types of processing likely interact to drive superior expert performance.

Currently, no empirical investigations have attempted to capture the decision-making processes of experts during fingerprint analysis. Think-aloud protocols could provide a valuable methodology to uncover the balance between intuitive and analytical strategies used by examiners.

1.3 | Protocol Analysis

Understanding fingerprint examiners' expertise requires examining the cognitive processes behind their judgements. A common method for gathering information on exceptional performance involves asking experts to reflect on their skills and behaviours (Lintern et al. 2018; Van de Wiel 2017). While introspection and self-reporting methods can provide some insight into human cognition, they may not always yield accurate information (Nisbett and Wilson 1977). This is because people often have limited conscious access to their own mental processes, raising concerns about the reliability of these methods (Nisbett and Wilson 1977).

Protocol analysis provides an alternative to introspective methods, aiming to capture real-time cognitive processes with greater scientific rigour (Ericsson and Simon 1980, 1993; Fox, Ericsson, and Best 2011). Unlike introspection, which relies on retrospective accounts, protocol analysis captures concurrent verbal data through the 'think-aloud' method. This method involves asking subjects to verbalise their thoughts while performing a task, offering a more systematic way of capturing cognitive activity as it happens. While the think-aloud procedure has its limitations—it may not capture a complete record of cognitive processes—it is a nonreactive method (Fox, Ericsson, and Best 2011). This means it does not alter performance or the cognitive mechanisms that mediate task performance (Fox, Ericsson, and Best 2011). By studying the evolving thoughts of experts while they perform a task, we can gain detailed insights into the mechanisms that drive their performance.

Protocol analysis has been effective in illuminating various domains of expertise. De Groot's seminal work in 1946 established that chess experts do not instantly identify the best move; instead, they engage in elaborate planning and evaluation. Think-aloud protocols revealed that the players' memory skills—specifically, the ability to hold and manipulate chess positions in memory—were a critical aspect that contributed to their expertise (de Groot 1946/1978; Ericsson, Patel, and Kintsch 2000). The think-aloud methodology has also been applied to medical reasoning and diagnosis, another traditional area for studying expertise. Think-aloud protocols have shown that expert medical practitioners often engage in 'inductive forward reasoning,' where they make a tentative diagnosis based on pattern recognition of symptoms, while novices employ more resource-intensive approaches to reasoning (Schmidt and Rikers 2007). Similarly, think-aloud studies in the game of bridge reveal discernible differences between experts and novices. Expert bridge players, when planning a hand, perceive problems and constraints more accurately and plan more extensively than their less skilled counterparts (Charness 1989).

1.4 | Current Research

The insights gleaned from think-aloud protocols across various fields suggest that this approach may prove invaluable in understanding the cognitive processes at play in fingerprint analysis. Our study aims to apply this technique to illuminate how fingerprint examiners navigate their decision-making. While we recognise that this method may not capture a complete record of cognitive processes, it presents an opportunity to explore elements of the practice that could benefit from increased transparency and scrutiny.

This research focuses on fingerprint examiners within the Australian forensic system, where training and protocols are guided by national standards (e.g., ANZPAA 2017, 2019). While many aspects of these practices overlap with those in other jurisdictions, such as the United Kingdom or the United States, Australia's framework has its own unique guidelines (e.g., an emphasis on peer review consistency). These findings should be understood in this specific context, with caution applied when generalising to other jurisdictions.

By capturing how Australian examiners make their judgements, we aim to identify the features they prioritise, the criteria they use to assess the sufficiency of evidence, and how they navigate the complexities of challenging prints. Through this exploration, we seek to contribute insights that can inform the development of more standardised, objective and transparent methods in the field of fingerprint analysis. By better understanding the cognitive processes and expertise of examiners in one specific context, we can begin to construct a more comprehensive picture of the decision-making strategies used by fingerprint experts worldwide and work towards establishing best practices in this critical area of forensic science.

2 | Method

2.1 | Design

This experiment was preregistered before data collection (see: https://osf.io/kz475/). We used the Think Aloud technique to investigate the evolving thought processes of experts and novices as they engaged in fingerprint analysis. This method, proposed by Ericsson and Simon (1980), allows for the capture of sequential thought processes as participants verbalise their decision-making process. We adhered to the recommended instructions and procedures for protocol analysis, which included initial instruction, warm-up procedures, reminders to keep talking and directing the participant to focus on the presented task rather than introspect and describe their thought processes (Ericsson and Simon 1980, 1993).

2.2 | Participants

A total of 44 qualified practising fingerprint experts (25 females and 19 males, mean age = 43.64 years, SD = 8.41, mean experience = 14.89, SD = 7.75) from the Australian Federal Police, Victoria Police and New South Wales Police participated in this experiment. We also tested 44 novices (25 females and 19 males, mean age = 43.64 years, SD = 8.67) who were matched with the experts in terms of age, gender and education but had no formal experience with fingerprints. Both experts and novices completed this experiment, along with a suite of other tasks during the same testing session. Each novice received the identical trial sequences and completed each task in the same random order as their matched expert. Novices were recruited from The University of Queensland, The University of Adelaide, Murdoch University communities and the general Australian public and were paid $20 for their participation. To further motivate novice participants, we offered them an additional $5 if they could reach or exceed the performance of their matched expert.

2.3 | Materials

Our fingerprints were sourced from the NIST Special Database 300 (Fiumara et al. 2018). The collection included both plain and fully rolled impressions that were originally obtained in real-world policing settings, ensuring a representative variability in print quality. We also used latent ('crime scene') and rolled ('arrest') prints from the Forensic Informatics Biometric Repository (Tear, Thompson, and Tangen 2010). These high-quality prints with known ground truth were collected from undergraduate students. Rolled exemplars were captured using ink onto standard 10-print cards, fully rolling each finger from nail-edge to nail-edge. Latent prints were lifted from common crime scene surfaces (identified through examiner consultation) including gloss-painted timber, smooth metal, glass and plastic. To approximate real crime scene variation, participants made contact by actions like 'pushing on timber to open a door' or 'safely grabbing a knife blade.' Thus, interacting with objects generated realistic latent prints. In summary, the fingerprint stimuli comprised forensically relevant latent and rolled prints collected under controlled conditions from student volunteers.

2.4 | Procedure

Participants were filmed using two Zoom Q2n video recorders. Before starting the experimental tasks, they were given pre-recorded video instructions on how to 'think aloud'. They were also presented with a series of warm-up problems, which allowed them to receive feedback on their verbal reports. They then proceeded to complete six trials, during which they were instructed to verbalise their thoughts while deciding whether two fingerprint impressions belonged to the same finger or two different fingers. The trials varied in difficulty (easy, medium and difficult), with half of them containing prints that matched, and the other half containing non-matching pairs. The easy pairs consisted of two fully rolled prints, while the medium and difficult comparisons presented a crime-scene print on the left and a rolled print on the right. The difficulty levels were assigned by an experienced fingerprint examiner based on typical criteria used in forensic practice, such as the clarity of ridge patterns, the presence of distortions and the amount of detail available. This subjective categorisation was intended to provide a diverse set of fingerprints to participants, reflecting a range of real-world conditions.

2.5 | Data Analysis

Each participant's verbal reports were transcribed verbatim by an independent transcriber, ensuring accuracy and consistency in capturing the spoken content. The transcripts were then segmented and checked for relevance by one of the researchers who also served as a rater in the later stages of analysis. The segmentation involved breaking down the transcripts into simple statements, each representing a distinct phase of the participant's speech. The purpose of this segmentation was to isolate meaningful units of analysis for coding, ensuring that each segment captured a complete thought or concept relevant to the task. Portions of the text that did not reflect verbal thoughts, such as when the participant read the task instructions, as well as any other verbalisation not relevant to the task were eliminated from the segments. In addition, fillers such as 'ehh', 'umm' were removed.

We conducted a thematic analysis using an inductive coding approach. Initially, two independent raters immersed themselves in a subset of the data, our 'training set', to develop preliminary codes. This involved engagement with the data to identify recurring patterns and establish an initial coding scheme. These raters then compared and discussed their independently developed codes, reaching a consensus on a refined set of codes. This consensus-based coding scheme was subsequently applied to the remaining data, designated as our 'test set'. Both raters independently coded this test set, and consistency checks were performed to ensure agreement and refine the codes further. Whenever discrepancies arose in this phase, the raters engaged in discussions to reach a mutual agreement, ensuring a consistent application of the coding scheme across the dataset. Throughout this process, the coding was conducted blind to the participants' expertise level (experts or novices).

3 | Results

3.1 | Performance Accuracy Analysis

We started our analyses by comparing the performance of experts and novices in correctly identifying whether two fingerprints belonged to the same person or not. The average performance of each participant across the six trials is represented in Figure 1. We performed a one-sample z-test for proportions to compare the mean proportion of correct scores of experts and novices to a chance level performance (50%). Our results show that both experts ($M = 93\%$, SD = 11%) and novices ($M = 71\%$, SD = 22%) performed significantly above chance as indicated by the dotted line in Figure 1, $z = 18.39$, $p < 0.001$, $h = 1.97$, 95% CI [85.73%, 100.63%] and $z = 13.53$, $p < 0.001$, $h = 1.37$, 95% CI [57.83%, 84.59%], respectively. A two-sample z-test for differences in proportions revealed that experts' performance was significantly better than novices' on this fingerprint identification task, $z = 2.69$, $p = 0.007$, $h = 0.60$, 95% CI [5.98%, 37.95%] suggesting that experts excel compared with novices at determining whether two prints belong to the same person.

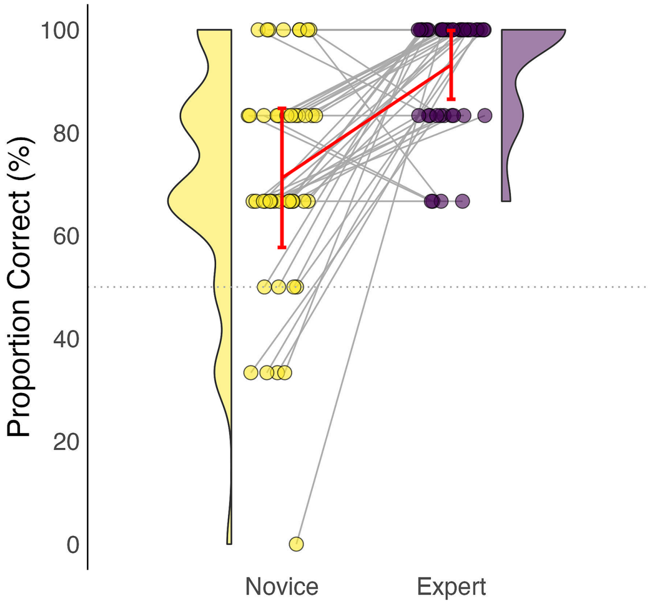

FIGURE 1 | Performance of experts and novices on fingerprint matching. Each data point represents the score of one participant, with experts in purple and novices in yellow. The half-violin plot shows the distribution of scores for each group, with a wider area indicating a higher density of scores. The dotted line indicates chance performance (50%), and the error bars represent the standard deviation. Each line connects two data points that belong to the same expert-novice pair, who completed the identical set of six trial sequences.

3.2 | Verbal Protocol Analysis

Next, we analysed the think-aloud data. We identified 12 codes covering four general dimensions within the verbal protocols: feature search and comparison, planning and knowledge application, problem representation and decision-making processes. A definition and example of each code are presented in Table 1 and the pie charts in Figure 2 represent the proportional use of different codes in the verbal protocols of experts and novices. The verbal protocols were analysed in two phases: (1) a quantitative phase, where we calculated the proportion of each code used by expert and novice participants; (2) a qualitative phase, where we closely examined the content of each code to understand the nuanced differences in the decision-making strategies

TABLE 1 | Detailed summary of codes, their definitions and example verbalisations.

Code	Definition	Example
Feature search and comparison		
Global characteristics	The subject is iteratively analysing and comparing global characteristics of the two prints (e.g., pattern type, flow of ridges, general references to the core or delta, or references to other, global, yet superficial aspects of the images including size, colour and orientation).	• 'The shape of the arch is consistent' (Novice_35) • 'Not the same, different colours, one brighter than the other, even though the centre pattern looks similar' (Novice_52)
Local characteristics	The subject is iteratively analysing and comparing local characteristics of the two prints (e.g., specific ridge details such bifurcations, dots, ridge endings, as well as the presence of scars, sweat pores and other finer level details).	• 'I can see a bifurcation coming down between, probably halfway between the delta and the core that doesn't appear in the other one, not on the same spot anyway' (Expert_47)
Planning and knowledge application		
Declarative and Procedural Knowledge	The subject states facts, rules and/or explicit procedures from memory to clarify decision-making.	• 'I can't tell the pattern type. However, I do have a delta, so I can rule out that it's an arch. Arches don't have deltas' (Expert_24)
Planning and goal setting	The subject verbalises a goal or plan to proceed with their judgement, which may be tied to a specific or non-specific aspect of the print. It encompasses both the formulation of detailed, feature-oriented plans and broader, more general strategies that are not localised or feature-directed.	• 'Now, it's possible that this may be a lot higher than where I'm looking, so I'm not getting a start down here, I'm going to look a lot higher' (Expert_01)
Problem representation		
Print quality	The subject mentions an obstacle which makes it more difficult to make an identification, including reference to the poor quality of prints such that they are distorted, smudged, blurry, partial or otherwise difficult to analyse.	• 'There's also a bit of a loop attached to the arch and it's hard to tell if that loop is also attached to the arch on the left-hand side because it's a bit smudged or something' (Novice_27)
Task limitations	The subject acknowledges a barrier related to the limited availability of imagine manipulation tools such as altering the image colour, size, orientation or marking up the prints.	• 'That's a big problem, not being able to zoom in nor make any adjustments' (Expert_55)
Uncertainty	The subject is unsure how to move forward with the judgement process, expresses confusion, uncertainty or is contradicting themselves.	• 'They're similar but I don't think that they're the—oh, I don't know' (Novice_27)
Causal factor	The subject states a causal factor; factors that play a role in producing a problem/outcome. Causal factors may involve the biology of ridge formation, or the mechanics of fingerprint deposition, which include creasing, slippage and distortion.	• 'These ridges are definitely a lot fatter than these ones, which makes you think there's added bit of pressure' (Expert_25) • 'The core seems to be on a slightly different angle ... That may or may not mean anything, because when a finger moves on a surface, it's always the core that grips' (Expert_16)

(Continues)

TABLE 1 | (Continued)

Code	Definition	Example
Decision-making processes		
Hypothesis generation	The subject makes an initial guess or makes a tentative decision typically early on in their analysis and often before they have considered all the evidence. The language used may be speculative, less decisive or confident but not final.	• 'By first impression, I'm thinking not from the same person' (Expert_04)
Validating data	The subject identifies information needed to come to a solution to the problem or they are seeking more data to verify whether information gathered is correct and complete.	• 'I'm already loving these two, but I love to be sure, so I want to find something unique' (Expert_62)
Falsification and alternative explanation	The subject is looking for evidence to disconfirm their initial hypotheses or they identify an alternative explanation to a problem.	• 'Now going to continue looking for differences to see if there is any differences before I would call an ident …' (Expert_34) • 'Sometimes, differences can occur in the core because it's—of the pliability of fingerprints. Sometimes, staples could look like curves, vice versa, so I need something more solid than just looking at staples' (Expert_16)
Critical appraisal	The subject is weighing the evidence, making a statement about the quality, reliability and strength of the evidence or their level of confidence or risk associated with the evidence.	• 'It's got quite unusual ridge formations, and here, in this sort of platform ridge area, we call it, and that's quite unusual. That gives me a lot of confidence that, yes, it's an identification' (Expert_16)

and thought processes of experts and novices. The results of both analyses are reported in turn.

3.2.1 | Quantitative Analysis

To compare the use of each code by group, we first calculated the proportion of each code used by each individual and then calculated the average across each group. We used between-groups *t*-tests, applying Bonferroni correction, to compare the average proportion of each code used by experts with the average proportion of each code used by novices. During preliminary testing of assumptions for parametric statistical analysis, several variables failed to meet the criteria of normality and/or homogeneity of variances. These variables include hypothesis generation, validating data, uncertainty, declarative and procedural knowledge, global characteristics and critical appraisal. For these variables, we employed Welch's *t*-test, which does not require the assumption of equal variances, instead of the conventional *t*-test. The numerical analyses are summarised in Table 2 and Figure 3 visually captures the code frequency of experts and novices.

3.2.1.1 | Feature Search and Comparison. The proportion of statements categorised under *local characteristics* was significantly higher in the Expert group ($M=44.07\%$, $SD=9.01\%$) compared with the Novice group ($M=21.31\%$, $SD=13.06\%$), $t(76.38)=9.52$, $p<0.001$. Conversely, for *global characteristics*, the Expert group ($M=15.67\%$, $SD=7.39\%$) had a significantly lower proportion of statements compared with the Novice group ($M=37.64\%$, $SD=14.75\%$), $t(63.29)=-8.83$, $p<0.001$.

3.2.1.2 | Planning and Knowledge Application. For the category *planning and goal setting*, there was no significant difference between the proportion of statements from the Expert group ($M=9.75\%$, $SD=5.92\%$) and the Novice group ($M=8.29\%$, $SD=7.59\%$), $t(81.21)=1.01$, $p=1.00$. Experts made significantly more *declarative and procedural knowledge* statements ($M=1.37\%$, $SD=2.34\%$) compared with novices ($M=0.16\%$, $SD=1.08\%$), $t(60.44)=3.12$, $p=0.036$.

3.2.1.3 | Problem Representation. For statements referencing *print quality*, Experts ($M=9.53\%$, $SD=4.49\%$) had a significantly lower proportion of statements compared with Novices ($M=13.78\%$, $SD=7.14\%$), $t(72.43)=-3.34$, $p=0.017$. No significant difference was observed in the *task limitations* category between the Expert ($M=1.85\%$, $SD=2.54\%$) and Novice groups ($M=1.36\%$, $SD=2.28\%$), $t(85.01)=0.95$, $p=1.00$. For the *uncertainty* category, the Expert group ($M=0.74\%$, $SD=1.99\%$) had significantly fewer statements compared with the Novice group ($M=5.45\%$, $SD=6.20\%$), $t(51.74)=-4.79$, $p<0.001$. In the *causal factor* category, Experts ($M=3.99\%$, $SD=3.67\%$) made significantly more statements than Novices ($M=1.45\%$, $SD=3.82\%$), $t(85.87)=3.18$, $p=0.027$.

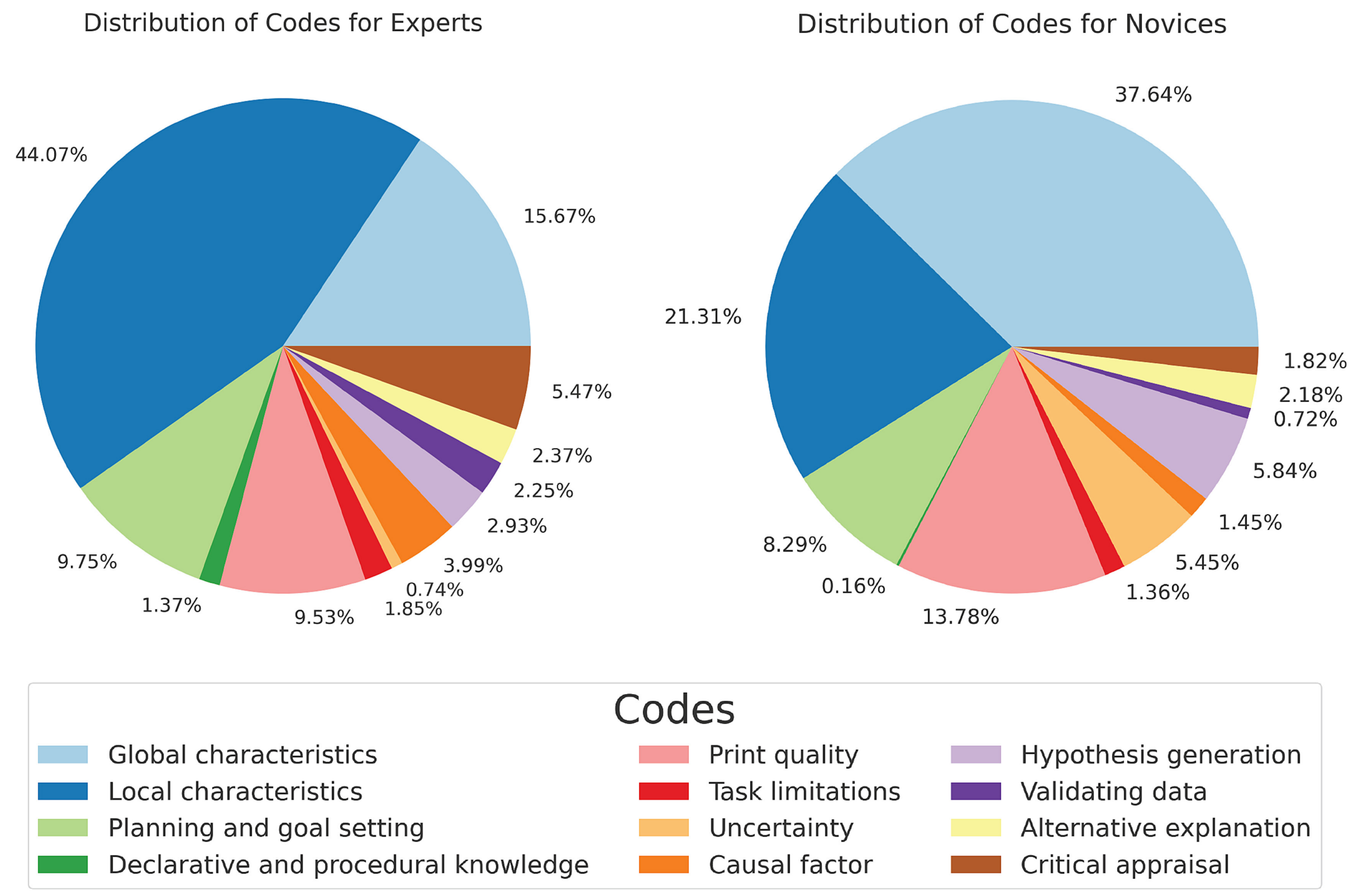

FIGURE 2 | The pie charts depict the percentage distribution of codes used by experts and novices.

TABLE 2 | Means, standard deviations and *t*-test results of code usage by Expert and Novice groups.

Measure	Novice mean (SD)%	Expert mean (SD)%	df	*t*	*p*	Cohen's *d*
Global characteristics	37.64 (14.75)	15.67 (7.39)	63.29	−8.83	< 0.001	1.91
Local characteristics	21.31 (13.06)	44.07 (9.01)	76.38	9.52	< 0.001	2.05
Planning and goal setting	8.29 (7.59)	9.75 (5.92)	81.21	1.01	1.00	0.22
Declarative and procedural knowledge	0.16 (1.08)	1.37 (2.34)	60.44	3.12	0.036	0.67
Print quality	13.78 (7.14)	9.53 (4.49)	72.43	−3.34	0.016	−0.72
Task limitations	1.36 (2.28)	1.85 (2.54)	85.01	0.95	1.000	0.21
Uncertainty	5.45 (6.20)	0.74 (1.99)	51.74	−4.79	< 0.001	−1.03
Causal factor	1.45 (3.82)	3.99 (3.67)	85.87	3.18	0.025	0.69
Hypothesis generation	5.84 (8.34)	2.93 (3.41)	56.95	−2.14	0.443	−0.46
Validating data	0.72 (1.44)	2.25 (2.57)	67.49	3.45	0.012	0.74
Alternative explanation	2.18 (3.15)	2.37 (2.30)	78.77	0.32	1.000	0.07
Critical appraisal	1.82 (2.53)	5.47 (4.31)	69.47	4.84	< 0.001	1.04

Note: The column *p* reports the *p* value corrected for multiple comparisons using the Bonferroni method.

3.2.1.4 | **Decision-making Processes.** Experts also made significantly more *validating data* statements ($M = 2.25\%$, $SD = 2.57\%$) than Novices ($M = 0.72\%$, $SD = 1.44\%$), $t(67.49) = 3.45$, $p = 0.013$. There was no significant difference between the Expert ($M = 2.37\%$, $SD = 2.30\%$) and Novice groups ($M = 2.18\%$, $SD = 3.15\%$) in the *falsification and alternative explanation*

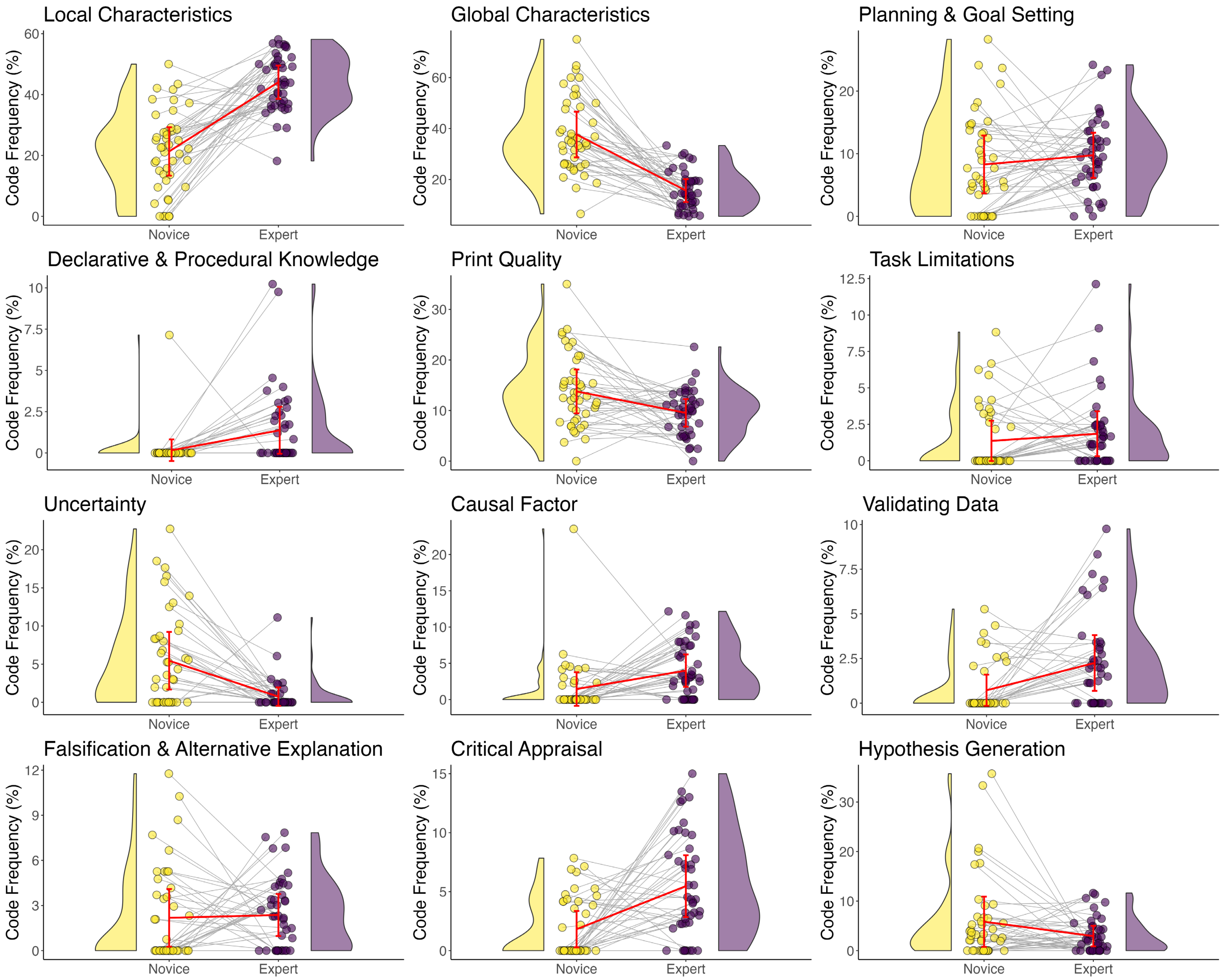

FIGURE 3 | Code frequency of experts and novices. Each data point represents the average use of each code used by one participant, with experts in purple and novices in yellow. The half-violin plot shows the distribution of scores for each group, with a wider area indicating a higher density of scores. The error bars represent the standard deviation. Each line connects two data points that belong to the same expert-novice pair, who completed the identical set of six trial sequences.

category, $t(78.77) = 0.32$, $p = 1.00$. In the *critical appraisal* category, the Expert group ($M = 5.47\%$, SD = 4.31%) had a significantly higher proportion of statements than the Novice group ($M = 1.82\%$, SD = 2.53%), $t(69.47) = 4.84$, $p < 0.001$. Last, in the category *hypothesis generation*, the Expert group ($M = 2.93\%$, SD = 3.41%) did not differ significantly from the Novice group ($M = 5.84\%$, SD = 8.34%) in the proportion of these types of statements made, $t(56.95) = -2.14$, $p = 0.480$.

3.2.2 | Qualitative Analysis

3.2.2.1 | Feature Search and Comparison. Feature search and comparison were crucial for both experts and novices in their fingerprint analysis. Most of their statements involved detailed examination and comparison of fingerprint features, categorised as either *global characteristics* or *local characteristics*. Global characteristics included pattern type (e.g., arch, loop, whorl), ridge flow, often called Level 1 features and broader aspects like size, colour, and orientation. Local characteristics focus on unique Level 2 and Level 3 features such as minutiae (ridge endings, dots, bifurcations), sweat pores, incipient ridges and the presence of scars.

Key differences emerged between experts and novices, particularly in terminology and focus. While experts used precise, domain-specific language, novices often used inaccurate or vague terms when describing fingerprint features, indicating their unfamiliarity with the terminology. Experts also took a more granular approach, more often noting local characteristics and their position relative to other features like the core or delta. On the other hand, novices largely focused on global characteristics compared to experts. The differences in emphasis on global versus local features highlight a fundamental expertise gap. Novices' focus on global characteristics, such as pattern type, reveals a reliance on features that offer limited specificity for identification. In contrast, local characteristics, such as minutiae, are more detailed and crucial for accurate identification.

Novices were also distracted by irrelevant factors like colour and size. Some novices were misled by scars or distortions, treating them as unique identifiers when they can change over time or between impressions. Experts exercised caution in such cases, as Expert_49 noted: 'I need to look at an area away from that to do the comparison because of the scarring. If it's not permanent or if the latent was before the scar, we can't always tell'. This shows the depth and specificity experts bring compared to novices' relatively broad, surface-level approach.

3.2.2.2 | Planning and Knowledge Application. Both experts and novices used *planning and goal setting* during their fingerprint analysis with comparable frequency. However, the depth, quality and adaptability of these plans differed greatly between the two groups. Experts demonstrated adaptability in their problem-solving approach, characterised by continual reassessment and adjustment of goals and tactics. For instance, Expert_12's statement 'I'm going to go in the other direction because I'm not really happy with the clarity down there' shows a willingness to modify goals based on specific challenges like clarity. Similarly, Expert_13's comment 'Now, it's possible that this [feature] may be a lot higher than where I'm looking, so I'm not getting a start down here; I'm going to look a lot higher' reveals readiness to question and revise assumptions. Additionally, their planning involves a detailed evaluation of evidence. Statements like 'I'm going to go somewhere else because I like this crease here, that's standing out to me as well to follow from there' (Expert_12) and 'Trying to find a start point, I'll be looking for things on both prints that are more discriminating ... details that might have a bit more discriminating power' (Expert_19) show how experts set specific goals related to feature quality and significance, not just searching for features but are considering their utility. They are prepared to shift their focus based on the perceived importance.

In contrast, novices exhibited more limited adaptability, often coupled with uncertainty. Unlike experts who continually re-evaluate and adjust plans based on a comprehensive understanding of evidence, novices generally react to immediate obstacles as they arise. For instance, the comment 'It doesn't feel right. Even the line—maybe if I go further out, further out, in that direction' (Novice_15) indicates novices often modify course only when directly confronted with challenges or uncertainties. Other statements like 'I'm just trying to look for something in there that might be similar' (Novice_02) reflect novices' more observational approach rather than strategic, proactive planning. This lack of proactive planning results in a more fragmented, less structured problem-solving process.

Beyond adaptability in planning and goal setting, experts also stood out through their reliance on *declarative and procedural knowledge*. Experts occasionally drew upon a well-defined set of facts, rules and procedures to guide their decision-making, planning and goal-setting. One example is the concept of 'delta danger', a term experts used to indicate risks associated with analysing deltas. Deltas on different fingers can often have very similar minutiae patterns, increasing the risk of false positives. As Expert_50 noted, 'Once again, starting off at the delta—and we always talk about delta danger, so we work out from the delta very quickly'. This reflects the expert's use of both declarative knowledge (the concept of 'delta danger') and procedural knowledge (the strategy of working out from the delta quickly) to mitigate risks.

Additionally, experts typically adhered to established protocols, beginning their investigation with the 'crime-scene' fingerprint and identifying potential features to compare with the known exemplar. Their analysis often shifted systematically from global to local features, highlighting a structured approach. Statements like, 'We're supposed to work from the not known source to a known source so that you're not biased by one or the other' (Expert_05), reveal this rule-based approach. However, some experts acknowledged exceptions, valuing flexibility when warranted. For instance, one expert remarked, 'Although I know we are supposed to go from the latent to the ink set, you can't help but notice—and it saves time—if you notice there's another one on there' (Expert_12). Another added, 'Traditionally, even though I normally would like to go from the left to the right, the right is actually clearer, so I'm taking a few more visual clues from the right than I am on the left' (Expert_15).

Novices tended to base their knowledge and procedures on personal experiences and intuitions formed during the experiment. Their statements often reflected more experiential understanding compared to experts' formalised knowledge. Their direction of analysis was inconsistent and the focus between global and local features varied. This lack of a rule-based framework resulted in a more fragmented, less structured problem-solving process. This disparity demonstrates the importance of formal training and experience in shaping the cognitive strategies for complex tasks such as fingerprint analysis.

3.2.2.3 | Problem Representation. How problems and obstacles were conceptualised and articulated emerged as key in fingerprint analysis for both experts and novices. This covered four dimensions: *task limitations*, *print quality*, *uncertainty*, and *causal factors*. Both groups acknowledged *task limitations* like the inability to manipulate images at similar rates. However, proportional differences were observed in the other three dimensions. Novices made a higher proportion of statements expressing *uncertainty*, indicating greater ambiguity in their decision-making process. Similarly, more of their comments related to *print quality*, focusing on perceptual challenges like smudging on the images.

Experts, instead, made more statements referencing *causal factors*—elements contributing to a problem or outcome. These factors encompassed biological and mechanical aspects, from ridge formation biology to fingerprint deposition mechanics, including creasing, slippage, pressure and distortion. For example, Expert_10 noted causal factors for ridge thickness variations: 'some of the ridges are a lot thicker, some of them are a lot lighter down here', attributing them to 'movement and pressure distortions'. This shows that experts go beyond identifying challenges to understand their root causes.

3.2.2.4 | Decision-Making Processes. The process of reaching a decision during fingerprint analysis involves several cognitive steps across four key dimensions: *hypothesis generation*, *validating data*, *falsification and alternative explanations*, *and critical appraisal*. Both experts and novices formed initial hypotheses or preliminary decisions early on. While experts often used qualifiers like 'By first impression' and 'My initial thoughts' to preface their initial hypotheses, novices tended to be more definitive. This difference in tentativeness could indicate experts' deeper cognitive evaluation and awareness of complexities. The novices' less tentative approach may compromise their adaptability when confronted with new evidence, thus potentially affecting the quality of their final decisions.

While novices may settle prematurely on initial judgements, experts were more often engaged in process *validation*. For example, Expert_25 emphasised the importance of performing a ridge count to confirm the spatial relationships between identified markers: 'One thing that I do make sure that I do, even though I'm already convinced that it's an identification, is that I will do a ridge count just to make sure that the spatial relationship between the markers that I've identified is actually correct'. Similarly, Expert_16 seeks further evidence to verify initial conclusions, stating, 'At this stage, I'm thinking it's not an identification, but I also like to double-check other characteristics and to double-check my conclusions'. This practice indicates a more methodical, evidence-driven approach by experts.

Both experts and novices engaged in falsification and considered alternative explanations similarly. This reflective approach involved seeking disproving evidence and exploring different possibilities before settling on a conclusion. However, for experts, these considerations were grounded in their understanding of causal factors, relying on a deep knowledge of the underlying mechanics and biology of fingerprint formation. For instance, one expert's exploration of alternative explanations was influenced by factors like distortion and movement, as shown in the statement: 'Coming down from that, I've either got a lake or just another short ridge. It will depend on any sort of distortion or movement when I examine the rolled impression' (Expert_50).

In contrast, novices' alternative explanations seemed rooted in a sense of uncertainty and low confidence rather than deep understanding. Furthermore, their statements often concerned more superficial elements of print variation, like image angle or rotation. For example, statements like, 'The general slant here, it looks slightly different here to me, but that could be that it's on a different angle' (Novice_09) reflect this surface-level focus. Unlike experts, who examine causal factors behind observed features, novices seem to fixate on less relevant attributes.

Last, *critical appraisal* separated experts and novices. This process guided their evaluations and final decisions through careful weighing of the evidence, assessment of its quality and reliability and consideration of associated confidence levels or risk. Many experts expressed needing a 'sufficient number of points', suggesting a quantitative element to evidence appraisal, though no exact numerical threshold was specified. This indicates a qualitative dimension also exists, where experts deem certain features more significant in their evaluations.

Experts used multiple methods to evaluate evidence, pointing to a qualitative nature of evidence appraisal. One expert weighted reliability based on location, noting, 'This looks like a bifurcation, but my weighting is quite low because it's the edge of a print' (Expert_62). Others considered factors like pressure movements and slippage. As one said, 'Taking into account if there's pressure movements, slippage, that's something you've got to take into account … You don't just discount it, you keep looking. That's not about making it fit; it's knowing that we can be comfortable as far as whether that's an identification or not' (Expert_60).

Some experts also integrated statistical understanding into their critical appraisal. Mentions of 'rare', 'unique' and 'discriminating', or conversely 'common' and 'typical' features or sequences factored into how much weight was assigned to that piece of evidence. As one expert elaborated, 'This sequence is very rare and unique, so I'm starting to think…that it's the person, but I need to get a sufficient amount of points to call it a positive identification' (Expert_13).

The experts' evidence appraisal shows metacognition, continuously evaluating not just the evidence but also the quality and reliability of their analytical processes. This involves adjusting confidence levels and using various evaluation methods. Novices' relative lack of critical appraisal highlights the cognitive depth and complexity expertise brings to analytical tasks like fingerprint analysis.

4 | Discussion

In domains like fingerprint analysis that heavily depend on examiner judgements, understanding experts' reasoning and decision-making is vital to pinpointing their expertise and advancing best practices. Think-aloud protocols give invaluable insights into the cognitive processes underlying complex skills and expertise (Ericsson and Simon 1980, 1993; Fox, Ericsson, and Best 2011). By eliciting examiners' thoughts as they analyse fingerprints, we aimed to gather information about the mental processes separating novices from experts. These insights can indicate training and practice improvements to enhance the validity, reliability and transparency of this complex, high-stakes field.

Our analysis of think-aloud protocols revealed clear quantitative and qualitative differences between experts' and novices' cognitive processes and strategies. Experts more frequently referenced local print characteristics, showed greater strategic planning, and occasionally cited facts and rules guiding their analysis. Novices focused more heavily on global and surface-level characteristics, had more reactive and less structured planning, and based knowledge and procedures on personal experiences and intuitions formed during the experiment. Further differences emerged in problem representation and decision-making. While both groups recognised task limitations similarly, novices more frequently referenced print quality as a source of challenge. In contrast, experts identified more nuanced biological and mechanical factors creating analysis challenges. Experts also engaged in more tentative hypothesis formation and rigorous validation processes, critically appraising evidence and displaying metacognitive awareness.

Overall, these differences highlight the substantial role of domain knowledge in shaping experts' cognitive processes, which can be attributed to the extensive training fingerprint examiners receive on minutiae and specific comparison strategies (OSAC 2020; Stanley and Horswell 2004; SWGFAST 2012a, 2012b). However, expertise development likely involves a complex interplay between rule-based knowledge and pattern recognition abilities, with the latter developing through prolonged exposure to a wide range of cases (Brooks 2005; Kahneman and Klein 2009; Norman, Young, and Brooks 2007; Norman et al. 1989).

The findings of our study highlight key themes that provide valuable insights into fingerprint expertise. In the following sections, we explore these themes in depth, focusing on how they can inform efforts to enhance the validity, reliability and transparency of fingerprint analysis through targeted improvements in training and practice. By identifying current gaps in understanding and offering evidence-based recommendations, this discussion aims to contribute to the ongoing advancement of this complex and critical domain of forensic science.

4.1 | Experience Shapes Attentional Focus

One key theme was that experience shapes attentional focus in fingerprint examination. Expert and novice analyses differed noticeably in depth and specificity. Experts focused more on localised characteristics, while novices concentrated on broader global features, sometimes mistaking temporary or superficial details as unique identifiers. Experts' capacity to disregard ambiguity and surface structure of similar prints to accurately discriminate them may be a hallmark of their expertise.

Research in perceptual expertise shows that with training and experience, there is an evolution in perceptual focus. Over time, experts typically become sensitive to features that best discriminate categories in their domain (Goldstone 1998; Kellman and Garrigan 2009). Indeed, fingerprint training highlights minutiae like lakes, trifurcations and spurs as pivotal for identification, whereas global patterns like loops offer less discrimination (OSAC 2020; Stanley and Horswell 2004; SWGFAST 2012a, 2012b). Our findings reveal a clear shift in the attentional strategies employed by experts compared with novices, with experts focusing more on local characteristics and minute details, while novices tended to concentrate on global features.

Other research suggests that attentional shifts are not just strategic; they are also deeply rooted in perception. Perceptual learning studies show diagnostic dimensions become distinctive through relevance, while irrelevant details become less distinguishable (Goldstone 1998; Haider and Frensch 1996; Honey and Hall 1989). In other words, with training, experience and feedback, experts' attention heightens for important features while reducing for less important ones, likely explaining analysts' and novices' differing attentional strategies.

One interesting study shows this difference in attentional focus. Expert fingerprint examiners more often than novices failed to recognise a gorilla embedded in a print's global features (Robson and Tangen 2023). This supports the idea that experts may have a narrower attentional window than novices. If largely attending to specific ridge details, experts may be 'blinded' to irrelevant global features like a gorilla or other surface details, constraining their attention to what is important. While our findings are consistent with the idea of a perceptual shift, further research using complementary methodologies like eye-tracking (Busey and Vanderkolk 2005) alongside verbal data is needed to further understand the nature and extent of perceptual differences between experts and novices in fingerprint analysis. Additionally, exploring how training might influence the development and maintenance of these attentional strategies could provide valuable insights into enhancing fingerprint expertise.

4.1.1 | Implications

Fingerprint training practices vary across different countries and jurisdictions; however, they often focus on formal classification and identification rules (OSAC 2020; Stanley and Horswell 2004; SWGFAST 2012a, 2012b). While this training likely refines examiners' focus and skills, the impact of formal training on expertise development in fingerprint examination is not yet fully understood. Research on the effectiveness of training in forensic domains presents a complex picture. In the field of fingerprint examination, no studies have directly tested the impact of professional fingerprint training programmes. However, some research has shown improvements in performance over time. Searston and Tangen (2017a) tracked the performance of fingerprint trainees over 12 months and observed improvements in accuracy. While this study could not isolate the specific factors contributing to improvement, Growns et al. (2022) demonstrated in a controlled setting that statistical feature training led to modest improvements in fingerprint-matching accuracy for both novices and professionals. Similarly, Searston and Tangen (2017b) showed that feedback, labelling and exposure to contrasting examples can contribute to perceptual learning in fingerprint examination, though the improvements were generally small.

Research from a related field hints at the importance of extensive on-the-job practice in the skill development of facial examiners (Towler et al. 2019, 2021). Professional training programmes do not improve facial identification accuracy (Towler et al. 2019), and while targeted training on specific diagnostic features can slightly improve performance, the gains were limited (Towler et al. 2021). These findings collectively suggest that multiple factors contribute to the development of expert performance, including formal training, on-the-job experience and potentially other elements not fully captured in short-term training studies.

While formal training is undoubtedly important, we suspect that informal on-the-job experience may play a significant role in the development of fingerprint expertise. In our study, experts averaged almost 15 years of experience, viewing hundreds of prints daily. This exposure is not just about quantity but diversity. Examiners encounter many superficially similar prints from different people, distinct prints from the same source and everything in between (Dror and Mnookin 2010; Fagert and Morris 2015; Pankanti, Prabhakar, and Jain 2002; Towler et al. 2018). This varied exposure aligns with theories suggesting that such diversity refines perceptual skills over time (Brooks 1978, 2005; Norman, Young, and Brooks 2007). The exemplar theory of categorisation proposes that categorisation becomes easier for people who have encountered many examples from various categories. Their experience allows them to categorise new items based on similarities to examples they've seen before (Brooks 1978, 2005). Importantly, people can learn to recognise patterns and variations within categories without intentional training, especially when they receive accurate and timely feedback (Hogarth 2001; Kahneman and Klein 2009).

Automation potentially downsides this exposure. Automated fingerprint identification systems, such as the National Automated Fingerprint Identification System in Australia (Australian Criminal Intelligence Commission 2018), and the Integrated Automated Fingerprint Identification System (IAFIS) in the United States (Moses et al. 2011), use algorithms to search and compare fingerprints against large databases, generating candidate lists for examiners to review. While these systems have significantly advanced the fingerprint identification process, their design for efficiency may limit the variety of prints presented to examiners, dampening the extensive exposure crucial for expertise development (Brooks 1978, 2005; Norman, Young, and Brooks 2007).

Therefore, in the face of automation, it is essential to design training programmes that emulate this rich, diverse experience. The importance of variability in training materials for enhancing learning and generalisation has been well-established across various domains (Healy et al. 2012; Raviv, Lupyan, and Green 2022; Ritchie and Burton 2017). In fingerprint examination, expanded training sets with diverse matching and nonmatching prints are likely to be beneficial. However, exposure to variability alone may not be sufficient; accurate and timely feedback is also crucial for improving learning outcomes (Hogarth 2001; Kahneman and Klein 2009; Krasne et al. 2013; White et al. 2014). While the most effective design is uncertain for fingerprint examination, a programme incorporating these principles along with formal rules would likely focus attention on pertinent characteristics more efficiently (Moxley et al. 2012). In other words, this training approach would reinforce the attentional focus that is developed through extensive on-the-job practice. However, attentional focus is just one facet; how experts employ analytical processes also critically distinguishes their performance.

4.2 | Experts Harness Analytical Processes

Analytic processes involve systematic, controlled evaluation of evidence before reaching conclusions, thought to improve decision-making by reducing biases (Kahneman 2011). In various fields, focusing on articulable features significantly improves accuracy for both experts and novices (Norman et al. 1996; Brooks and Hannah 2006). Fingerprint guidelines similarly advocate methodically marking and comparing features before drawing conclusions (Ashbaugh 1999; SWGFAST 2012a, 2012b).

In our study, experts demonstrated these broader analytical processes and feature-based strategies. They verbalised both declarative and procedural knowledge, systematically applying established rules and protocols to identify key fingerprint features. This focused approach likely aids in categorising and interpreting complex information more effectively (Brooks, LeBlanc, and Norman 2000). Yet, our study also points to the complexities within this seemingly structured approach.

While guidelines recommend initiating the analysis with the crime-scene print to minimise potential biases (PCAST 2016), experts often revisit their initial findings after scrutinising a known sample. This iterative approach was defended by some experts in our study as indispensable for refining their analyses. There might be cases when the known exemplar helps to highlight a significant feature or shows that a detail in the latent print, initially considered as 'noise' (i.e., irrelevant or misleading information), is actually meaningful 'signal' (i.e., information that contributes to the identification decision). While this reinterpretation may be justified in some cases, it also risks confirmation bias, where examiners interpret new information in a way that confirms their initial hypotheses (Kassin, Dror, and Kukucka 2013). This can lead to circular reasoning, with examiners potentially seeing what they expect to see, rather than what is objectively there.

To mitigate these risks, linear sequential unmasking has been proposed as a technique to minimise exposure to potentially biasing information (Dror et al. 2015). This technique involves separating the examination of the crime-scene print from the comparison with the suspect's print, with the latter only being introduced after the initial analysis is complete. However, strict adherence to such procedures may limit the potential insights gained from iterative analysis. So, it is a double-edged sword: prohibiting examiners from revisiting their initial findings may limit biases but could also limit valuable insights.

While rules and protocols have merits and drawbacks, our study suggests experts possess additional analytical understanding. This deeper, more nuanced grasp goes beyond procedural and declarative knowledge. It involves a robust conceptual understanding of causal factors impacting print variation, like pressure effects on ridge detail. This conceptual depth enables experts to better distinguish meaningful 'signal' from irrelevant 'noise'. For instance, experts in our study could recognise that an apparent ridge ending, which could be mistaken for a significant feature, is actually 'noise' caused by movement during the printing process. This understanding may be crucial when experts reinterpret their initial findings. Their grasp of causal factors allows a more critical evaluation of whether to reconsider a latent print detail that had been previously discounted. For example, a faint ridge line, initially dismissed as 'noise' or a smudge, might actually be a significant feature (i.e., 'signal') when considered in the context of varying pressure during print deposition.

4.2.1 | Implications

Given the balance between analytical rigour and the flexibility offered by revisiting initial findings, how can fingerprint examination processes be refined? The answer lies not just in rigidly adhering to established protocols but in enriching them with the deep understanding that experts bring to the table. While structured approaches, such as the ACE-V process (Ashbaugh 1999), provide a valuable foundation, the think-aloud protocols revealed that experts often engage in a more adaptive and iterative analysis, revisiting initial findings and adjusting their analyses based on new information. This adaptability appears to be an essential component of the expertise that is not fully captured by the structured ACE-V framework indicating a need for research aimed at developing new protocols that allow for this flexibility while safeguarding against pitfalls.

One immediate step is enhancing transparency and accountability in analysis. Examiners should disclose any adjustments to their initial examinations. Specific procedures like documentation could track iterations for transparency about challenges and biases that may influence analysis. Existing suggestions already call for such documentation (Dror and Kukucka 2021; Langenburg and Champod 2011; Quigley-McBride et al. 2022). In this context, our findings on experts' understanding of causal factors are invaluable here. If examiners reinterpret their initial findings, documentation should include specific causal factors, like pressure variations or surface conditions, that led them to reconsider an initial 'noise' classification as 'signal'. Further depth could come from documenting preliminary hypotheses about ambiguous features and grounding this in their understanding of the causal factors leading to print variation. For instance, they could document a hypothesis that a feature might be either a lake or a short ridge depending on distortions caused by pressure or movement. This additional causal documentation serves as both an internal check and external transparency, helping to make the case for the validity and reliability of revisiting initial findings in the analysis process.

While these suggestions are compelling in theory, operational constraints like heavy caseloads and time pressures must be considered. Nonetheless, our findings point to needing an approach to fingerprint analysis that balances analytical rigour with expert flexibility. The ideal system would blend a robust rule-based framework with specific bias-mitigating documentation procedures while still allowing experts to draw on their extensive experience and domain-specific knowledge to revisit and refine their initial analyses.

4.3 | Navigating Biases for Reliable Conclusions

Our study also revealed explicit ways experts and novices attempt to navigate biases, specifically through validation, falsification and alternative explanations. Early on, both form initial hypotheses. However, for experts, this is done cautiously. While early hypotheses can be useful starting points, they also risk premature closure, reaching conclusions without considering all evidence. To counter this, experts meticulously review and verify information through validation. Yet solely seeking aligning information can lead to confirmation bias.

Both experts and novices engaged in falsification and considered alternative explanations before reaching conclusions, but the depth and basis of these processes differed significantly between the two groups. Experts' alternatives were grounded in a comprehensive understanding of causal factors like distortions in fingerprint impressions. Novices based their alternative explanations on surface-level features like image angle or rotation, often stemming from uncertainty rather than a deep understanding of fingerprint analysis principles. These differences, while expected given varying levels of training and experience, highlight specific areas where expertise manifests in fingerprint analysis.

4.3.1 | Implications

Falsification and alternative explanations rigorously test conclusions rather than taking them at face value. Research supports these practices to enhance objectivity and mitigate bias in a variety of decision-making contexts (Fahsing, Rachlew, and May 2023; Lord, Lepper, and Preston 1984; Morewedge et al. 2015). Incorporating documentation of initial hypotheses and falsification attempts into reports could directly counter confirmation bias through transparency. Structured protocols combining falsification and alternatives may further reduce subjectivity risks. However, formal research is still needed on integrating documented falsification and alternatives into fingerprint analysis. Controlled studies could assess whether mandated documentation of hypotheses and falsification attempts improves accuracy and transparency. Developing evidence-based guidelines would promote reliable practices.

4.4 | Evidence-Based Evaluation

Moving beyond methodological safeguards, understanding experts' evidence evaluation in practice is crucial. Our study shows fingerprint examiners exhibit far greater critical appraisal than novices. This involves weighing evidence, assessing its quality and reliability and associated confidence or risk levels. Across the board, experts displayed behaviours consistent with a metacognitive approach, frequently verbalising reassessments of the reliability of their analytical methods. However, different evaluation approaches and conclusion thresholds were evident among experts.

While there are established legal standards and a substantial body of literature regarding the number of minutiae required for identification decisions, there remains a lack of agreement across nations and even jurisdictions regarding the required number of points of correspondence to determine a "match" (Champod 2009; Cole 2004; Ulery et al. 2013). For instance, standards have varied between 7 points in Russia to 16 points in the UK (Champod 2009). Our findings indicate that current practice in Australia often leaves the determination of sufficiency to the expertise of the individual analyst. Indeed, many experts still expressed the need to identify a sufficient number of points for a conclusion, but the exact number was not noted. This indicates that the thresholds for these decisions can vary among examiners. Some experts emphasise the quality of evidence, taking into account distortions like pressure movements and slippage during their critical appraisal. They might discount seemingly different details if they can be explained by factors such as distortion.

Others incorporate statistical reasoning, considering the rarity of certain features as they evaluate a print. For instance, prints sharing a rare feature like a 'trifurcation' are more likely to match than those with a common feature like a 'bifurcation' (Gutiérrez-Redomero et al. 2012). While the true statistical rarity of fingerprint patterns and minutiae remains unknown, evidence suggests that experts develop an intuitive statistical understanding from extensive casework. For instance, Growns et al. (2023) found that fingerprint examiners were better able to discriminate between rare and common broad fingerprint patterns (e.g., tented arch vs. left plain loop) than novices. Similarly, Mattijssen et al. (2020) observed that fingerprint examiners are better able to rank the frequency of fingerprint categories than novices. These studies indicate that experts may possess an intuitive grasp of statistical information in fingerprint analysis, which aids them in efficiently discerning and processing critical patterns.

4.4.1 | Implications

Fingerprint analysis guidelines do not provide a metric for specification of which features should be used for comparison nor does it provide criteria for judging whether the evidence is sufficient to support a conclusion (Haber and Haber 2007; Mnookin 2008; National Research Council 2009). Rather than formal methods or quantified rubrics, examiners in our study relied on their training and experience to make these judgements.

In light of this, there is an urgent need for rigorously collected data that can better inform the weighting of features during the interpretive process. Existing attempts to quantify the rarity of specific fingerprint features (e.g., Gutiérrez-Redomero et al. 2011, 2012) are fraught with limitations. For one, there is an absence of standardised terminology for identifying and categorising fingerprint features. This results in disparate methodologies across studies, making it challenging to compare or consolidate findings. Additionally, many existing data sets do not account for the positioning of specific features relative to core and deltas, leading to potentially misleading results. For instance, some areas of a fingerprint, like the core and delta, naturally produce a higher number of features.

Given these limitations, research is urgently needed to standardise minutiae terminology and develop more nuanced statistical models. Models should incorporate variables like minutiae density and location, especially relative to cores and deltas, to better capture fingerprint features' true discriminatory power. Such advancements could enable more effective cross-study comparisons and robust model development.

Achieving this could provide two key benefits. First, empirically-grounded guidelines would enable examiners to make more reliable identifications. Second, more robust statistical models would offer the judicial system an improved scientific foundation for evaluating fingerprint evidence reliability, ensuring legally sound decisions.

5 | Conclusion

In summary, our think-aloud study provides important insights into the cognitive processes underlying fingerprint expertise, highlighting key differences between experts and novices. A key caveat is that these findings are based on the verbalised thoughts of Australian fingerprint examiners. While these insights provide information on their cognitive processes, they may not fully capture the practices of examiners in other jurisdictions. Differences in training systems and regional practices could influence the nature of expertise in fingerprint analysis. Additionally, the nature of think-aloud data presents another limitation to our study. Although think-aloud methods offer a useful window into complex decision-making, they do not capture all internal processes. This approach allows us to access thought processes that might otherwise remain hidden, but it may not fully represent the unconscious or automatic aspects of expert performance.

Notwithstanding these limitations, our study yields valuable findings that contribute to our understanding of fingerprint analysis expertise. Our findings highlight key themes related to attentional focus, analytical strategies, bias navigation and evidence evaluation to help advance training, practice and research in fingerprint analysis. Key implications include balancing intuitive and analytical skills in training, enhancing the precision, accountability and transparency of documentation procedures, incorporating falsification practices and driving statistical advancements to aid in evidence weighing.

More research is required to deepen understanding and translate findings into evidence-based policies and frameworks, particularly research that examines these processes across different forensic systems and jurisdictions. The identified themes serve as a foundation for future empirical work aimed at clarifying expertise and strengthening forensic examination.

Author Contributions

Conceptualization: B.J.C. and J.M.T. Data curation: B.J.C. and J.M.T. Formal analysis: B.J.C. Funding acquisition: J.M.T. Investigation: B.J.C. Methodology: B.J.C. and J.M.T. Project administration: B.J.C. and J.M.T. Resources: B.J.C. and J.M.T. Software: B.J.C. and J.M.T. Supervision: J.M.T. Validation: B.J.C. and J.M.T. Visualisation: B.J.C. Writing – original draft: B.J.C. Writing – editing and reviewing: B.J.C. and J.M.T.

Acknowledgements

We would like to acknowledge the fingerprint examiners in Australia who contributed their time to participate in this research. Special thanks go to Amy Adams, Eden Clothier, Madeleine Graham and Chloe Smith for helping to collect novice data, and again to Amy Adams for her work as an independent rater. Additional thanks to Duncan McCarthy for support in material creation, and to Gary Edmond and Kevin Eva for their roles in obtaining funding. Finally, we appreciate the collaboration from partners across the country, who contributed ideas and support, making this research possible.

Ethics Statement

This study was cleared in accordance with the ethical review processes of The University of Queensland and within the guidelines of the National Statement on Ethical Conduct in Human Research (Approval Number: 2018001369).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data for each novice and expert participant, and the code used to produce our results and plots, are available on the Open Science

Framework, with the exception of identifiable demographic information (https://osf.io/kz475/).

References

Ashbaugh, D. R. 1999. *Quantitative-Qualitative Friction Ridge Analysis: An Introduction to Basic and Advanced Ridgeology*. Oxford: CRC Press. https://doi.org/10.1201/9781420048810.

Australian Criminal Intelligence Commission. 2018. "Biometric and Forensic Services." https://www.acic.gov.au/biometric-and-forensic-services.

Australia New Zealand Policing Advisory Agency (ANZPAA). 2017. "An Introductory Guide to Evaluative Reporting."

Australia New Zealand Policing Advisory Agency (ANZPAA). 2019. "Double Blind System Testing: A Model Framework for Forensic Science Laboratories."

Brooks, L. R. 1978. "Non-Analytic Concept Formation and Memory for Instances." In *Cognition and Categorisation*, edited by E. Rosch and B. Lloyd, 3–170. New York: Routledge.

Brooks, L. R. 2005. "The Blossoms and the Weeds." *Canadian Journal of Experimental Psychology* 59, no. 1: 62–74. https://doi.org/10.1037/h0087462.

Brooks, L. R., and S. D. Hannah. 2006. "Instantiated Features and the Use of "Rules"." *Journal of Experimental Psychology. General* 135, no. 2: 133–151. https://doi.org/10.1037/0096-3445.135.2.133.

Brooks, L. R., V. R. LeBlanc, and G. R. Norman. 2000. "On the Difficulty of Noticing Obvious Features in Patient Appearance." *Psychological Science* 11, no. 2: 112–117. https://doi.org/10.1111/1467-9280.00225.

Busey, T. A., and J. R. Vanderkolk. 2005. "Behavioral and Electrophysiological Evidence for Configural Processing in Fingerprint Experts." *Vision Research* 45, no. 4: 431–448. https://doi.org/10.1016/j.visres.2004.08.021.

Champod, C. 2009. "Presentation to the Fingerprint Enquiry Scotland." http://www.webarchive.org.uk/wayback/archive/20150428163143/htt.

Champod, C., C. Lennard, M. Stoilovic, and P. Margot. 2016. *Fingerprints and Other Ridge Skin Impressions*. Boca Raton: CRC Press.

Charness, N. 1989. "Expertise in Chess and Bridge." In *Complex Information Processing: The Impact of Herbert A. Simon*, edited by D. Klahr and K. Kotovsky, 183–208. Hillsdale NJ: Erlbaum Associates.

Cole, S. A. 2004. "Grandfathering Evidence: Fingerprint Admissibility Rulings From Jennings to Llera Plaza and Back Again." *American Criminal Law Review* 41, no. 3: 1189–1276.

Cole, S. A. 2005. "More Than Zero: Accounting for Error in Latent Fingerprint Identification." *Journal of Criminal law & Criminology* 95, no. 3: 985–1078. http://www.jstor.org/stable/3491332.

de Groot, A. D. 1946/1978. *Thought and Choice in Chess*. The Hague: Mouton.

Dror, I. E., and J. Kukucka. 2021. "Linear Sequential Unmasking–Expanded (LSU-E): A General Approach for Improving Decision Making as Well as Minimizing Noise and Bias." *Forensic Science International: Synergy* 3: 100161. https://doi.org/10.1016/j.fsisyn.2021.100161.

Dror, I. E., and J. L. Mnookin. 2010. "The Use of Technology in Human Expert Domains: Challenges and Risks Arising From the Use of Automated Fingerprint Identification Systems in Forensic Science." *Law, Probability, and Risk* 9, no. 1: 47–67. https://doi.org/10.1093/lpr/mgp031.

Dror, I. E., W. C. Thompson, C. A. Meissner, et al. 2015. "Context Management Toolbox: A Linear Sequential Unmasking (LSU) Approach for Minimizing Cognitive Bias in Forensic Decision Making." *Journal of Forensic Sciences* 60, no. 4: 1111–1112. https://doi.org/10.1111/1556-4029.12805.

Edmond, G., B. Found, K. Martire, et al. 2016. "Model Forensic Science." *Australian Journal of Forensic Sciences* 48, no. 5: 496–537. https://doi.org/10.1080/00450618.2015.1128969.

Ericsson, A., and H. A. Simon. 1993. *Protocol Analysis: Verbal Reports as Data*. Rev. ed. Cambridge: MIT Press.

Ericsson, K. A., V. Patel, and W. Kintsch. 2000. "How experts' Adaptations to Representative Task Demands Account for the Expertise Effect in Memory Recall: Comment on Vicente and Wang (1998)." *Psychological Review* 107, no. 3: 578–592. https://doi.org/10.1037/0033-295X.107.3.578.

Ericsson, K. A., and H. A. Simon. 1980. "Verbal Reports as Data." *Psychological Review* 87, no. 3: 215–251. https://doi.org/10.1037/0033-295X.87.3.215.

Fagert, M., and K. Morris. 2015. "Quantifying the Limits of Fingerprint Variability." *Forensic Science International* 254: 87–99. https://doi.org/10.1016/j.forsciint.2015.07.001.

Fahsing, I., A. Rachlew, and L. May. 2023. "Have You Considered the Opposite? A Debiasing Strategy for Judgment in Criminal Investigation." *Police Journal (Chichester)* 96, no. 1: 45–60. https://doi.org/10.1177/0032258X211038888.

Fiumara, G., P. Flanagan, J. Grantham, B. Bandini, K. Ko, and J. Libert. 2018. *NIST Special Database 300: Uncompressed Plain and Rolled Images From Fingerprint Cards* (NIST TN 1993; p. NIST TN 1993). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.TN.1993.

Fox, M. C., K. A. Ericsson, and R. Best. 2011. "Do Procedures for Verbal Reporting of Thinking Have to Be Reactive? A Meta-Analysis and Recommendations for Best Reporting Methods." *Psychological Bulletin* 137, no. 2: 316–344. https://doi.org/10.1037/a0021663.

Gegenfurtner, A., E. Lehtinen, and R. Säljö. 2011. "Expertise Differences in the Comprehension of Visualizations: A Meta-Analysis of Eye-Tracking Research in Professional Domains." *Educational Psychology Review* 23, no. 4: 523–552. https://doi.org/10.1007/s10648-011-9174-7.

Gobet, F., and H. A. Simon. 1996. "Templates in Chess Memory: A Mechanism for Recalling Several Boards." *Cognitive Psychology* 31, no. 1: 1–40. https://doi.org/10.1006/cogp.1996.0011.

Goldstone, R. L. 1998. "Perceptual Learning." *Annual Review of Psychology* 49: 585–612. https://doi.org/10.1146/annurev.psych.49.1.585.

Growns, B., A. Towler, J. D. Dunn, J. M. Salerno, N. J. Schweitzer, and I. E. Dror. 2022. "Statistical Feature Training Improves Fingerprint-Matching Accuracy in Novices and Professional Fingerprint Examiners." *Cognitive Research: Principles and Implications* 7, no. 1: 60. https://doi.org/10.1186/s41235-022-00413-6.

Growns, B., E. J. A. T. Mattijssen, J. M. Salerno, N. J. Schweitzer, S. A. Cole, and K. A. Martire. 2023. "Finding the Perfect Match: Fingerprint Expertise Facilitates Statistical Learning and Visual Comparison Decision-Making." *Journal of Experimental Psychology: Applied* 29, no. 2: 386–397. https://doi.org/10.1037/xap0000422.

Gutiérrez-Redomero, E., C. Alonso-Rodríguez, L. E. Hernández-Hurtado, and J. L. Rodríguez-Villalba. 2011. "Distribution of the Minutiae in the Fingerprints of a Sample of the Spanish Population." *Forensic Science International* 208, no. 1: 79–90. https://doi.org/10.1016/j.forsciint.2010.11.006.

Gutiérrez-Redomero, E., N. Rivaldería, C. Alonso-Rodríguez, et al. 2012. "Are There Population Differences in Minutiae Frequencies? A Comparative Study of Two Argentinian Population Samples and One Spanish Sample." *Forensic Science International* 222, no. 1–3: 266–276. https://doi.org/10.1016/j.forsciint.2012.07.003.

Haber, L., and R. N. Haber. 2007. "Scientific Validation of Fingerprint Evidence Under Daubert." *Law, Probability, and Risk* 7, no. 2: 87–109. https://doi.org/10.1093/lpr/mgm020.

Haider, H., and P. A. Frensch. 1996. "The Role of Information Reduction in Skill Acquisition." *Cognitive Psychology* 30, no. 3: 304–337. https://doi.org/10.1006/cogp.1996.0009.

Healy, A. F., V. I. Schneider, and L. E. Bourne Jr. 2012. "Empirically Valid Principles of Training." In *Training Cognition: Optimizing Efficiency, Durability, and Generalizability*, edited by A. F. Healy and L. E. Bourne Jr., 13–39. New York, NY: Psychology Press.

Hogarth, R. M. 2001. *Educating Intuition*. Chicago, IL: University of Chicago Press.

Honey, R. C., and G. Hall. 1989. "Acquired Equivalence and Distinctiveness of Cues." *Journal of Experimental Psychology: Animal Behavior Processes* 15, no. 4: 338–346. https://doi.org/10.1037/0097-7403.15.4.338s.

Kahneman, D. 2011. *Thinking, Fast and Slow*. New York: Farrar, Straus and Giroux.

Kahneman, D., and G. Klein. 2009. "Conditions for Intuitive Expertise." *American Psychologist* 64, no. 6: 515–526. https://doi.org/10.1037/a0016755.

Kassin, S. M., I. E. Dror, and J. Kukucka. 2013. "The Forensic Confirmation Bias: Problems, Perspectives, and Proposed Solutions." *Journal of Applied Research in Memory and Cognition* 2, no. 1: 42–52. https://doi.org/10.1016/j.jarmac.2013.01.001.

Kellman, P. J., and P. Garrigan. 2009. "Perceptual Learning and Human Expertise." *Physics of Life Reviews* 6, no. 2: 53–84. https://doi.org/10.1016/j.plrev.2008.12.001.

Krasne, S., J. D. Hillman, P. J. Kellman, and T. A. Drake. 2013. "Applying Perceptual and Adaptive Learning Techniques for Teaching Introductory Histopathology." *Journal of Pathology Informatics* 4, no. 1: 34. https://doi.org/10.4103/2153-3539.123991.

Langenburg, G., and C. Champod. 2011. "The GYRO System: A Recommended Approach to More Transparent Documentation." *Journal of Forensic Identification* 61, no. 4: 373–384.

Lintern, G., B. Moon, G. Klein, and R. R. Hoffman. 2018. "Eliciting and Representing the Knowledge of Experts." In *The Cambridge Handbook of Expertise and Expert Performance*, edited by K. Ericsson, R. Hoffman, A. Kozbelt, and A. Williams, 165–191. Cambridge: Cambridge University Press. https://doi.org/10.1017/9781316480748.011.

Lord, C. G., M. R. Lepper, and E. Preston. 1984. "Considering the Opposite: A Corrective Strategy for Social Judgment." *Journal of Personality and Social Psychology* 47, no. 6: 1231–1243. https://doi.org/10.1037/0022-3514.47.6.1231.

Mattijssen, E. J. A. T., C. L. M. Witteman, C. E. H. Berger, and R. D. Stoel. 2020. "Assessing the Frequency of General Fingerprint Patterns by Fingerprint Examiners and Novices." *Forensic Science International* 313: 110347. https://doi.org/10.1016/j.forsciint.2020.110347.

Mnookin, J. L. 2008. "Of Black Boxes, Instruments, and Experts: Testing the Validity of Forensic Science." *Episteme* 5, no. 3: 343–358. https://doi.org/10.3366/e1742360008000440.

Morewedge, C. K., H. Yoon, I. Scopelliti, C. W. Symborski, J. H. Korris, and K. S. Kassam. 2015. "Debiasing Decisions." *Policy Insights From the Behavioral and Brain Sciences* 2, no. 1: 129–140. https://doi.org/10.1177/2372732215600886.

Moses, K. R., P. Higgins, M. McCabe, S. Prabhakar, and S. Swann. 2011. "Automated Fingerprint Identification System (AFIS)." In *The Fingerprint Sourcebook*, 6-1–6-33. Washington, DC: U.S. Department of Justice, National Institute of Justice.

Moxley, J. H., K. A. Ericsson, N. Charness, and R. T. Krampe. 2012. "The Role of Intuition and Deliberative Thinking in experts' Superior Tactical Decision-Making." *Cognition* 124, no. 1: 72–78. https://doi.org/10.1016/j.cognition.2012.03.005.

National Research Council. 2009. *Strengthening Forensic Science in the United States: A Path Forward*. Washington, DC: National Academies Press.

Nisbett, R. E., and T. D. Wilson. 1977. "Telling More Than We Can Know: Verbal Reports on Mental Processes." *Psychological Review* 84, no. 3: 231–259. https://doi.org/10.1037/0033-295X.84.3.231.

Norman, G. R., L. R. Brooks, R. G. Regehr, R. M. Marriott, and R. V. Shali. 1996. "Impact of Feature Interpretation on Medical Student Diagnostic Performance." *Academic Medicine* 71, no. 1: S108–S109. https://doi.org/10.1097/00001888-199601000-00059.

Norman, G., M. Young, and L. Brooks. 2007. "Non-Analytical Models of Clinical Reasoning: The Role of Experience." *Medical Education* 41, no. 12: 1140–1145. https://doi.org/10.1111/j.1365-2923.2007.02914.x.

Norman, G. R., and L. R. Brooks. 1997. "The Non-Analytical Basis of Clinical Reasoning." *Advances in Health Sciences Education: Theory and Practice* 2, no. 2: 173–184. https://doi.org/10.1023/A:1009784330364.

Norman, G. R., D. Rosenthal, L. R. Brooks, S. W. Allen, and L. J. Muzzin. 1989. "The Development of Expertise in Dermatology." *Archives of Dermatology (1960)* 125, no. 8: 1063–1068. https://doi.org/10.1001/archderm.1989.01670200039005.

Organization of Scientific Area Committees (OSAC) for Forensic Science. 2020. "Standard for Friction Ridge Examination Training Program." Version 1.0.

Pankanti, S., S. Prabhakar, and A. K. Jain. 2002. "On the Individuality of Fingerprints." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, no. 8: 1010–1025. https://doi.org/10.1109/TPAMI.2002.1023799.

President's Council of Advisors on Science and Technology. 2016. "Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods." Executive Office of the President President's Council of Advisors on Science and Technology. https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf.

Quigley-McBride, A., I. E. Dror, T. Roy, B. L. Garrett, and J. Kukucka. 2022. "A Practical Tool for Information Management in Forensic Decisions: Using Linear Sequential Unmasking-Expanded (LSU-E) in Casework." *Forensic Science International: Synergy* 4: 100216. https://doi.org/10.1016/j.fsisyn.2022.100216.

Raviv, L., G. Lupyan, and S. C. Green. 2022. "How Variability Shapes Learning and Generalization." *Trends in Cognitive Sciences* 26, no. 6: 462–483. https://doi.org/10.1016/j.tics.2022.03.007.

Ritchie, K. L., and A. M. Burton. 2017. "Learning Faces From Variability." *Quarterly Journal of Experimental Psychology* 70, no. 5: 897–905. https://doi.org/10.1080/17470218.2015.1136656.

Robson, S. G., and J. M. Tangen. 2023. "The Invisible 800-Pound Gorilla: Expertise Can Increase Inattentional Blindness." *Cognitive Research: Principles and Implications* 8, no. 1: 33. https://doi.org/10.1186/s41235-023-00486-x.

Robson, S. G., J. M. Tangen, and R. A. Searston. 2021. "The Effect of Expertise, Target Usefulness and Image Structure on Visual Search." *Cognitive Research: Principles and Implications* 6, no. 1: 16. https://doi.org/10.1186/s41235-021-00282-5.

Schmidt, H. G., and R. M. J. P. Rikers. 2007. "How Expertise Develops in Medicine: Knowledge Encapsulation and Illness Script Formation." *Medical Education* 41, no. 12: 1133–1139. https://doi.org/10.1111/j.1365-2923.2007.02915.x.

Schriver, A. T., D. G. Morrow, C. D. Wickens, and D. A. Talleur. 2008. "Expertise Differences in Attentional Strategies Related to Pilot Decision Making." *Human Factors* 50, no. 6: 864–878. https://doi.org/10.1518/001872008X374974.

Scientific Working Group on Friction Ridge Analysis, Study and Technology (SWGFAST). 2012a. "Standards for Minimum Qualifications and Training to Competency for Friction Ridge Examiner Trainees Version 2.0." https://www.nist.gov/system/files/documents/2016/10/26/swgfast_qualifications-training-competency_2.0_121124.pdf#:~:text=Prior%20to%20becoming%20a%20friction,are%20already%20trained%20to%20competency.

Scientific Working Group on Friction Ridge Analysis, Study and Technology (SWGFAST). 2012b. "Standard for the Documentation of Analysis, Comparison, Evaluation, and Verification Version 2.0." http://clpex.com/swgfast/Documents.html.

Searston, R. A., and J. M. Tangen. 2017a. "Expertise With Unfamiliar Objects Is Flexible to Changes in Task but Not Changes in Class." *PLoS One* 12, no. 6: e0178403. https://doi.org/10.1371/journal.pone.0178403.

Searston, R. A., and J. M. Tangen. 2017b. "The Emergence of Perceptual Expertise With Fingerprints Over Time." *Journal of Applied Research in Memory and Cognition* 6, no. 4: 442–451. https://doi.org/10.1016/j.jarmac.2017.08.006.

Stanley, S., and J. Horswell. 2004. "The Education and Training of Crime Scene Investigators: An Australian Perspective." In *The Practice of Crime Scene Investigation*, edited by J. Horswell, 89–98. Boca Raton: CRC Press.

Tangen, J. M., M. B. Thompson, and D. J. McCarthy. 2011. "Identifying Fingerprint Expertise." *Psychological Science* 22, no. 8: 995–997. https://doi.org/10.1177/0956797611414729.

Tear, M. J., M. B. Thompson, and J. M. Tangen. 2010. "The Importance of Ground Truth: An Open Source Biometric Repository." In *Proceedings of the 54th Annual Meeting of the Human Factors and Ergonomics Society*, vol. 54, 1464–1467. San Francisco, CA: United States of America. https://doi.org/10.1177/154193121005401923.

Thompson, M. B., and J. M. Tangen. 2014. "The Nature of Expertise in Fingerprint Matching: Experts Can Do a Lot With a Little." *PLoS One* 9, no. 12: e114759. https://doi.org/10.1371/journal.pone.0114759.

Thompson, M. B., J. M. Tangen, and D. J. McCarthy. 2014. "Human Matching Performance of Genuine Crime Scene Latent Fingerprints." *Law and Human Behavior* 38, no. 1: 84–93. https://doi.org/10.1037/lhb0000051.

Towler, A., R. I. Kemp, A. Mike Burton, et al. 2019. "Do Professional Facial Image Comparison Training Courses Work?" *PLoS One* 14, no. 2: e0211037. https://doi.org/10.1371/journal.pone.0211037.

Towler, A., M. Keshwa, B. Ton, R. I. Kemp, and D. White. 2021. "Diagnostic Feature Training Improves Face Matching Accuracy." *Journal of Experimental Psychology. Learning, Memory, and Cognition* 47, no. 8: 1288–1298. https://doi.org/10.1037/xlm0000972.

Towler, A., D. White, K. Ballantyne, R. A. Searston, K. A. Martire, and R. I. Kemp. 2018. "Are Forensic Scientists Experts?" *Journal of Applied Research in Memory and Cognition* 7, no. 2: 199–208. https://doi.org/10.1016/j.jarmac.2018.03.010.

Ulery, B. T., R. A. Hicklin, G. I. Kiebuzinski, M. A. Roberts, and J. Buscaglia. 2013. "Understanding the Sufficiency of Information for Latent Fingerprint Value Determinations." *Forensic Science International* 230, no. 1–3: 99–106. https://doi.org/10.1016/j.forsciint.2013.01.012.

Ulery, B. T., R. A. Hicklin, M. A. Roberts, and J. Buscaglia. 2014. "Measuring What Latent Fingerprint Examiners Consider Sufficient Information for Individualization Determinations." *PLoS One* 9, no. 11: e110179. https://doi.org/10.1371/journal.pone.0110179.

Ulery, B. T., R. A. Hicklin, M. A. Roberts, and J. Buscaglia. 2015. "Changes in Latent Fingerprint Examiners' Markup Between Analysis and Comparison." *Forensic Science International* 247: 54–61. https://doi.org/10.1016/j.forsciint.2014.11.021.

Ulery, B. T., R. A. Hicklin, M. A. Roberts, and J. Buscaglia. 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.

Van de Wiel, M. 2017. "Examining Expertise Using Interviews and Verbal Protocols." *Frontline Learning Research* 5, no. 3: 94–122. https://doi.org/10.14786/flr.v5i3.257.

Vogelsang, M. D., T. J. Palmeri, and T. A. Busey. 2017. "Holistic Processing of Fingerprints by Expert Forensic Examiners." *Cognitive Research: Principles and Implications* 2, no. 1: 15. https://doi.org/10.1186/s41235-017-0051-x.

White, D., R. I. Kemp, R. Jenkins, and A. M. Burton. 2014. "Feedback Training for Facial Image Comparison." *Psychonomic Bulletin & Review* 21, no. 1: 100–106. https://doi.org/10.3758/s13423-013-0475-3.